Knowledge aggregation from experts and customers: A contribution to new product innovation with artificial intelligence techniques

Pooja Mohanty

http://hdl.handle.net/10803/672428

DOCTORAL THESIS

Title	Knowledge aggregation from experts and customers: A contribution to new product innovation with artificial intelligence techniques
Presented by	Pooja Mohanty
Centre	Esade Business School
Department	Marketing
Directed by	Dr. Núria Agell Dr. Mònica Casabayó

C. Claravall, 1-3 | 08022 Barcelona | Tel. 93 602 22 00 | Fax 93 602 22 49 | info@url.edu | www.url.edu

To my beloved Mama (my mother's mother)

Acknowledgments

For those who want to make the world a better place with their effort and knowledge.
- Anonymous

First of all, I would like to thank my supervisors Dr. Núria Agell and Dr. Mònica Casabayo for their constant support, encouragement and inspiration for the PhD. Even in the busiest moments, Núria and Mònica always had time for me, and I would thank them for believing in me and my PhD project. I would like to thank Dr. Francois Collet for selecting me for MRes, and Dr. Maria Galli and Dr. Ana Maria Valenzuela for conforming the selection to ESADE. I am grateful to the wonderful teachers during the MRes, Dr. Alfons Sauquet for introducing me to the world of management science and the reality of academia in the first semester. I would like to thank Dr. Jaime Bonache for teaching me epistemology of science. I will never forget Aristotle's cave and I hope to remain in the light for the rest of my life. I am grateful to Dr. Jan Brinkmann and Dr. Dimo Ringov for encouraging me to excel in strategy and entrepreneurship, and sowing the seed for multi-disciplinary research in me. I would like to thank Dr. Vicenta Sierra for being a constant support through out the PhD, you and Pilar were always there to help, no matter how small or big the issue was.

I am thankful to have met some brilliant minds and kind academics who inspired and guided me through parts of the PhD journey. Foremost, Dr. Ken Matsuno for sharing your feedback on my earliest PhD work, Dr. Frank Germann for showing how can rigorous methodology make research interesting and giving me timely feedback when you had 12 projects to finish. I would like to thank Dr. Stefan Markovic for inviting me to Copenhagen Business School for the international visit. Your kindness, maturity and empathy made my stay in Denmark unforgettable. I am also humbled by people whom I have met on occasions and they have inspired me to be part of the academic community. It gives me immense pleasure to have made acquaintance with Dr. Dev Jennings, Dr. Myrto Chilova, Dr. Laia Pujol, Dr. H. Dennis Park, and Dr. Johny Holmström for your feedback on different projects and your encouragement to continue research in the last four years.

It was wonderful to have a fun MRes cohort, Souhayla, Lucie, Anna, Julius, Asma, Lina and Natalia. I enjoyed our running- and yoga-group's company, and talking about practically anything under the sun. It was also a great experience meeting the ever expanding PhD cohort that made lunches so interesting. I would like to thank Ben, Atieh, Sahar and Menna for some great advice and delightful conversations. My family in Sant Cugat without whom I would not have survived and enjoyed my time in Spain.

José and Carmé, you are my true friends, who not only gave a house to live but treated me as your guest. Clara, Adria, Alex and Abril became my family that I didn't know I wanted in my life. Thanks for making me feel at home. To my Spanish language group, Trasy, John, Vicky and others, with whom I learned a new language and forged friendships that I truly treasure. I am touched by the Spanish warmth, "*you are here (in Spain), so you are part of us now*" - I will never forget these words. I would like to thank Sayantani for being there when I couldn't move my arm, you cooked for me, helped me survive and otherwise added joy to my life. I am grateful to my family, especially my mother who was uncharacteristically understanding through out the PhD years and who motivated me to do my best. Last but not the least, my partner, Sandipan who was always there (mostly virtually) for me to discuss subjects ranging from politics, science, programming languages and superhero etc., but most importantly being there for me as my moral support.

Knowledge aggregation from experts & customers: A Contribution to new product innovation with Artificial Intelligence techniques

ABSTRACT

New product innovation with customer participation has helped reinforce the paradigm shift from organisation to customer centric logic. The benefits of customers' participation for cocreation in new products are specifically prolific, giving rise to interesting phenomenon of crowdsourcing with tournaments, broadcasting and contests. However, with increased participation, firms face challenges in filtering the best solutions or ideas for their problems. Acknowledging these facts, we employ a customer-adoption perspective to investigate the problem. First, we identify 72 articles that examine customer adoption of innovation in a systematic literature review of 30 years across multiple disciplines. By synthesising the existing knowledge, we propose a conceptual framework linking Innovator Group (IG) customers to new product performance. Similar to lead users, the role of IG customers is crucial in New Product Development (NPD) process, and we propose ways to identify these customers and link them to new product performance by utilize their knowledge for preference dissemination, idea generation or new product information diffusion. Despite their contribution and usage in the innovation process, identification of these early adopters has remained unsystematic. Prior research has included netnography, pyramiding and screening techniques with surveys to capture the potential IG. However, there remains a gap in the systematic way to identify these customers.

To bridge the gap, we explore specific ways to select these customers from Big-Data for mitigating the challenges of overcrowding in cocreation process. We developed an intelligent system framework by combining knowledge from experts and knowledge on customer behavioural information. By employing supervised machine learning models, we help identify and predict the early adopters from the firm's database. By doing so, we show managers that they can develop Artificial Intelligence (AI) models to utilise the Big-Data they have on customers for early adopter identification. For academia, we show with knowledge aggregation from experts and customers, AI techniques perform in identification better than the existing methods. We also advance the knowledge on the key factors that affect the early adopters the most in their new product adoption decisions. Theoretically, we contribute to the NPD and customer classification literatures with applied machine learning algorithms. We also provide insights and suggest future directions for advancing knowledge in the cocreation and crowdsourcing research.

Contents

List of figures

List of Tables

List of acronyms

AHP Analytic Hierarchy Process.

AI Artificial Intelligence.

ANN Artificial Neural Network.

B2B Business to Business.

B2C Business to customers.

CART Classification and Regression trees.

CV Cross Validation.

ENC Emergent Nature Consumers.

FMCG Fast Moving Consumer Goods.

GDM Group Decision making.

IG Innovator Group.

JPIM Journal of Product Innovation Management.

kNN k Nearest Neighbour.

KU Krishnan and Ulrich.

MCDM Multi Criteria decision making.

ML Machine Learning.

MLP Multi Layer perceptron.

NPD New Product Development.

R&D Research and Development.

ROI Return on Investment.

RPIT Relationship between product and individual traits.

SLR Systematic literature review.

SMOTE Synthetic minority oversampling technique.

TFN Triangular Fuzzy Numbers.

TOPSIS Technique for Order Preference by Similarity to Ideal Solution.

WOM Word-of-mouth.

Chapter 1

Introduction

1.1 Introduction to the topic of the PhD thesis

New product innovation is a multi-disciplinary research field that encompasses diverse perspectives from business and engineering. Since its origin, new products have been regarded as a source of revenue [162], firm value [231] and sustaining firm performance [10, 133]. After initial conception, new products were developed as strategic interventions for sustaining competitive advantage in the market [51]. The pressure to compete motivated organisations to model new product innovation as process efficiency problem, where product design and execution were benchmarked with speed to market [86, 144]. In concurrence, organisations realised that cross-functional teams and top management team are important components in new product innovation machinery [26, 96]. Not long after, learnings from new product launches in high-tech industry showed an integrated approach of process efficiency, team management, product advantages and market potential to be more effective in launching successful products than any single approach [25, 101, 175].

Since the scope of New Product Development (NPD) literature is diverse and extensive due to application of multi-disciplinary approaches, it makes the aggregated knowledge rather difficult to comprehend [7, 51, 101, 110, 128, 150, 187, 191, 247]. Nonetheless, the scattered research provides key findings to the managers, yet the explanatory power of the factors have declined over the last decade (2001-2011) and can explain around 5% of new product success factors [76]. Nonetheless, paradigm shifts from product to service logic, and firm centric to customer centric perspectives in the early 2000s have transformed new product innovation. Customer participation, cocreation, co-production and user innovation highlighted the new benefits of involving customers for value creation and solving problems for the firms.

In the recent developments in information technology and digital platforms, cocreation for new product innovation has become easier with mass access to internet. This has given firms opportunity to attract crowds for solving their problems. With the easy access, the tournaments or contests for idea generation attracts a large number of participants [113]. Further, selection of good quality idea requires expertise in understanding the particular problem (why ideas were called for) and the solution space (what constitutes a good idea) [217]. Increasingly the selection process has become cumbersome and resource expensive. The goal of contests/tournaments was to attract a large crowd so that it contains some individuals that may have the solution in their neighbourhood search [258]. Scholars in the crowdsourcing and cocreation predominantly studied ways to collect ideas with multiple medium and understanding motivational factors of the crowd. Surprisingly, researchers have not focused on identifying the right customers who can cocreate with good quality inputs [1].

Lead users, Emergent Nature Consumers (ENC), market mavens and Innovator Group (IG) customers have been studied for different innovation settings [97, 107, 203, 218, 257]. Lead users are way ahead of the population in terms of their need, and they differ much from the population [257]. ENC are those who have innate ability to reimagine products [97, 107]. Market maven like to collect available information on products/prices [218]. Innovator group are the earliest to adopt new products and have influence over the later adopters [143]. Innovator group are critical because they identify new products early, take risks to purchase, and act as opinion leaders for new products [122, 205], and not including them at pre-launch and post-launch activities create negative attitude among early adopters [49]. Among all customer groups, IG customers fall in between lead users and majority customers. Since product cocreation problems are specifically designed by firms to generate ideas or to select prototypes, and not intended mostly for new inventions, we argue that Innovator groups with their domain knowledge are the right choice for a select crowd. Since Roger's definition of the customer groups [203], an update on classification of Innovator group is needed because, with the changes in technology and environment, the group must have evolved too. With customer specific data available to most consumer goods firms, these customers can be identified as a select crowd, who has specific domain knowledge, wide variety of customer preference knowledge, cognitive frames to solve some specific problems, and can influence later adopters. This also presents an opportunity for NPD researchers to improve new product's success by capitalising innovator group's knowledge and capability to help cocreate better products and to reduce overall product failure [228].

- Accordingly, the first overarching objective of the PhD thesis is to define innovator group customers and link them to new product performance from the adoption-decision making and customer participation perspectives. To achieve this research objective, a systematic literature review study and a fuzzy group decision making study were conducted to define the innovator group with their adoption factors with weights assigned to each factor on the degree of importance.

However, the current research focus of cocreation for products and crowdsourcing have increasingly moved towards valuing democracy of ideas and individual contributions. In such an environment, rapid customer participation is increasing with the access to internet, and firms are open to collaborate in all phases in the NPD process. This has generated a conducive environment for over-crowding for idea generation and information overload for the firms in particular [282]. Accordingly, many firms have started to employ third party platforms to manage idea collection and selection, some with Artificial Intelligence (AI) techniques. However, various scholars have recognised the value creation and democratisation of NPD process with the right customers [1, 110], there still remains a gap in crowdsourcing research to manage the process with selection of these customers (or best ideas). Additionally, the use of BigData and AI techniques help customer preference knowledge and decision-making more accessible to firms [226, 239], it is surprising that cocreation literature has sparingly used these relevant technology [285].

- Accordingly, the second overarching research objective of the PhD thesis is to identify and predict future innovator group customers from their transactional, behavioural and demographic data with a combination of supervised machine learning algorithms. With Artificial Intelligence techniques, the study identifies the group, and compares predictive accuracy among the algorithms.

1.2 Structure of the thesis

This PhD thesis adopts the form of a monograph, that do not have to be published yet. A brief overview and a detailed structure of the thesis is presented in this section.

- Chapter 2 contains the overarching framework of the PhD thesis. It discusses the theoretical background in detail, elaborates on the research gaps, and presents specific research objectives that forms the base for three articles presented in the subsequent chapters.

- Chapter 3 is concerned with the first research objective of the PhD thesis. It aims to propose a conceptual framework that links new product performance to innovator group's adoption decision-making. Moreover, it intends to evaluate most of the investigated adoption factors of the customer group, and categorize them into a organised structure. By doing so, the chapter sheds light on the innovator group's characteristics that helps define/identify them. The article that composed this chapter is titled as "*A conceptual framework that links new product performance to innovator group's adoption decision-making*", which is co-authored with Prof. Núria Agell and Dr. Mònica Casabayó.

- Chapter 4 is connected to the first research objective of the PhD thesis. It aims to validate and refine the part of the conceptual framework that focuses on the innovator group's adoption factors. Moreover, by validating the importance of thee factors with knowledge from industry experts and employing a combination of fuzzy logic based group decision making techniques with the experts' opinion allowed us to capture the imprecise and tcit information of the experts. That article that composed this chapter is titled as "*A Fuzzy decision-making approach to define a framework for understanding Innovator-group customers*", which is co-authored with Prof. Núria Agell and Dr. Mònica Casabayó.

- Chapter 5 addresses the second overarching research objective of the PhD thesis. In this chapter, a real-world application of the innovator group framework is conducted. With several supervised machine learning algorithms, including ensemble methods, innovator group is identified from the labelled data. A comparative analysis of the accuracy of the techniques informs the academic and managers the efficacy of a decision support system for selecting innovator group customers for cocreation purposes. That article that composed this chapter is titled as "*Understanding and predicting Innovator group customers in consumer goods industry: An Artificial Intelligence approach*", which is co-authored with Prof. Núria Agell and Dr. Mònica Casabayó.

- Chapter 6 presents the conclusion of the PhD thesis. Additionally, it provides an integrated discussion on the theoretical contributions, managerial implications, future research directions and limitations of the articles that comprise chapter 3, 4 and 5.

At the end of the PhD thesis, a combined list of academic references for all the chapters is presented.

Chapter 2

Overarching Framework

The overall aim of the thesis is to contribute to literatures in new product innovation and cocreation with applied artificial intelligence techniques. Specifically, our focus is to create new insights on understanding Innovator group from their adoption decision making process, and to identify Innovator group customers for new product cocreation process with the application of Artificial Intelligence (AI) on customer data. This chapter briefly discusses on new product innovation, cocreation knowledge and AI applications within the framework of knowledge aggregation from customers and experts. This chapter discusses the theoretical background, identifies research gaps, presents specific research objectives and methodologies that will be presented in chapters 3, 4 and 5.

2.1 New product innovation: evolution over the decades

New product innovation has evolved since Schumpeter introduced concepts of radical and disruptive innovations in 1930s [220]. Innovation research diverged into two separate dimensions: macro level or industry and country specific, and micro level or organisation and strategy specific. Both levels focused on exploring innovation through new products within their respective boundaries. We focus on the micro or organisational level of innovation. Academic research on new product at this level was influenced by product dominant logic or producer as organiser, until mid-1990's. In the era of product dominant logic, new product innovation further bifurcated into three distinct research perspectives such as, *rational planning*, *communication web* and *problem solving* [25].

Within, rational planning perspective, scholars suggested organisational efficiency in project management and superior product development in-accordance-to market needs to be key determinants for success. The studies retrospectively analysed success and sometimes failure factors for new products, and broadly attributed successful products to efficient project execution and effective product advantage. This discipline remained a-theoretical, Explorative and prescriptive in their findings [25, 51, 52]. In this period, scholars focused more on team related processes, project management, time to market for new products then customer or supplier involvement.

Communication web perspective accentuated on communication within teams, organisation, external partners and even, market. This line of research disproportionately studied effects of communication of information during New Product Development (NPD), especially on performance of teams and on new product performance. Top management team also played a significant role in the internal and external communications, according to this perspective. In contrast, problem solving perspective anchored itself to process and team management, product effectiveness, and top management team influencing both team and product decisions. Altogether, the underlying themes, that new product innovation research in this era followed can be organised into a) *process performance* with team coordination, communication and efficiency, b) *product effectiveness* with superior products in accordance to market and c) *market environment* with growth potential [25, 175].

Within process management research, many researchers formed cohesive themes on process efficiency in R&D, speed of completion, resource management and decision making within the projects. The overall line of research also contributed to project management research, where teams were employed as enablers for the projects [139]. In particular, decision framework perspective analysed NPD process within an organisation and assimilated findings from marketing, operations and strategy viewpoints to form a cohesive layout for new product development [139]. The decisions that were made in different stages of the NPD process was broadly divided into: *concept development*, *supply-chain design*, *product design*, *prototype testing* and *product launch*, and were organised according to specific core activities of the new products. Therefore, organisations explicitly refrain themselves from dividing the decisions into function specific activities. At its core, the decision-making perspective tried to combine process, product and market factors into five phases, where concept phase requires all three domains to build core product concepts. Similarly, prototype testing and product launch phases require coordination of process, product and market knowledge for successful product launches. Scholars started to explore within the five phases of the NPD process. Product design innovation and marketing competence research supports that superior product designs affect customer more than any technical aspect of the product [177, 177]. Linking marketing and engineering decisions for product design yielded more profitable products than involving either of the department [171]. Within this perspective, suppliers in supply-chain design and customers participation in concept development and prototyping were highlighted for their valuable contribution to NPD. The final products' attractiveness has been a consolidated part of new product success and it was acknowledged by both research and practice [25, 52, 54]. Whereas product design remained within the organisations because, information stickiness and technical capability were hard to find externally at one locus [258]. Whereas, idea for conceptualisation was recognised to have multiple loci, both internally (employees) and externally (customers or suppliers) [274]. Since conceptualisation phase was recognised as critical for its cascading effect on other phases, *If the core idea fails, the final product fails*, having multiple sources of problem solving helped firms advance rapid product development.

By late 1990s, human centric strategy took a stronghold within organisations along with paradigm shifts from product to service [123, 227] and firms to stakeholders [66]. The research following the paradigm shifts in the new millennium resulted in some new streams in new product innovation such as, top management's role in NPD [54], organisational culture [176], cross-functional team collaboration [230] and customer participation [274]. Organisations recognised customer knowledge as a resource, which propelled customer centric research into divergent themes of customer co-creation [198], prosumerism [201], user-innovation [259], customer engagement [98, 215], knowledge management and crowd sourcing [194] to name a few. New products were well pronounced among all the applications that organisations create with customer's knowledge [259]. In particular, von Hippel's [242, 257] research on lead users for creating innovative products and Prahalad and Ramaswamy's co-opting customers competence [196] augmented co-creation and customer engagement research on dimensions of value creation [125, 195], open innovation [48, 264], contests [240], brand management [116], product innovation [211] and service innovation [141, 234].

Simultaneously, project and process management continued to dominate the new product research. In particular, process characteristics and product characteristics received increased attention from academia between 1999-2011, indicating a search for newer significant success factors for NPD [76]. However, the explanatory power of these factors declined over the decades. This created opportunities for academic research to explore new ways to further the NPD field [76].

The perspective of "knowledge from customers" (co-creation) and "knowledge about customers" (customer relationship management) differentiated research within the customer centric domain. By late 2000s, research showed evidence in support of impact of customer cocreation on NPD process, including new product ideas, speed-to-market, customer satisfaction, and reduction in development cost [41, 183, 259]. Furthermore, by exploring new product's advantages (disadvantages) from a customer's point of view, enabled firms to understand customer's preference knowledge. This exploration was achieved either by extensive market research [235] or by direct customer interactions [89, 257]. Although knowledge generated by individuals for innovation was recognised by both management and marketing scholars [208], further focus on typology of individuals who can innovate amplified and interest arose about their location - internal or external to organisation for e.g. employees, star scientists, lead users, regular customers and partners. Conducive innovation environment at organisations also showed to accentuate creativity but individual's creativity functions both in the presence and absence of organisational environment [208].

2.2 Customer participation in New product innovation: The current state (2010-present)

Increasingly, there is growing evidence of shift in the locus of innovation from organisations to customers [97, 168]. Customer participation research has exploded since its inception [195, 196, 254]. The result of customer's involvement at various stages of NPD contributed to organisations - both public and private - for increasing their market and social welfare [83, 102]. Later scholars established the link between value creation to customer -individual and group (crowd)- participation and progressed the customer centric paradigm further [7, 44, 75, 82, 194]. In doing so, the research in last decade has embraced the philosophy of a) *democracy of ideas*, b) *value creation from contribution of regular people* to business and public services. The collaborative paradigm has resulted in co-creation in design [213], consumption experience [13], brand community [23], prosumption [201], and co-creation in product innovation [110, 215].

Cocreation is the most widely practised and recognised form of customer participation, where customers contribute to the product/service creation process continuously. Cocreation in NPD has also accelerated in practice where collaborative product development happens actively by customers and firms [110, 183]. Crowdsourcing and Open innovation are built on similar philosophy of value creation and democratisation [259].

Customer involvement in product innovation also varies according to the nature of the involvement, for e.g. customers as information source can help in incremental innovation or *exploitation*. Customers as co-innovators, contingent upon firm's technological capabilities [55], can help in radical new products or *exploration*. Similarly, knowledge *from* customers is contingent upon the individual's expertise and interests. As certain phases in NPD requires *sticky information* which is a combination of information and expertise [258], application of customer's knowledge is also contingent upon the nature of the problem. Hence, quality of cocreation-knowledge that is required for each specific phase [139] may differ between product design and core conceptualisation. Additionally, non-participating customers perceive any firm favourably, who co-create and empower customers in the NPD process [82]. The complex balance of customer knowledge management with problems of NPD process, makes the cocreation process rather nuanced.

Since individual customer became the central focus for new product innovation, research on typology of customers has increased [97, 107, 257]. In general, customers who take active part in NPD process or who may have some ability to create or select new products, are sought by firms. They are categorized into: *Innovators*, *Lead users*, *Emergent Nature Consumers (ENC)*, and *market mavens*. However these customers are not monolithic in nature and they are not identical either demographically or characteristically, or even in terms of motivation [110].

Lead users are highly creative individuals who can create new products themselves when their need is unmet by the market offerings [257]. They are distinct because of their need to create and they can be recruited by NPD teams based on their previous inventions. However, there are some disadvantages with employing lead users for product innovation: first, they are domain specific experts and are not generalists; second, since they are ahead in terms of need, they may not share preferences with later customers who are the majority of the population [161]. Also, it is hard to identify them for various categories of products where they have potential to create but have not innovated or shared their inventions with the world.

Innovator group are customers who are the earliest to purchase new products (within first 7 days from the launch - early adopters and innovators combined) [12, 203] and may or may not have lead user's motivation to invent themselves. Nonetheless, they purchase ahead of the majority in time and have potential to understand the new products better than the majority. They also share information on products with their social network. IG customers are closer to the majority in terms of preferences and needs, than lead users are. Innovator group customers are also more generalists than lead users. However, they are also attracted to some domains for e.g. a computer hardware enthusiast may not be a food connoisseur.

Emergent Nature Consumers are those who can apply judgement and intuition to change the core concepts of product, and the majority customers may find their changes appealing and useful [107]. These customers are creative in re-purposing: they come up with new use cases for existing products. However, they may not be able to create new (product) ideas by themselves. On the other hand, *Market mavens* are customers who may not buy early or innovate or change, but they do collect market information on products that interest them. They seem to be attracted to information on price and promotions more then other product features [11], and are willing to share information with the population [218].

Employing individual customers in various NPD projects have resulted in innovative products and services for firms. In general, motivation for customers' participation varies from creative satisfaction, altruism to economic gain [81, 225]. Nonetheless, organisations have benefited from their idea generation, evaluation of prototypes, and building core concepts for new products by reducing internal organisational cost [19]. The cocreation approach has helped organisations to gather interesting ideas and knowledge which resides external to their processes and employees. Application of knowledge from customers has also turned out to be highly effective for firms where customers add value with their innovations, for e.g. mountain bicycle, kites, shoes, workout clothes, watches, toys are some of the successful consumer goods product created by customers and adopted by firms [104, 159, 160, 193, 257, 260].

Cocreation as a concept is based on collaboration of customer(s), producer(s) and supplier(s) to either create value or solve problems [258]. The external agents are required because, firms creating new products can reduce cost in ideation, launch or prototype; generate novel solutions and improve brand value. Additionally, customers can bring tacit preference knowledge into the solution space that remains a difficult task for the organisations. This information is hard to collect and particularly has some key implications for creating new products that matches with customer's needs, for launching campaigns, and for managing brands. With the advent of internet and digital platforms, firms approach customers easily by open calls, contests or tournaments to gather a group of customers or crowd [75].

2.3 New product innovation with knowledge *of* and *from* customers

Particularly for NPD research with customer cocreation, idea generation has received most attention because of the spread of information technology, customer participation has become easier. Scholars argued that individuals are not the only source of innovation, rather a collection of individuals or *crowd* can also contribute to cocreation process [109]. After 15 years of crowdsourcing research, some key challenges have emerged. The primary reason for the challenge is with the misalignment of the problem expectation (knowledge and specific needs) with the solutions or outcomes.

The underlying issue is the *quality of customer innovation* which has direct impact on product (un)feasibility and implementability by firms due to cost, technical and engineering aspects of NPD [113, 139]. Second in line is the issue of quality control, as it is not easy to gauge customer's ideas. Since, different products need different levels or areas of knowledge from customers to co-create, for e.g. radical and incremental products vary in knowledge type and expertise, identifying *the good contributors* remains a challenge [112, 113, 194]. Additionally, identification of specific groups of good contributors for cocreation that can form a crowd with right combination of capabilities and knowledge, especially conducted in an unbiased manner also remains a challenge.

Discussing the challenges sequentially: First, the feasibility of customers' ideas for innovation (to be produced and launched) is contingent upon the expectation-outcome equation, especially for complex products. For simple products like t-shirts or furniture, customers' innovation helps acceptance by non-participants for signaling openness of the firm, customer empowerment or user's own need alignment [82]. Whereas for complex products, such as software or scientific measurement equipments, where technical knowledge precedes over users' preference knowledge, only a specific group of customers can self-select to co-create [94, 224]. For complex product scenarios, ideas generated by regular customers may be more novel and radical in nature [151], but they lack in product feasibility criteria [112, 140, 194]. This gap highlights the *sticky information* principle where the locus resides outside most customer's knowledge space, and it lies in the neighborhood space of technical experts and some select domain expert customers [1, 258].

Second, the quality control of customers' ideas is a challenge. Meta analytic research has established that customers contribution in valuable ideas in conceptualisation stage, is more effective than their contribution in either product design or production stage [43]. However, with ideas or any creative process, quality control is a prerequisite to achieve good outcomes. For example, in Linux's open source community, code writing remains open to all. However, only those programmers who have adequate expertise to contribute to the complex product participate in the software development. Since, open-source community also cultivates transparency, *open-to-correct* provision helps quality control. Not all NPD projects have similar self-correcting mechanism that are also cost-effective [113].

Third, there exists products that are too complex to disentangle customer's preference, and to create novel products by learning the preference criteria. For example, food or music, though considered generic, are complex products because of their tacit customer preference knowledge. Conjoint analysis method has been used to extract latent structure for customer preferences [47, 93]. However, if the product portfolio is diverse for a manufacturer, then the process of NPD becomes more difficult in translating exact preferences into the final products [214]. Nonetheless, knowledge of customer preference is implicit in nature and contains tacit information. From the behavioral decision making perspective, customer preference is a local search for the customer but it is a global search for the firms. Since the knowledge is tacit and the search is local [1, 56], customers have this important sticky information with them [258]. Therefore, by selecting a group of customers, who are capable of innovating and accessing this knowledge, may become attractive for organisations to create new complex products (without explicitly disentangling their preferences).

Fourth, some individuals can innovate because, they have information (knowledge), experience, problem solving cognitive frames and absorptive capacity that are required for a specific set of problems [1, 56]. Not all customers have the right combination, and to find the right customers is not an easy task for the firms. Attracting the crowds by calls, idea contests or tournaments are created to solve this problem with online platforms and communities. The assumption is that by including a population, there is higher chance to locate individuals who fit the profile to solve specific problems. However, *Crowdsourcing* has its own set of challenges.

2.4 Selection of customers for cocreation in NPD

With cocreation becoming a major source of idea generation, firms have employed different ways to involve customers in the process. Lead users are approached by their previous inventions [160, 257], market mavens and ENC are selected with surveys [11, 97, 107], and regular customers are invited by public announcements of tournaments or crowdsourcing contests for NPD [65, 87]. In particular, because of the concept of crowdsourcing (anyone can participate), crowdsourcing projects often receive a disproportionally large number of ideas. Screening for good ideas among all of the ideas requires expertise, which can be expensive in terms of cognitive resources and time [1, 217]. Quality of the ideas are poor when number of ideas are greater in the ”idea-market” which leads to overcrowding. Customers are unaware of the cost-structure for the implementation and lack self-awareness about their ability to cocreate [113] . Additionally, self-selection of customers may lead to a pool of same customers participating in different contests (similar to AmazonTurk). Moreover, not all customers who have the solution in their neighbourhood space have equal resources such as internet, money for membership fee for the online platforms (e.g. Chaordix, Hyve, Wazoku, InnoCentive). This leads to a crowd with a limited socio-economic background with low diversity and heterogeneity. This is contradictory to the core principle of crowdsourcing. Moreover, crowds may act as a heterogeneous population, cocreation takes place at the customer level [1, 87, 113].

Since the locus of innovation lies with the individual customer, the search for these customers has also intensified over the years. Research shows that internal lead users' created ideas are of lower quality than the external lead users', strengthening the value that customers bring for the NPD process [221, 282]. Recognizing this, past research has looked into ways to identify the lead users, innovator group, emergent consumers and market mavens with different methods. Mass screening with surveys [160, 219], pyramiding [261], focus group discussions, Netnography of online communities [14], and crowdsourcing with online platforms [87]. Crowdsourcing aims to include a heterogeneous group to benefit from attracting those few customers who possess exact criteria that a firm needs for a specific problem solving. Moreover, not all customers have the solution in their local search space [1].

Although the methods have a common goal to filter customers who can cocreate, they also have some common weaknesses. First, they are time consuming and resource intensive in nature. Second, in all of the methods, customers self-declare (and self-select) to be creative or innovative [113] or refer someone *who they think* is an expert [14]. On one hand, self-selection affects crowdsourcing contests or online platforms with *overcrowding*. On the other hand, barriers such as participation fee or internet access may prevent potential good contributors from participating. Hence, the final sample may not serve the purpose of finding good contributors. Additionally, in B2B products, managers don't value crowdsourcing ideas especially from online platforms more than the traditional marketing research because, managing, validating and integrating information from crowds seems to be cumbersome for their NPD process [282]. Whereas, the online communities may have their own goals and grow organically to non-innovation activities. Some (product) online communities are less innovation oriented in their discussions than *communities of creation* that grow out of common interests e.g. “Harley owners group” [193].

Altogether, selection process of customers for cocreation is cumbersome, resource intensive, suffers from self-selection bias, and time- and resource-expensive. Surprisingly, few scholars have conducted research on selecting *right customer* for different cocreation projects with Big-Data, which can be less expensive and less biased for firms who have some digitalization in place [285]. This is a relevant research gap because, while it is good to know customer's motivations, experiences, knowledge domain etc, managers also need to know how to manage cocreation without information overload from crowds, in a cost effective manner and still can select the right customers who could provide good quality solutions for their problems [1, 110].

Among all expert customer groups, IG customers fall in between the lead users and the majority customers. Since product cocreation problems are specifically designed by firms to generate ideas or to select prototypes, and not intended for new inventions, Innovator groups with their domain knowledge can be the right choice for a *select crowd*, as compared to the lead users. Innovator group are early to identify new products and take risks to purchase them, but they may or may not be as motivated to create new products from the scratch. They have some overlapping characteristics with ENC and market mavens. Since Roger's definition of the specific customer groups [203], an update on classification of IG customers is needed because, with the changes in technologies and environment, the group may have evolved. Additionally, Innovator groups for different product category may differ. With customer specific data available to many consumer goods firms, these customers can be identified as *a select crowd* which will have specific domain knowledge (similar to lead users), wide variety of customer preference knowledge (common with majority), and cognitive frames to solve some specific problems as they are often self-motivated. This also presents an opportunity for NPD researchers to improve new product's success by capitalise innovator group's knowledge and capability to create better products, and to reduce overall product failure [228] by cocreation from this select crowd.

2.4.1 Research objective 1

- "*A conceptual framework on linking new product performance to Innovator group's decision-making*" (refer chapter 3). Specifically, this article aims to define the innovator group's adoption decision making factors. Since, these customers are unique because of their early adoption of new products, authors goal was to collect most of the investigated factors, from the extant literature, on their adoption process, and propose a starting framework to link with the new product performance. The contribution of the study is the classification of the innovator group with the characteristics that are factors that help innovator group adopt. In doing so, authors help link the customers to cocreation and innovation diffusion processes, that ultimately leads to the new product's performance for the firms.

- "*A Fuzzy decision-making approach to define a framework for understanding Innovator-group customers*" (refer chapter 4). In this study, we propose and refine the conceptual framework to identify Innovator Group (IG) customers from their most important adoption decision-making factors. We validate the importance of these factors with knowledge from 16 industry experts who have extensive experience in new product launch. A combination of fuzzy logic based group decision making techniques with the experts' opinion allowed us to capture the imprecise and tacit information of the experts. The final ranking reveals the most crucial factors of IG customers' adoption for new products tend to be perceptual, visual and innovation driven.

2.5 Artificial Intelligence: Tool for augmenting cocreation in NPD

In the last decade, resurgence of Artificial Intelligence (AI) techniques has overtaken all other technological trends in the industry [92]. Academic research has also explored AI's role in management [136]. Some researchers consider AI as a threat to organised society as automation will perish millions of jobs [272]. Other researchers look at AI in a benevolent manner and have explored how can AI augment managerial work [27], and some argue in favour of a balanced automation and augmentation approach in management [197]. The dichotomy about AI is reflected in the society, and debate is ongoing whether AI will become a general purpose technology in the future [27, 59].

Nonetheless, the applications of AI technology, especially machine learning, are considered better suited to handle large amount of data with higher dimensionality and complexity [4, 152] than traditional analytic methods. They can process voluminous data on real-time with higher precision to predict or forecast. The most used techniques that are already showing results are primarily supervised or unsupervised machine learning and to some extent, reinforced machine learning. Training with examples, labelled or unlabelled, is the core learning method for the machines. Supervised machine learning with labelled data, that are identified by the domain experts, yield superior insights [78].

The reasons why AI techniques are well suited for the high dimensional customer data are manifold. From technical perspective, AI techniques are anticipated to show prowess in process optimisation [121], automation of administrative tasks [136], error detection and accurate predictions [15] to name a few. Generation of valuable insights from analysis has been used by data-driven firms to build new products, services and processes. For e.g. AirBnB, Netflix, Google and Amazon have extracted customer's past behavior information from their existing customer base to recommend or suggest offerings products [184]. This is attributed to AI techniques' adaptable algorithms and their training with labelled examples. In a way, supervised machine learning is successful because of human knowledge from which the machine learns. How it learns is a different topic altogether.

From the managerial perspective, the speed of execution and diligent processing of (customer) data in real-time are core strengths of AI techniques [199, 265]. Additionally, finding intricate correlational insights on customer behaviour, frauds, future predictions and customer decision making with clever algorithms which are unmatched in accuracy, precision or speed, makes AI attractive for consumer goods sector [243]. The information acts as resources for customer relationship management and customer value creation.

In summary, AI can automate and augment managerial tasks. It still lacks in creating something from scratch but it helps artists in creating art [8]. Idea generation and new product creation falls into this category. Since AI with human guidance can learn from any complex data, considering the overload of information for managers with crowdsourcing data (ideas), AI can help this NPD stage for managing and integrating with internal processes. Changing consumer preferences adds additional challenge for the NPD managers who need aid in navigating the NPD process with a competent support system. Surprisingly, researchers have not exploited the AI techniques in finding the *right customers* for crowdsourcing and cocreation purposes. A scarce research has explored cocreation with lead users and ENC by identifying them with surveys from an online community platform [256]. Although the selected customers performed better than the regular customers for cocreation tasks, the research did not explore a less cost intensive and more data reliant method (with less self-selection bias) for selecting the *right customer*.

On one hand third party online platforms may save time and hassle for new product managers while integrating ideas from the crowd [81]. On the other hand, using external online platforms has some challenges: high volume, high variety and lack of diversity. The cost also remains high because, tournaments/contests attract high volume in ideas from a crowd which may become stagnant pool of contributors over the years and loses diversity and heterogeneity [190]. Moreover, for the selection of high volume of ideas, managers face the problem of absorptive capacity. Hence, many platforms use AI techniques for the selection process where algorithms are built with experts' knowledge, which needs updating or customisation according to each client's goals [18]. Many managers find cost and managing the online platforms for crowdsourcing to be prohibitive, and prefer traditional marketing research [221].

This is a relevant research gap in the crowdsourcing and cocreation fields where from managerial perspective, selection of right customers is a critical first step for cocreation in all of the NPD stages. Additionally, the firms can reduce cost by not employing external (idea contest) platforms and by using internal infrastructure which can strategically protect their innovations from imitation by competition [263]. The proposed way can enable the firms to explore with creative customers more while not incurring the (cognitive) cost of formulating difficult problems or incurring information overload of absorptive capacity. This second overarching research objective is addressed in the following article:

2.5.1 Research objective 2

- "*Understanding and predicting Innovator group customers in consumer goods industry: An Artificial Intelligence approach*". In this study, building on previous literatures on cocreation, adoption of innovation and AI, we propose a framework that can help identify and predict future innovator group customers from their transactional, demographic and behavioural data. With supervised machine learning algorithms, the study helped identify and predict future innovator group customers for a consumer good firm. In doing so, we further the customer cocreation with crowdsourcing research to include a *select crowd* of innovator group customers. The results indicate that combining experts' knowledge in determining features and their weights, with AI techniques can help managers identify the right customers with higher accuracy and less bias, and then they can approach these customers for cocreation to form a special crowd of knowledgable customers.

Table 2.1: Research objectives and methodologies

Chapters	Overarching research objective	Specific research goal	Methodologies
Ch. 3	To understand and classify innovator group customers with the help of their adoption decision	To propose a conceptual framework to organise innovator group customer's adoption decision factors, and linking the identification to the NPD performance through cocreation at ideation, prototype and post-launch diffusion mechanisms	Systematic literature review, with statistical trend analysis
Ch. 4	To understand and classify innovator group customers with the help of their adoption decisions	To find out the weights of the adoption decisions of innovator group with the help of industry experts	Fuzzy group decision making techniques - Analytical Hierarchy Process and Technique for Order Preference by Similarity to Ideal Solution
Ch. 5	To understand and identify innovator group customers as a select crowd for the cocreation purposes	To identify innovator group customers from the structured data with AI algorithms	Supervised machine learning algorithms with statistical analysis

Chapter 3

A conceptual framework on linking new product performance to Innovator group's decision-making

Abstract
New product failure is a major concern for manufacturers across industries. Extensive research has been conducted on New Product Development (NPD) in multiple research disciplines, and studied from many levels of analysis. As a result, the accumulated knowledge is difficult to interpret, and new product failure persists. Acknowledging these facts, we propose a customer-centric perspective to investigate the phenomenon. We identify 72 articles that examine customer adoption of innovation in a systematic literature review of 30 years across multiple disciplines. We propose a conceptual framework linking Innovator Group (IG) customers to new product development performance. The role of IG customers is crucial in NPD process, and we propose ways to utilize the customers' knowledge before launch, and then use their social influence post-launch to mitigate new product failure. We also provide insights and suggest future directions for advancing knowledge in cocreation research.

3.1 Introduction

New products are failing in consumer goods industry despite the best efforts from practitioners and academics [228]. Extensive research projects to understand this phenomenon have remained inconclusive as researchers have tried to answer questions such as: what is the speed to market of any new product? What is the role of integration between departments in a firm on New Product Development (NPD)? What are the success factors for NPD? However, the new product success rate has stagnated at 60%, and by inference the failure rate too [228]. Product failure is far more difficult to investigate than product success and far less focus has been placed on this area [128, 187]. Despite awareness about failures, NPD launch and post-launch diffusion are the least well-managed NPD phases in firms [114, 175]. Surprisingly, relatively few studies have investigated successful ways to manage NPD launch and post-launch diffusion to improve NPD performance [43, 63, 144, 147]. Overall, NPD performance research emphasizes success rather than failure Journal of Product Innovation Management (JPIM) published only six articles that contained "failure" in the titles between 2009 and 2019), and even meta-analyses [76, 101] show a lack of attention to failure factors [62, 175].

A different approach is needed to look at the problem: not from a firm's perspective, but from a customer's decision-making perspective. Supporting this view, marketing research calls for a new theoretical approach in NPD performance research [76]. Especially, owing to informational and technological evolution [95], customers have become central in firm strategy - customer orientation [254] is as important as market orientation [135]. Based on customer knowledge and experience, customers' adoption decision-making plays a greater role in product success. The literature has examined several constructs related to new-product adoption by customers (such as the product search process, information processing, inference making, and buying behaviour) and customer participation (recommendation, word-of-mouth, co-creation). Results show that customers are heterogeneous in nature when they search, evaluate, or recommend. So, the next question is: is there any group of customers who could specifically help us improve product success? Innovators and early adopters are the earliest customers to adopt new products chronologically [203]. Therefore, understanding these two groups (together known as the Innovator group) [161] and their adoption decision-making may help firms reduce failure.

The research gap needs to be addressed from the perspective of customers' adoption of innovation, which remains a fragmented but immensely important area of research for NPD. Hence, by focusing on both NPD performance and customer adoption of new products, we could improve understanding of the failure phenomenon. Therefore, the aim of this study is to understand new product performance from the perspective of the Innovator group customer, which leads to our primary research questions:

- What is a useful conceptual framework to understand IG customers' adoption decision-making affecting new product performance?
- What are the most influential factors affecting IG customer decision-making that impact new product performance?

In this paper, we provide a systematic literature review and descriptive analysis of cross-disciplinary new-product adoption research with citations to 72 relevant papers across specific research areas. Further, we synthesize the revealed categories from the literature review into a conceptual framework that links new product performance to IG customer adoption. Finally, we spell out the most important variables influencing new products' adoption by the IG customers.

Globally, we suggest changing the conversation on NPD research from a firm and organization perspective to a customer perspective, particularly when addressing product failure. Our focus on new product adoption is driven by the motivation to reduce failure by understanding customer adoption behaviour. Industry or country level models on NPD may highlight important aggregated insights; but from a managerial perspective, managers' actions are limited to their own firms. Therefore, our framework is constructed to be a practical starting point for managers to impact their firms by focusing on IG customers for NPD performance, while also suggesting possible future research directions.

3.2 Systematic literature review methodology

A Systematic literature review (SLR) is adopted as the most suitable method for this study to collect, synthesize, interpret, and provide valuable insights that go beyond summarizing [32]. Content analysis is conducted because it is an objective coding scheme to truncate data and to make it comparable using classifications and levels of analysis [166, 222]. The selected articles reviewed for this study were analysed using content analysis, following classifications similar to other systematic literature reviews [181].

3.2.1 Inclusion criteria

We searched the SCOPUS (Science-Direct) database as it is the largest full-text multidisciplinary academic database with citation analysis (others include Google Scholar, Web of Science, and Business Source Premier). The primary search terms were a combination of "early adopters", "innovators", "customers", and "retail". The principal reason to add retail into the search terms was to include all consumer and service products (B2C), and exclude industrial products (B2B). Our research focus is on consumer and service products or simply "products" as defined by Kotler [138](page.11) and supported by the NPD research [128, 187].

In terms of detailed inclusion criteria, we firstly limited our discussion to studies that measure customer adoption in new-product commercialization, making customer adoption our unit of analysis. Secondly, we retained articles about customer adoption of new products (rather than firm/industry NPD strategy adoption) between 1988 to October 2017. We did so to capture NPD research evolution four years after the launch of JPIM, and to capture the impact of new products such as personal computers, mobile devices, and the internet. Thirdly, given the growing consensus among NPD researchers that all consumer products and services should be considered as *products* [101, 138, 139], we did not exclude services from our search. Fourthly, in order to gain comprehensive insights, we included multi-disciplinary academic articles from "decision sciences", "computer science", "business, management, and accounting", "economics, econometrics, and finance", and "social science" in the search. Finally, we restricted our search to published peer-reviewed academic articles in support of our aim to include quality, proven, and value adding knowledge materials. We avoided the file drawer problem by excluding books, trade magazine articles, conference papers, doctoral theses, and other non-reviewed scholarly works.

3.2.2 Selection of articles

We employed all the search words in combinations of title or abstract or keyword for the articles in the SCOPUS database, and checked with results from other databases such as Google Scholar, Web of Science, and IEEE. Initial findings included 276 articles. Articles were discarded if they had no author names, different level of analysis, or focused on product characteristics. In total, 112 articles were identified for full-article reading. Thereafter, seven new articles were included from three editorial articles, increasing the total selection to 116 (by removing editorial articles). After comprehensive reading, 72 relevant articles (see Table 3.7) were selected from a total of 283 initial articles, meaning a 25% acceptance rate (which is in accordance with most of academic literature reviews).

Coding of articles

Each selected article was given a code. Classifications used in the content analysis were easy to determine by reviewing the articles. Following similar literature reviews [181], we broadly classified research study design into conceptual or empirical, and further into qualitative or quantitative, or mixed method. Similarly, the nature of empirical studies (survey, experiments, models, or simulations) and conceptual studies (conceptual framework, case studies, and literature reviews) were easily determined. Location of the study was based on real locations where data was collected and not based on the principal author's nationality. Each study has two codes: country and continent. Channel categorization was straight-forward as the study focused on e-commerce, mobile-apps, online and physical stores, or omni-channel.

For more subjective and tricky classifications, studies were twice rechecked to ensure the correctness of the categories. Theoretical foundations were assigned by analysing principal arguments, hypotheses, and research questions, and whether they are based on any established theory. Each article's theoretical background was coded to a contributing disciplinary domain. In certain cases, the initial publication of the theory was checked for certainty (an exception was lead user theory, published in Management Science as a marketing theory). Some articles contained multiple theories, in such cases, the theories were coded as dominant theories given that they were equally important. In this process, the number of theories exceeded the total article count. The categories that emerged from the theoretical background were one-theory, multi-theory, and A-theory respectively. Similarly, for industry categorization, we analysed the implications/scope of the studies. For example, an article "tailoring website appeal to customers" was assigned to the IT industry instead of retailing because the objective of the study was to explore the features of the website to attract new users.

To observe the changes in categories over time, the 30-year period was selected from 1988 to 2017 and this was divided into two periods: P1 for 1988-2002 and P2 for 2003-2017. The year range, or the terms P1 and P2, are used throughout the study to elucidate trends over time. Chi-square or Mann-Kendall tests are conducted on the major classifications to show the trend analysis, their relevance, and significance level.

3.2.3 Findings

This section shows the findings from the SLR and establishes the importance and relevance of adoption of innovation research.

Descriptive analysis

Journals and article growth. The growth in adoption of innovation research is clearly seen over the last 30 years (Fig. 3.1, Fig. 3.2), and the journal name abbreviations are in accordance with the Web of Science abbreviation list [275]. The 3-year average trend is statistically significant (Table 3.1), and the slope indicates that research output has grown by a factor of four over the study's total time period.

Figure 3.1: Trend in published articles 1988-2017

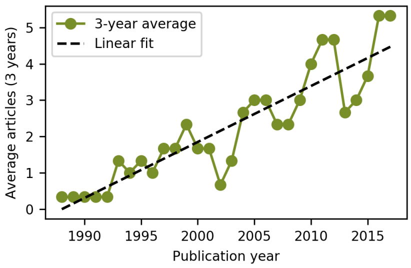
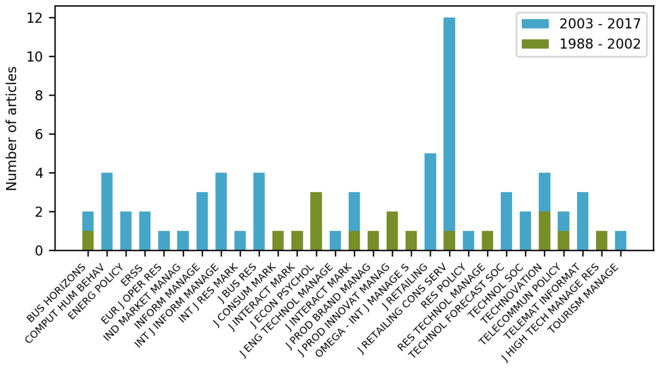


Figure 3.2: Journal distribution 1988-2017 (over two periods)



The selected journals are grouped into established research disciplines (Table 3.8). Marketing journals account for 35% of the articles; information management account for 19%; followed by innovation research journals at 14%. Articles on IG adoption in each of the 29 journals have either increased or maintained their research level over the periods.

Research design. We broadly classified research design into conceptual and empirical designs, following previous research on SLR [181, 187].

The categorized articles were analysed (see Table 3.2), and this revealed an increasing statistically significant trend for quantitative-empirical and mixed methodologies, and a declining trend for qualitative-conceptual research design. Interestingly, a trend analysis indicates salience for the first two methodologies with an increased focus on individual-level granular data availability and analysis. *Theoretical background/knowledge domains*.

Table 3.1: Journal trend (Mann-Kendall trend test) for articles published in 1988-2017

Kendall's tau	0.54
S	230
Var(S)	3037
p-value (Two-tailed)	3.3×10^{-5}
α	0.05

Table 3.2: Research methodology distribution 1988-2017 (over two periods)

Research Design	1988-2002	2003-2017	P2 over P1	1988-2017	
Conceptual - qualitative	7	6	-14%	13	18%
Empirical - qualitative	1	3	200%	4	6%
Conceptual - quantitative	0	6	NA	6	8%
Empirical - quantitative	8	34	325%	42	58%
Mixed methodologies	1	6	500%	7	10%
Total	17	55	224%	72	100%

The trend for theory shows that multi-theory articles have steadily increased compared to one-theory articles (see Table 3.3). Rogers' diffusion of innovation theory was the most referred theory (21% of articles) followed by A-theoretical (18% articles). Statistically significant trends for single and multiple theories from sociology and marketing disciplines are clearly seen in the adoption of innovation research (Table 3.2). The low attribution to A-theory (18% compared to a standard of more than 45% in other disciplines [28, 271]) could be ascribed to the maturity of the research domain. In summary, the trend on theory indicates that customer adoption of innovation research is progressing towards a strong multi-theoretical grounding.

Insights from content analysis

This section synthesizes the data obtained from the SLR, and the findings form the core of our proposed framework on IG customers. During the SLR process, one of the coded categories studied was the influencing factors of IG decision-making. A total of 103 unique factors were collectively studied in the 72 articles (see Table 3.6). After analysing these factors, it was evident that they could be meaningfully organized into four broad categories: *individual traits*, *product traits*, *environment*, and *relationship between product and individual traits*.

Individual traits. Based on the psychological and socio-psychological theories, the individual trait category is formed by factors that operate at the individual level. These personal traits affect adoption decision-making by IG customers. Personal innovativeness is one of the key distinguishing characteristics of IG customers, and it depends on a customer's degree of innovativeness and this acts independently from the communi-

Table 3.3: Mann-Kendall trend test for dominant theory types

	One theory	Multi-theory	A-theoretical
Kendall's tau	0.363	0.462	0.269
S	131	137	81
Var(S)	2640	1993	2165
P-value (two-tailed)	0.011	0.002	0.086
Alpha	0.05	0.05	0.05

Table 3.4: Summary of influencing factors on IG customer decision-making

Category	#Unique IF	% Share of IF	#Articles studied IF	% Share of articles
Environment	13	13%	34	10%
Product traits	21	20%	44	13%
Individual traits	37	36%	124	37%
Relationship between individual & product traits	32	31%	129	39%
Total	103	100%	331	100%

cated experience of others [172]. Multiple studies on the adoption of new technologies have investigated the level of innovation among individuals [6, 156, 188, 241] and found that personal innovativeness plays a major role in the intention to use new technology, especially for IG customers.

Other key individual traits include domain specific innovativeness, risk taking behaviour, rule breaking behaviour, and the free-spirited nature of IG customers. Some of the other salient personal traits that distinguish IG customers (during adoption) include self-motivation, do-it-yourself attitude, information-seeking attitude, hedonism, expertise, and inquisitiveness. The combination of these personality traits portrays IG customers as opinion leaders and experts in specific areas of knowledge, who willingly share their knowledge with social communities, and exhibit a strong inquisitiveness towards new consumer products.

Product traits. The research tradition of marketing helps categorize factors that are associated with products under the product trait category. Adoption literature emphasizes product characteristics as a major component for the adoption of innovation [175, 205] and primarily emphasizes product attributes, product advantage, and additional new product features. Although the intended usage and the advantage of the products are envisioned by manufacturing firms, they are perceived differently by end-users. Hence, product attributes clearly belong to the product trait, whereas the perception of product-related factors is placed under *relationship between product and individual traits*. The use of prototypes, trial offers, availability of products at retail stores (both online and offline), availability of choices, and the quantity of products are among the key factors that encourage IG customers to try new products.

Environment. Researchers grounded in sociological perspectives examined the social aspects of customers regarding the adoption of new products. All the factors categorized under the environment category operate externally to individual customers, and are socially embedded. Most of the sociological theories explain implicit social impact on individuals and the accumulation of social capital [154]. Theories such as social network theory, network externality, and social contagion [20, 122, 130] provide mechanisms for operationalizing these influences. Mass-media/brand advertisements, social norms, and social images also affect IG customers as they form subjective norms, values, cultural norms, social approval, and biases in general. Some macro-level institutional factors like regulations, commerce treaties, and competition may indirectly impact individual customers that go beyond the scope of this study.

Relationship between product and individual traits (RPIT). For the adoption of new products, customers engage and interact with products, collect information, interpret brand communications, gain experience, and on the basis of satisfaction, form perceptions about products and brands [153]. One of the critical findings of this study is the emergence of RPIT as the most impactful category in the adoption process. This is a new category to our knowledge that has not been proposed as a separate category in any previous study. The fact that RPIT turns out to be the most impactful category indicates that IG customers' perception, involvement, and interaction with products, create a strong influence during the adoption of new products. Innovation is subjectively judged, and the degree of involvement is conditional on the perceptions of IG customers, and on previous experiences. Similarly, the perceived benefits or advantages of a product for a customer differ from those perceived by the NPD team/manufacturer. Perception in general is a multi-level construct that is based on an individual's experiences, personal characteristics, and social influences.

Involvement with products is also a multi-level construct [283]. For example, involvement arises because of interest in information gathering that motivates the processing of complex information. Involvement may be higher in customers for their preferred firms/brands [5, 31](page 515). Similarly, trust plays a big role in indicating an IG customer's belief in the brand/product, and their willingness to make early investments [16, 167]. For an IG customer, trust is a pre-requisite for exploring innovation and newness. Similarly, price sensitivity is also perception based, and depends on brand credibility [73], product categories, and customer interactions with these brands.

The value of a product and price sensitivity are highly subjective in nature, as the valuation of a product and its price sensitivity are inversely related, i.e. the higher the perceived value, the lower is price sensitivity and vice versa. Satisfaction is another key construct that concerns product usage satisfaction and expectation compared to experience. Comparison with other products in a similar product domain, and overall brand experience are also perceived. Numerous brand touch points could enhance or diminish customer experience, and hence satisfaction level for IG customers varies during the entire customer purchase-journey [149].

Table 3.5: Mann-Kendall trend test for influencing factors on IG customers

	Environment	Individual trait	Product trait	RPIT
Kendall's tau	0.385	0.412	0.311	0.492
S	156	177	131	212
Var(S)	2698	3030	2905	3059
P-value(two tailed)	0.0028	0.0014	0.0159	0.0001
Alpha	0.05	0.05	0.05	0.05
Trend	increasing	increasing	increasing	increasing
Z	2.98	3.20	2.41	3.82

In the NPD and innovation literature, product and individual characteristics are considered more critical than external social influences and perceptions. In contrast to this belief, we observed that the most influential factors on adoption behaviour are perception about products and customer involvement with products that shape IG customer decision making. Some 39% of the SLR articles studied factors belonging to the *relationship between product and individual traits* category and with a statistically significant increasing trend (P2 vs P1) (see Tables 3.5, 4.3). This indicates the importance of *relationship between product and individual traits* in IG customer adoption behaviour. We argue that a gap exists in understanding customer decision-making because of the interplay between product and individual characteristics, which has not been fully explored. Hence, personal characteristics, and early formation of relationships with products differentiate IG customers' buying behaviour from that of the late adopters. The relationship formed between individuals and product traits are more valuable than either the product or individual traits (see Table 4.3).

3.3 Developing conceptual framework

In this section, we introduce a conceptual framework that links new product development performance to IG customer decision-making. Before that, we discuss the main constructs of the proposed conceptual framework to establish the background.

NPD performance. NPD performance refers to "the success of new product development efforts" [248](page.136) and this is considered to be a multi-dimensional construct

[128]. NPD efforts can be broadly divided into pre-launch and post-launch activities, and the performance is operationalized by external outcome (financial/marketing) or by operational outcome (product/process/market/strategy) [33, 101]. New product financial performance is measured by Return on Investment (ROI) and profitability; and marketing performance is measured by customer satisfaction, loyalty, sales/market share, and customer life-time value. In this study, we refer to financial outcome as the new product performance, following an NPD meta-analysis [43].

Adoption of innovation. In this context, we refer to innovation as very new products. Traditionally, adoption of innovation is an integral part of the diffusion of innovation model, and refers to acceptance by a population over a certain period of time [203]. Additionally, adoption of innovation depends on the innovativeness of the customers and products [205]. Some other research weakly supports innate innovativeness of customers with new-product adoption [117, 172], while some contradict the claim [91].

In the adoption literature, researchers also refer to adoption as acceptance of innovation. The technology acceptance model outlines the factors of new technologies perceived by customers. Customer acceptance as a construct is tested with behavioural intention (theory of reasoned action) and planned behaviour (theory of planned behaviour). Adoption and acceptance have been empirically tested by purchase intention, buying behaviour, and number of products owned [117] along with the time of adoption [203].

Innovator Group (IG) customers. IG customers are defined by the Rogers and Bass models of diffusion [12, 203]. This group is the first to adopt new products and also influences other customers when transmitting information. Greater product knowledge, expertise, and subject-matter interest (among other factors) make IG customers domain specific experts and their need for innovativeness acts as a core source for motivation [2, 157]. Studies have also shown that expectation of innovation-related benefits and need for early innovation are intrinsic qualities of lead users [160, 257]. IG customers and lead users are theoretically separated at the time of launch [242], and they share personal traits such as knowledge gathering, expertise, need for innovativeness – yet IG customers differ from the lead users because of their non-invention attitude.

Acquiring useful information about customer preference and need is crucial for NPD success in consumer products [129, 262]. As a consequence, user innovation has been incorporated into the NPD process [124, 182, 195]. Considering the benefits of including customers' knowledge, usage preferences and ideas [68], IG customers can be highly beneficial for firms trying to improve their products (idea generation and prototype improvement) and reduce product failures.

IG adoption decision-making. Customer journey literature delineates the entire purchase-process into three distinct stages: *awareness* (consideration); *evaluation* (purchase); and *experience* (post-purchase) [72, 149]. In general, decision making by an individual refers to a process of making choices (Oxford and Cambridge dictionaries), and customer decision-making is primarily related to buying choices. However, depending on the complexity of a product, information processing related to the product and evaluation among alternatives make the buying process rather difficult. In these situations, customers use heuristics [60, 127, 250], with brand names signalling quality [3], and take recommendations from experts and knowledgeable individuals they trust [155].

Nonetheless, IG customers buy new products without recommendations because they invest time and cognitive resources on domain interests, deliberately acquiring knowledge by considering complex information and developing intuition [58, 108, 127, 232]. The acquisition of knowledge makes IG customers knowledgeable and experts, and leads to sharing on their social networks [103, 146]. However, there are other influencing factors working on these customers [117, 172, 202] that shape their adoption decision-making. From the literature review (section 4.4.1), we collect the most investigated factors affecting IG customers in adoption decision-making.

Network analysis and knowledge sharing attributes are heavily researched and suggest that early adopters (IG) are central to social contagion [122, 253]. Studies also show that early adopters (here IG) are opinion leaders [40, 252, 267]. Therefore, researchers assign similar meaning to early adopters, Innovator group, opinion leaders, and initiators, and refer to them interchangeably in the context of social contagion.

The voices of customers and early feedback before large-scale production help reduce product failures [53]. Instead of involving random customers, the involvement of knowledgeable and domain-specific experts, i.e. IG customers, is better for testing new ideas. Although a substantial amount of research has been conducted to understand NPD performance and the characteristics of IG customers separately, no meta-analysis or SLR has been made to completely understand the phenomenon of new-product failure. To the best of our knowledge, we are the first to consolidate most of the investigated attributes on IG customer behaviour and link this to the NPD performance.

3.3.1 Conceptual framework

The primary motivation behind this paper is to help reduce new-product launch failure. Based on our in-depth SLR analysis (see Section 4.4.1), we conceptualized a framework to link new product performance to IG customer decision-making. To describe the framework clearly, we elaborate on three main aspects, referred to as sub-sections: identification of adoption decision-making factors of IG customers; pre-launch NPD process; and post-launch NPD commercialization (see Figure 3.3).

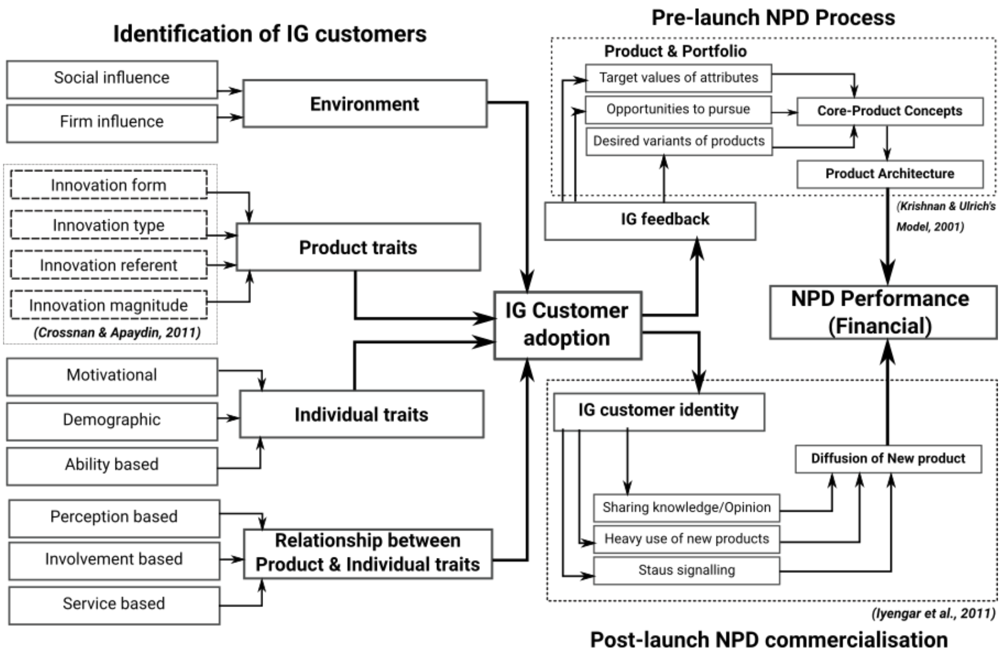


Figure 3.3: Conceptual framework

Identification of adoption decision-making factors of IG customer

IG customers are better positioned in terms of early need for new products and services as they are experts and knowledgeable customers – but they do not take part in actual innovation/invention. Non-participation of customers at new product developmental stage is beneficial for process optimization and speed to market for firms [43]. However, understanding the antecedents to first-user behaviour has value to the NPD process. Additionally, their (IG) decision-making process has vital information for NPD researchers about preferences, behaviour, and needs.

Since IG customers are crucial for new product performance, the next question is: "how do we identify their adoption decision-making factors?" To answer this question we examine factors used in existing literature to analyse IG customer decision-making. Within this context, we have synthesized four categories of influencing factors for new product adoption: *individual traits*; *product traits*; *environment*; and *relationship between product and individual traits* (see Section 3.2.3 and Figure 3.3).

Individual traits include some intrinsic factors that affect IG customers at an individual level, and can be divided into three sub-categories: *motivation*; *ability*; and *demographics*. Similar to the theories on information-search, we observe that IG customers are motivated to search, decipher, and share complex information. The motivational factors for IG customers (for adoption) that distinguish them from other customers are often cultivated attitudes rather than innate abilities. Risk-taking, hedonistic, do-it-yourself, variety-seeking, novelty-seeking, and information-seeking attitudes impel IG customers to make new product purchases (see section 3.2.3).

Ability or innate characteristics that are specific to individuals (psychological and personality based) also influence IG customer decision making. Researchers have studied some of these abilities in the adoption context, such as learning new subjects, gaining expertise, inquisitiveness, aesthetic sense, sharing ability, personal innovativeness, and domain-specific innovativeness. Extant research shows that opinion leaders (IG customers) generally harvest more information from the same source – and this demonstrates their superior information gathering ability [200]. These characteristics speak volumes about the personality of IG customers, and broadly depict them as knowledgeable who update their knowledge with learning, and who are cognitively capable of processing information about new products to develop domain specific (niche) expertise.

Demographic variables are generally treated as control variables for decision-making research but these factors are critical for IG customers. Research shows that IG customers are well educated, earn good incomes, and tend to be older. However, the age factor has changed dramatically in recent decades. As new technology and the ability to decipher new information has become imperative, age has declined in importance. Some studies have found that young adults are more innovative than older adults, and other studies show that when experience matters, older individuals are more likely to be IG customers [158, 186, 216].

Our second category of factors affecting IG adoption decisions is called *product traits*. Out of 103 influencing factors identified in our SLR, we grouped 21 factors to the product traits category, which we further sub-grouped into *attributes*, *availability* and *innovativeness*. Attractiveness of new features, visual appeal, functionality, size of a product belong in the *attribute* sub-category. The availability of choices at the point-of-sale, both online and physical stores, is an important determinant for adoption decision-making. Availability of trial products and prototypes offer IG customers the possibility of direct contact and interaction with new products, which influence their adoption decisions. Innovativeness of new products or product related services and technology also influence IG adoption behaviour.

We group external factors that shape IG customer perceptions, values, and subjective evaluations of new products into our third major category labelled *environment*. 13 factors in this category were coded into sub-categories *social influences* and *firm influences*. Socio-psychological theories explain how social influences - acting on individuals - gradually lead to accumulated social-capital built on social-interactions and social-ties. Although IG customers are the first to adopt new innovations, previous social influences heavily impact forming subjective norms, cultural norms and measurement of social status with possession of certain products. Firm generated influences - advertisements, mass-media communications, social media community - shape the social image of products and brands, and help building familiarity and brand-trust.

Lastly, the *relationship between product and individual traits* is a complex construct that includes perceptual factors formed by exposure to products; service factors that stem from previous experiences and expectations; and involvement factors formed by interactions with the products. We include a total of 32 factors into this category (see Table 3.6) and further sub-divided it into *perception*, *involvement* and *service* sub-categories for better understanding. Perceived usefulness, perceived ease of use, perceived brand trust/image, relative advantage, compatibility and perceived risk affect IG customers while they evaluate new products. Level of involvement, satisfaction while comparing, expectations, loyalty and usage indicate relationship of IG customer with the products that shape their decision-making. Previous experiences with product, service, technology affect these customers the way they evaluate other new products. Collectively RPIT is subjective category that constitutes with factors that determine how an IG customer forms relationship with products and product related services. It is the most important category for IG customers during adoption decision-making (section 3.2.3).

Pre-launch NPD process

This part of the framework is aimed to extend an NPD model proposed by Krishnan and Ulrich [139]. In their model (henceforth KU model), the authors proposed co-ordination among teams within a firm for an efficient NPD process. However, the KU model included team participation for idea generation on new products, and excluded customer participation in this process. Our framework adds adoption decision-making factors of IG customers, their ideas and feedback into KU model to address the missing link.

We add IG customer involvement into the KU model, to refine the selection of opportunities; the selection of product variants; the filtering of core-product concepts for new products, and the selection of product attributes. We propose that IG customers should be involved in both the ideation and launch stages, but not in the developmental stage [43].

Involving IG customers at the ideation stage will lead to choices that are consumer-oriented, and that capture consumer (explicit/implicit) needs – provided that the NPD teams use their expertise to blend the information to create new products. Unlike lead users who invent/innovate products in the developmental stage of NPD without impacting NPD performance [43], IG customers will share their knowledge on products and this complements the technical knowledge of NPD team for creating new (feasible) products. IG customers' feedback on later stages of NPD for prototypes will help firms fine-adjust their final products. The feedback could help align firm and customer needs at a strategic level for portfolio selection, and even improve product architecture. We argue that when the necessary changes are made to core concepts, product variants, and prototypes, the final improved products will have a better chance of rapid customer adoption than otherwise.

Post-launch NPD commercialization

The post-launch commercialization of our framework links IG customers to the social contagion process. This section complements the social contagion mechanism demonstrated by Iyengar *et al.* [122]. In their seminal study of social contagion, the authors identified early adopters in a long field-study that was the most challenging phase of their research. In the social contagion study, social positions and links helped recognize medical practitioners (IG customers) who were well-positioned in their network for diffusing information or inspiring imitation among other doctors for new products.

In our context, establishing the social links of identified IG customers (or opinion leaders) is secondary to identification itself, because opinion leadership and heavy usage weigh more than social-network positions [122]. Heavy users are considered attractive viral seeding points and are more effective than non-user experts in the diffusion of information on new-products. Opinion leaders or IG customers also gather more information from the same source than others [200]. A combination of cognitive ability, interest, and high usage makes these customers credible sources for new product information, specifically at the evaluation stage for later adopters compared to the awareness stage [122].

Without the knowledge of the social network position, we argue that simply identifying and targeting all IG customers - only 16% of all customers, depending on the product type - for Word-of-mouth (WOM) propagation of information is sufficient for the diffusion of new products. We provide a comprehensive list of factors for adoption decision-making of IG customers and it is a way to identify these customers. Unlike the field experiments conducted to map social-network positions of influencers (IG), our proposition is simpler.

Identification and targeting IG customers for diffusion of new products is valuable for firms dealing with product failures. Incentivising IG customers with brand benefits or recognition could build positive associations with the firms. After convincing these customers with satisfactory new products and/or service experiences, firms could nudge these customers to spread positive WOM (i.e. network externalities), or display their possession to others (i.e. status signalling). The next step would be to devise appropriate marketing strategies involving IG customers (deciding among viral product designs, social media forums, blog/vlog platforms) and pricing strategies to attract later adopters.

All together, the diversified approach of our framework is aimed at a balanced NPD process for launching, commercializing, and measuring performance. We argue that all three segments of our framework are equally important for the new product success, and firms with an integrated approach can reduce the possibility of product failure. Since stochastic changes in the market and society can affect a new product's chances of success, no realistically applicable strategy can summarily eliminate product failures. However, our aim is to provide a starting point for managers to gather crucial customer information that will lead them to create better customer-oriented products, and position themselves for fewer failures than the current rate.

3.4 Discussion

This paper examines the phenomenon of new product failure in consumer goods industry. Several disciplines have approached the same problem from different perspectives. Our conceptual framework (see Figure 3.3) is based on a systematic literature review, and examines NPD performance through a customer-centric approach. We focus on identifying and categorizing influencing factors behind adoption decisions of IG customers to improve new product performance. Although some research has linked customers (lead users) to NPD process [257], to the best of our knowledge, this paper is the first to involve IG customers in NPD performance in a rigorous manner.

According to a meta-analysis on NPD research, the implementation frameworks in NPD performance represents only a few of the articles published in this domain [128]. NPD is a complex process and no single framework explains it completely. However, the search for best practices is an ongoing process, and given the lack of consensus among academics on NPD performance, there is a clear need for systematic integration of findings in the form of a conceptual framework [84, 175].

In our framework, the pre-launch NPD process is a way to extend Krishnan and Ulrich's NPD framework with IG customer involvement. It is aimed at impacting at the ideation stage and at the pre-launch stage for prototype assessments, to strategically make market-oriented new products [144]. In doing so, we help align customer needs with products for better adoption. The post-launch NPD commercialization section in our framework complements Iyengar's *social contagion model* [122] by providing a novel way to identify IG customers. However, in contrast to Iyengar's field experiment, we collect and synthesize customer adoption decision-making factors from the existing literature.

3.4.1 Implications for theory

Existing research has studied NPD process, performance, and commercialization for various consumer products from multiple perspectives, and still struggles to explain the high product-failure rate [228]. The explanatory power of success factors have also declined since Henard and Szymanski's meta-analysis [76]. Meanwhile, customer participation at specific stages on NPD has shown a positive impact on NPD performance [43], strengthening the case for a customer-centric approach in NPD. Therefore, one of the focus areas of this article is development of a customer-centric conceptual framework linking IG customer feedback in NPD pre-launch phase and social influence in post-launch phase to NPD financial performance.

By expanding two well established theoretical frameworks on NPD pre-launch [139] and social contagion [122], the study helps explaining new product performance through customer participation, more specifically IG customer participation. IG customers are, by definition, the first users of new products. They are known to be knowledgeable and spread new-product information to later adopters. Diffusion of innovation may fail if IG customers fail to purchase or if they do not spread WOM either by opinion leadership or by signalling status. We realize that both of these mechanisms are inter-dependent and their combination could explain the product failure.

Customer decision-making has vital information for researchers about customer preferences, behaviour, and needs [129, 251, 262]. Understanding the antecedents to first-users' (IG) decision-making provides insights on NPD process and performance. Therefore, our conceptual framework helps identify the most investigated factors influencing IG customers in the context of new product adoption.

We have identified 103 influencing factors on IG adoption decision-making, and synthesized the knowledge into a concise framework with four major categories and sub-categories. Individual trait, product trait, environment, and relationship between product and individual trait (RPIT) constitute the major categories of influencing factors. While it is difficult to obtain significant insights on individual influencing factors, our categorization - by grouping several factors into broader categories - reveals interesting statistical trends. Contrary to a common belief in the reviewed literature, we found that RPIT to be more important than either product trait or individual trait category. One of the critical findings of this study is the emergence of RPIT as the most impactful category in the adoption process.

To the best of our knowledge, no other research has explored the possibility of involving IG customers in the NPD process, and our paper is the first to interlink different research domains and conceptualize an aggregated framework to improve understanding on adoption of innovation as part of the NPD research. Since NPD and adoption of innovation are widely dispersed research fields with varying degrees of research, integrating and providing an over-arching framework, helps both of these research domains.

Furthermore, we present relevant results on trends of adoption of innovation research and organize major parts of our framework by interpreting the multi-disciplinary research trends. Following guidelines by Whetten [269], we highlight the *what* aspect: necessary factors for understanding IG customers' adoption decision-making, and emphasize the *how* aspect: relevance of IG customer participation in NPD process to reduce product failure.

3.4.2 Implications for managers

In our study, we aim to provide a path for managers to take micro-level actions grounded in theory-based conceptual framework to control NPD performance during the product launches. Our framework supports marketing activities, including targeting and promotion to IG customers who would, in turn, exert influence on other customers. Therefore, involvement of this group is greatly beneficial for both pre-launch and post-launch of NPD processes.

Customer information is key for managers. In the pre-launch phase, augmenting Krishnan and Ulrich's framework [139] with IG customer involvement in idea generation, managers can select new-product attributes corresponding to customer preferences, select opportunities well-suited to customers' need, and then select technically possible variants of products. Firms can align their core-concepts for new products with customer preference/need, hence improving chances for subsequent successful product launches.

For post-launch success, IG customers are proven to be opinion leaders and influencers who publicly disseminate information on products, and so signal low risk with their usage volume. They also signal superior social-status through tacit knowledge to later imitators. Managers should recognize the role of IG customers as experts and incorporate communication strategies accordingly. Notwithstanding the way in which IG customers influence – imitation or risk mitigation – managers should use their information sharing capability, WOM, and high status in social network to disseminate positive product information. Managers should be aware of the interplay between social network and mass media communication for new product diffusion. For example, it is observed that increased advertising reduces online WOM [79]. Therefore, managers should take a balanced approach in employing multiple tools to induce diffusion for new product launches.

3.5 Future research directions and limitations

The ultimate goal of a conceptual framework is to provide an integrated way to look at the new product failure problem [118, 223]. On the basis of the expectations and limitations of our conceptual framework, we suggest several directions for future research in the NPD performance and customer-adoption research domains.

Firstly, innovation and marketing literature has begun to study the inclusion of customers, including the lead users, into the NPD process with mostly positive outcomes [43, 124, 182, 195]. This on-going line of research may be continued to explore the effect of including IG customers at either ideation or pre-launch stages of the NPD process to generate customer-centric ideas and improve prototypes respectively.

Based on the current work, we expect that IG customers' participation has a greater effect on NPD performance than the average customers' participation. This proposition can be examined in a future empirical study. Future research could also look at IG customer participation at specific stages. For example, speed-to-market moderated by IG customer participation at the ideation stage may increase NPD (financial) performance more than participation by average customers.

Secondly, the aggregated framework posits the possible impact on NPD performance when considering feedback from IG customers on consumer goods. It will be interesting to explore the variability of performance gains due to IG participation between consumer goods and industrial goods.

Furthermore, not many studies have examined IG customers' negative social contagion effect on diffusion of new products. Current research focuses on positive WOM and neglects the importance of negative effects when IG customers impede NPD performance. To what extent does negative WOM by IG customers disrupt diffusion of a new product? Future researchers could also explore the effect of lack of IG customers' opinion leadership for new products. It would be interesting to investigate both of these mechanisms' impact on the NPD performance.

Much of the social contagion literature investigates the impact of involving early adopters (IG customers) in the post-launch diffusion process, but has limited knowledge on its direct impact on the product performance. Further research attention should be focussed on some specific mechanisms in the NPD post-launch phase, such as incentivize IG customers with early access to the new products or exclusive brand partnerships with them. Higher product performance may also be mediated by faster WOM by IG customers.

Finally, the application of the model with field data could provide specific insights on the relationship between IG customers and NPD performance with some appropriate metrics.

3.5.1 Limitations

Despite a careful selection and structure of the study, it has limitations. Firstly, our framework is limited to influencing-variables gathered from 72 studies that we included in the SLR after a careful selection process. We have proposed a *conceptual framework* not a *theoretical framework* to address the phenomenon of new-product failures in the consumer market. Therefore, our research may lack generalizability, especially in the industrial goods market.

Secondly, we did not include any peer reviewed research before 1987 because we decided to focus on the changing landscape of adoption of innovation over the last three decades, and especially on current practices in academia and industry. Nonetheless, we may have missed some important variables prior to 1988 or post 2017. However, these shortcomings may have been mitigated by following the best practices of academic rigour in the systematic literature review [180, 245]. We considered only published and peer-reviewed articles from multiple disciplines. The information lost may be in line with the accepted standard in the literature review process that ensures only the best knowledge sources are included [237].

Furthermore, our framework offers a theoretical linking of IG customer feedback to NPD process, and it does not invalidate other arguments that NPD customers are drastically different to other customers. Generalizing IG customer preferences for the entire population may not work for all products. Nonetheless, there is no empirical evidence supporting the claim, even in high-technology sectors [49, 205].

Another caveat to our framework is that we did not consider the cost of collecting data on IG customers. We assume that in the Big-Data era, most firms have systems to collect customer usage and buying-behaviour information. However, a study on big data by McKinsey [24] showed that discrepancy in data collection and storage exists among geographical regions. Hence, managers should be aware of the effort and cost of collecting data for using the proposed framework across countries.

Finally, our theoretical framework builds on the social contagion mechanism for the diffusion of new products for NPD performance that hinges on the premise of IG customer influencing other customers. In addition, the framework considers existing scenarios where IG customers positively respond to actions by firms by providing feedback. However, it does not explain scenarios where they are non-responsive or non-reactive to these efforts by the firms.

3.6 Appendix

Table 3.6: Detailed information on influencing factors of IG customer decision-making

Environment	Individual traits	Product traits	RPIT
Network externality	Personal innovativeness	Availability of choice	Perceived usefulness
Mass media influence	Hedonistic attitude	Incentives	Perceived benefits
Social approval	Income	Attributes of product	Perceived ease of use
Social image	Information seeking attitude	Quantity	Perceived brand image
Social norm	Trust	Speed of use	Perceived risk
Cultural norm	Age	Trial product offer	Price sensitivity
Number of initial adopters	Positive attitude towards innovation	Availability of choice	Perceived compatibility
Industry standard	Education	Functionality of product	Previous experience
Industry imperfections	Expertise	Product innovativeness	Relative advantage
Normative influence	Inquisitiveness	Store distance	Involvement
Number of followers	Risk taking attitude	Technical specifications	Frequency of purchase
Status signaling	Do-it yourself-attitude	Innovativenes of service	Expectation of product
Urbanisation	Domain specific innovativeness	Innovativeness of technology	Satisfaction while comparing
	Learning	Personalised offering	Usage
	Variety seeking attitude	Prototype	Store loyalty
	Aesthetic value	Radical innovation	Expectation of service
	Convenience seeking attitude	Search goods	Perceived complexity
	Knowledge sharing attitude	Size of product	Perceived guidance of information
	Novelty seeking attitude	Store reputation	Purchase intention (online)
	Self monitoring	Store service	Security concern
	Subjective norm	Technology generation	Adoption duration
	Absorptive capacity	Visual appeal	Evaluation of service
	Cognitive dissonance		Forming evaluative criteria
	Empathetic attitude		Observability
	Functionality attitude		Perceived brand trust
	Gender		Perceived web security
	Impulsive attitude		Previous experience with computer
	Innovative style		Previous experience with Internet

	Life event	Previous experience with social media
	Materialism	Privacy concern
	Media dependency	Search time
	Need for innovative products	
	Opinion leadership	
	Resistance to mass media	
	Self confidence	
	Self-efficacy	
	Self-initiated	
	Strong individuality	
	Values	

Table 3.7: Table with detailed Systematic literature review - article information

Article ID	Author name(s)	Year	Journal of publication
A1	Susan L. Holak	1988	Journal of Product Innovation Management
A2	Soyeon Shim, Marianne Y. Mahoney	1991	Journal of Direct Marketing
B217	A.L. Brannon, M.A. Schoenmakers, H.P. Klapwijk, K.B. Haley	1993	Omega International journal of Management Science
A3	Gordon R. Foxall, Seema Bhate	1993	Journal of Economic Psychology
A4	Fareena Sultan, Russell S. Winer	1993	Journal of Economic Psychology
B25	Steve Dunphy, Paul A. Herbig	1995	The Journal of High Technology Management Research
A5	Roland Pepermans, Gino Verleye, Sarah Van Cappellen	1996	Journal of Economic Psychology
A6	Richard W. Olshavsky, Richard A. Spreng	1996	Journal of Product Innovation Management
B26	Deborah Fain, Mary Lou Roberts	1997	Journal of Interactive Marketing
A7	Seema Bhate & Kevin Lawler	1997	Technovation
A9	Isabelle Szmigin & Gordon Foxall	1998	Technovation
B213	Supriya Singh	1999	Telecommunications Policy
A10	Jim Blythe	1999	Journal of Product and Brand Management
A11	Naoufel Daghfous, John V. Petrof & Frank Pons	1999	Journal of Consumer Marketing
A195	John H. Friar & R. Balachandra	1999	Research Technology Management
B163	Lawrence M. Bellman	2001	Business Horizons
B105	Marilyn Lavin	2002	Journal of Retailing and Consumer Services
A13	Natalie Muzinich, Anthony Pecotich & Sanjay Putrevu	2003	Journal of Retailing and Consumer Services
A14	Arthur W Allaway, David Berkowitz & Giles D'Souza	2003	Journal of Retailing

B261	Abhijit Biswas & Dipayan Biswas	2004	Journal of Business Research
A15	Andrew J Rohm & Vanitha Swaminathan	2004	Journal of Business Research
A16	Amit Bhatnagar & Sanjoy Ghose	2004	Journal of Retailing
A18	Margherita Pagani	2004	Journal of Interactive Marketing
A19	Leo R. Vijayasarathy	2004	Information & Management
B70	Michael Beverland & Michael Ewing	2005	Business Horizons
A20	Brian F. Blake, Kimberly A. Neuendorf & Colin M. Valdiserri	2005	Technovation
A21	Dale Littler & Demetris Melanthiou	2006	Journal of Retailing and Consumer Services
A22	Yifan Lu & Margaret Rucker	2006	Journal of Retailing and Consumer Services
A23	Chi Shing Yiu, Kevin Grant & David Edgar	2007	International Journal of Information Management
A24	Adam Faiers, Matt Cook & Charles Neame	2007	Energy Policy
A25	Hae Young Lee, Hailin Qu & Yoo Shin Kim	2007	Tourism Management
A26	Margherita Pagani & Charles H. Fine	2008	Journal of Business Research
A27	Jiunn-Woei Lian & Tzu-Ming Lin	2008	Computers in Human Behavior
B3	Nathalie T.M.Demoulin & Pietro Zidda	2009	Journal of Retailing
B49	Tommi Laukkanen, Suvi Sinkkonen, & Pekka Laukkanen	2009	International Journal of Information Management
A28	Mohammad Ali Zolfagharian & Audhesh Paswan	2009	Journal of Retailing and Consumer Services
A59	Robert J Kauffman & A. Angsana	2009	Telecommunication Policy
B12	Richard Clodfelter	2010	Journal of Retailing and Consumer Services
B159	Leo Pennings, Thijs Veugen, & Annemieke de Korte	2010	Technology in Society
A29	Kendra Fowler & Eileen Bridges	2010	Journal of Retailing and Consumer Services
A30	Hui-Chih Wang & Her-Sen Doong	2010	Information & Management
A31	Hanool Choi, Sang-Hoon Kim & Jeho Lee	2010	Industrial Marketing Management
A32	Renana Peres, Eitan Muller & Vijay Mahajan	2010	International Journal of Research in Marketing
B75	Chia-Liang Hung, Jerome Chih-Lung Chow & Tse-Ping Dong	2011	International Journal of Information Management
A33	Chuanlan Liu & Sandra Forsythe	2011	Journal of Retailing and Consumer Services
A34	Yaobin Lu, Shuiqing Yang, Patrick Y.K. Chau & Yuzhi Cao	2011	Information & Management
A35	Seunghyun Lee, Sejin Ha & Richard Widdows	2011	Journal of Business Research
B123	Akhter, Syed H.	2012	Journal of Retailing and Consumer Services
A36	Frank J. van Rijnsoever & Harmen Oppewal	2012	Technological Forecasting and Social Change
A37	Dominik Mahr & Annouk Lievens	2012	Research Policy
A38	Ivan Diaz-Rainey & Dionisia Tzavara	2012	Technological Forecasting and Social Change
B48	Erin H Green,Steven J Skerlos, & James J. Winebrake	2014	Energy Policy

B144	Ebru Uzunolu & Sema Misci Kip	2014	International Journal of Information Management
B152	Garry Wei-Han Tan, Keng-Boon Ooi, Siong-Choy Chong & Teck-Soon Hew	2014	Telematics & Informatics
A39	Shun Yin Lam & Venkatesh Shankar	2014	Journal of Interactive Marketing
A40	Chelsea Schelly	2014	Energy Research & Social Science
A41	Jing Li, Umut Konus, Koen Pauwels & Fred Langerak	2015	Journal of Retailing
A42	Tomi Nokelainen & Ozgur Dedehayir	2015	Technovation
A43	Jihyun Kim & Kim H.Y. Hahn	2015	Computers in Human Behavior
A45	Thanh-Thao T. Pham & Jonathan C. Ho	2015	Technology in Society
A46	Sonia San-Martín, Jana Prodanova & Nadia Jiménez	2015	Journal of Retailing and Consumer Services
B262	"Rebecca Jen-Hui Wang, Edward C. Malthouse & Lakshman Krishnamurthi	2015	Journal of Retailing
A48	Heetae Yang, Jieun Yu, Hangjung Zo & Munkee Choi	2016	Telematics & Informatics
A49	Denghua Yuan, Zhibin Lin & Ran Zhuo	2016	Computers in Human Behavior
A50	Guoyin Jiang, Pandu R. Tadikamalla, Jennifer Shang & Ling Zhao	2016	European Journal of Operational Research
A51	Stuart J. Barnes & Andrew D. Pressey	2016	Technological Forecasting & Social Change
A52	Jaewon Choi & Seongcheol Kim	2016	Computers in Human Behavior
B2	Genevieve Simpson & Julian Clifton	2017	Energy Research & Social Science
A53	Thamaraiselvan Natarajan, Senthil Arasu Balasubramanian & Dharun Lingam Kasilingam	2017	Journal of Retailing and Consumer Services
A54	Seok Chan Jeong, Sang-Hyun Kim d, Ji Yeon Park & Beomjin Choi	2017	Telematics & Informatics
A56	Liu Feng, Zhao Shaoqiong & Li Yang	2017	Journal of Retailing and Consumer Services
A57	F. Liébana-Cabanillas & M. Alonso-Dos-Santos	2017	Journal of Engineering and Technology Management

Table 3.8: Detailed information on disciplinary publication

Research discipline	Publication/Journal name	1988 - 2002		2003 - 2017		1988 - 2017	
		n	%	n	%	n	%
Economics, Econometrics and Statistics	Journal of Economic Psychology	3	18%	0	0%	3	4%
General Management	Business Horizons	1	6%	1	2%	2	3%
	Journal of Business Research	0	0%	4	7%	4	6%
Total		1	6%	5	9%	6	8%
Information Management	Computers in Human Behavior	0	0%	4	7%	4	6%
	Information & Management	0	0%	3	5%	3	4%

	International Journal of Information Management	0	0%	4	7%	4	6%
	Telematics and Informatics	0	0%	3	5%	3	4%
Total		0	0%	14	25%	14	19%
Innovation Research	Journal of Engineering and Technology Management	0	0%	1	2%	1	1%
	Technovation	2	12%	2	4%	4	6%
	Research Policy	0	0%	1	2%	1	1%
	Journal of Product Innovation Management	2	12%	0	0%	2	3%
	Journal of High Technology Management Research	1	6%	0	0%	1	1%
	Research Technology Management	1	6%	0	0%	1	1%
Total		6	35%	4	7%	10	14%
Marketing	Industrial Marketing Management	0	0%	1	2%	1	1%
	International Journal of Research in Marketing	0	0%	1	2%	1	1%
	Journal of Consumer Marketing	1	6%	0	0%	1	1%
	Journal of Interactive Marketing (formerly JDM)	1	6%	0	0%	1	1%
	Journal of Interactive Marketing	1	6%	2	4%	3	4%
	Journal of Product and Brand Management	1	6%	0	0%	1	1%
	Journal of Retailing	0	0%	5	9%	5	7%
	Journal of Retailing and Consumer Services	1	6%	11	20%	12	17%
Total		5	29%	20	36%	25	35%
Operations and Technology Management	European Journal of Operational Research	0	0%	1	2%	1	1%
	Omega:International Journal of Management Science	1	6%	0	0%	1	1%
	Technology in Society	0	0%	2	4%	2	3%
	Technological Forecasting and Social Change	0	0%	3	5%	3	4%
Total		1	6%	6	11%	7	10%
Sector research	Energy Policy	0	0%	2	4%	2	3%
	Energy Research & Social Science	0	0%	2	4%	2	3%
	Telecommunications Policy	1	6%	1	2%	2	3%
	Tourism Management	0	0%	1	2%	1	1%
Total		1	6%	6	11%	7	10%
Total		17	100%	55	100%	72	100%

Chapter 4

A Fuzzy decision-making approach to define a framework for understanding Innovator-group customers

Abstract
Consumer goods sector is predicted to grow in the future, despite the failure of majority of new products. The extant literature on New Product Development (NPD) has looked into the failure phenomenon intently to understand why still most of the new products launched are failing. We suggest that the core of the problem lies with customers i.e. customer's adoption/non-adoption of new products. In this chapter, we propose a conceptual framework to identify Innovator Group (IG) customers' most important adoption decision-making factors validated by industry experts. A combination of group decision making techniques with 16 experts' opinion with fuzzy logic allows capturing managers' imprecise and tacit knowledge. The final ranking reveals the most crucial factors of IG customers' adoption on new products tend to be perceptual, visual and innovation driven. We suggest implications of the study for both academia and managers at pre-launch and post-launch phases of NPD.

4.1 Introduction

A paradox is observed in consumer goods sector. Globally, 38,000 new products are launched every month in consumer goods industry [173], and irrespective of the best efforts from the firms, 60 % of new products fail within the first 3-years of their launches [228]. Managers are uncertain on exact reasons for these failures [99, 228], despite careful planning of the new products' developmental stages. Extant literature on New Product Development (NPD) is diverse and multi-disciplinary in nature making the aggregated knowledge rather difficult to comprehend [7, 51, 101, 110, 128, 150, 187, 191, 247]. Nonetheless, the scattered research provides explanation for 5% of new product success [76]. NPD, that includes new product performance, is a complex and expensive process, and it requires highly demanding resources. When new products fail to perform in the market, the collective resources invested in the process gets wasted. Considering these issues, the extant literature has provided some insights on NPD performance. However, investigation by prior research on determinants of new product success and failure has remained inconclusive - with lower explanatory power of the factors which in turn have declined over the last decade [76].

Although NPD research is important for academics and managers alike [43, 74], less attention is paid in order to reduce the new-product failure. Organisational theories in the context of NPD have not elucidated the failure phenomenon satisfactorily, even looking through team [106], project [238], firm, and industry level determinants for understanding the product performance extensively. And while some research has looked into customer participation in NPD phases [43], not much attention is paid to link customer adoption to NPD performance. Customer adoption perspective not only includes customer preference but also adds socio-cultural, environmental, perceptual and individual factors into the equation. Because of the complexity and importance of new product performance for firms, marketing community needs a new theoretical approach [71, 76] i.e. a customer adoption approach to better understand the failure phenomenon. This customer centric approach may help explain NPD failure phenomenon better than the other approaches because customers are the ultimate buyers [119, 170].

Innovator Group customers, who adopt new products first [161, 203] have been found to be better in generating ideas for NPD process [7, 97, 107, 257]. Additionally, based on the time of adoption, IG customers play a critical role in the overall product diffusion process to late adopters [203, 204]. The influencing factors that have a higher impact on the adoption decisions of IG customers may lead to reduce the number of new-product failures. Hence, in this chapter, our primary objective is to investigate/understand new-product's failure phenomenon from IG customer's perspective, which in turn aligns with the academia's suggestion to improve new product development research [71].

In a hyper connected world where information-overload and information-asymmetry exists simultaneously, customers need to navigate their purchase journey in this chaotic environment. This makes customers and their decision-making process highly complex as multitude of factors influence their behaviour. On the other hand, managers have different backgrounds and knowledge and they need to interpret reality from a vast amount of data/information. There is no unique way of seeing things: a tool based on *fuzzy logic* that is capable of capturing all this idiosyncrasies is required in the market. Moreover, managers can base their decisions on customers' needs and preferences and on the understanding that customers' behaviour is not black or white [34]. The wisdom-of-crowd research shows that a few experts can provide better knowledge than a large crowd of individuals [164]. A few industry experts who have considerable amount of experience in new-product launches can provide valuable knowledge on the failure phenomenon. Hence, taking their knowledge as input resource together with the results from a systematic literature review on adoption factors of IG customers and utilising a combination of Group Decision making (GDM) techniques, we can obtain the ranking on most important adoption factors. Therefore, the purpose of this paper is to provide a fuzzy group decision making approach combining Analytic Hierarchy Process (AHP) and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) methods to define a framework to help marketers in better understanding IG customers' adoption process. Additionally, to best of our knowledge, this is the first study to employ fuzzy hybrid GDM techniques to a marketing problem in consumer goods sector.

The paper is organised as follows. The preliminaries on the methodology of fuzzy logic, group decision making and proposed methodology for the current study are explained in detail in the following section 2. The conceptual background of new product development, adoption factors affecting IG customers and experts knowledge extraction are described in the section 3. In the section 4, results of the study is laid out, and in the last section, discussion and future research directions are suggested.

4.2 Methodology

Human judgements and opinions are embedded in complex socio-cultural settings. It is hard to translate human judgement into crisp ordinal numbers as practised by the survey methods, especially evaluating subjective values into objective terms [266]. However, there are ways to transfer subjective, imprecise, and tacit information into mathematical functions. Fuzzy logic that was proposed by Zadeh [281] allows capturing this tacit and impreciseness essential to human-judgments well. Application of fuzzy logic in management and supply chain has evolved into a big research field over the past decades.

4.2.1 Concepts on Fuzzy set theory

Fuzzy set theory is the underlying mechanism for fuzzy logic which is good in dealing with imprecise linguistic concepts, and the nature and functioning of fuzzy logic makes it a preferred choice for complex segmentation and identification problems. Detecting customer defection is an example for a fuzzy set application. In the retail setting, a study shows that identification of the most-probable customers to defect the firm was determined by fuzzy membership function [35]. In such a scenario, fuzzy membership is more suited instead of a classical membership (crisp) because it assigns each element (customer in this study) a degree of membership for defection whereas a crisp membership divides the elements into crisp partitions of either 1 or 0. In another study for industrial marketing planning, an agent based fuzzy cognitive map was used for strategic planning for the South Korean firms [145]. Due to the numerosity of variables for decision-making in manufacturing and the associated complexity, fuzzy based mapping was apt for finding the solutions.

A fuzzy set A, is defined by a membership function μ_A: $\Omega \rightarrow [0, 1]$ which represents the grading of an element to belong to the fuzzy set A. Ω considered as a reference set, containing the subset A and value of function $\mu_A(x)$ corresponds to a number in an interval 0 and 1. A triangular fuzzy number (TFN) is a particular case of fuzzy set where $\Omega = R$, and with the piecewise convex linear membership function μ_A defined by:

$$\mu_A(x) = \begin{cases} \frac{x-l}{m-l} - \frac{l}{m-l}, & l \leq x \leq m; \\ \frac{u-x}{u-m} - \frac{u}{m-u}, & m \leq x \leq u; \\ 0, & \text{otherwise.} \end{cases} \tag{4.1}$$

A triangular fuzzy number in general is denoted by (l, m, u) where the parameters l, m, and u are the lower bound value, the modal value, and the upper bound value respectively, Among the fuzzy system, triangular fuzzy numbers are widely used as membership function because of their computational efficiency [29, 46, 236].

In this chapter, we consider linguistic terms defined via triangular fuzzy numbers, because they offer an easy translation between ordinal linguistic terms and numerical ones for handling impreciseness [69]. Linguistic variables, in general, reduce the complexity of the model, specifically in cases where multiple experts are involved. An example of the type of linguistic variables is shown in Table 4.1.

Table 4.1: Linguistic scale for rating categories

Linguistic terms (importance)	Equally	Weakly	Strongly	Absolutely
Triangular fuzzy numbers	1,3,5	3,5,7	5,7,9	7,9,10

Introducing linguistic values for the quantification of a variable is motivated by the possibility of "computing with words" [69] i.e. using words or sentences rather than numbers because linguistic characterisations are in general less specific than numerical ones [277]. From the equation 4.1, operations for triangular fuzzy numbers can be defined through the pointwise operations over the interval [0, 1]. Given two fuzzy numbers such as A_1 and A_2 represented by (l_1, m_1, n_1) and (l_2, m_2, n_2) respectively, then the following operations can be defined as generalisation from respective operations in crisp sets.

1. Addition: $A_1 + A_2 = (l_1, m_1, u_1) + (l_2, m_2, u_2) = (l_1 + l_2, m_1 + m_2, u_1 + u_2)$
2. Multiplication: $A_1 \times A_2 = (l_1, m_1, u_1) \times (l_2, m_2, u_2) = (l_1 \times l_2, m_1 \times m_2, u_1 \times u_2)$
3. Division: $A_1 \,/\, A_2 = (l_1, m_1, u_1) \oslash (l_2, m_2, u_2) = (l_1 \,/\, u_2, m_1 \,/\, m_2, u_1 \,/\, l_2)$
   where, $l_1, l_2 > 0; m_1, m_2 > 0; u_1, u_2 > 0$
4. Reciprocal: $A_1^{-1} = (l_1, m_1, u_1)^{-1} = (1/u_1, 1/m_1, 1/l_1)$
   where $1/u_1 \,, 1/m_1 \,, 1/l_1 > 0$

In addition, in order to handle the fuzzy AHP method proposed by Chang [42], now we will introduce the method to obtain the priority vector from a Matrix whose values are TFNs. Let $S = T_i^j$, be a n x m dimensional matrix with TFNs, then the value of fuzzy synthetic extent with respect to the i^{th} row is defined as.

$$S_i = \sum_{j=1}^{m} T_i^j \odot \left(\sum_{i=1}^{n} \sum_{j=1}^{m} T_i^j \right)^{-1} \tag{4.2}$$

Note that, this synthetic extend gives us a normalized contribution of the i^{th} row to the global expression given by matrix S. In addition, the results obtained may not be symmetric TFNs.

Finally, we introduce a distance between triangular fuzzy sets because it is necessary in the application of fuzzy versions of some Multi Criteria decision making (MCDM) methods. There exists several ways to calculate the distance, for instance an extension of the Euclidean distance also known as *vertex method* [45] applicable for fuzzy triangular number is shown here:

$$d(A_1, A_2) = \sqrt{\frac{1}{3}[(l_1 - l_2)^2 + (m_1 - m_2)^2 + (u_1 - u_2)^2]} \tag{4.3}$$

4.2.2 Group decision making

Multi Criteria decision making (MCDM) is a sub-field in operational research which is useful to select, sort and rank alternatives. Selecting the best alternative based on some criteria, assigning weights to each criteria, and ranking alternatives on subjective scale are some of its applicability that broadens the scope to fields such as, business management [278], marketing [61, 276], supply-chain management [37, 278]. In particular, GDM is considered a sub-area of MCDM methods where selecting the best alternative is based on opinion of a group of experts [142].

Extant literature on wisdom-of-crowd suggests that depending upon the environment, a select crowd (group of experts) is selected for the best judgement against a whole crowd or one best-expert [164]. Instances where human expertise is imperative in solving complex multi-level problems is ideal for GDM techniques with fuzzy logic [105]. Because of fuzzy logic's effectiveness and flexibility in handling uncertainty that is inherent in experts' judgements, this study goes in this direction by utilising fuzzy linguistic terms to find the most important features for group decision-making from experts' knowledge.

For a real-world application of GDM methods, two major stages are involved: first, collection of information from the decision makers/experts on criteria weights and criteria values; second, aggregation by a method to rank, select or group the alternatives. The purpose of aggregation is to obtain collective preference value by combining all the individual preference values. There are some basic conditions that an aggregator operators should fulfil such as monotonicity, boundary conditions, associativity, neutral element and idempotency [30].

Aggregator operators can be further categorized into *first* and *last* aggregators depending upon when the weights are incorporated into the decision matrix of decision makers [206]. The accuracy of the methods depend on the variation in decision makers' opinions: when the variation in the opinions are low, *first aggregation* is preferred and when variation is high, *last aggregation* is preferred. On aggregation methods in fuzzy sets, in the context of our paper we use additive and multiplicative aggregation operators and *last aggregation* is performed according to opinions among the decision-makers [45, 236].

Fuzzy Analytic hierarchy process (FAHP)

Analytic Hierarchy Process (AHP) is a structured technique to solve complex MCDM problems by analysing pair-wise comparisons provided by the decision-makers. It is a relative measurement technique applied to get best order (output) among multiple attributes with judgements (inputs) from experts [209, 210]. It is best suited when exact comparison between choices is impossible to make but relative evaluation is still possible for the decision makers. Applications of AHP include various decision-making scenarios such as selection, planning, resource allocation among alternatives [50]. It is heavily used in public policy and operational research while use-cases in Marketing are limited considering application of AHP was first conceptualised in marketing [273]. Decision making, choice-making and selection are some of the most common problems in marketing discipline for advertisement, pricing, consumer preference, choice-decision, and selection in general [209].

In the first step, individual decision-maker's matrices are constructed where the comparative values are placed on the upper-right triangle of the matrix. The reciprocal values for each comparison is inserted into the lower-left triangle to complete the matrix. When the priority scale is simple in understanding for the decision-makers, the chance of ambiguity among experts can be easily ruled out [120]. However, when experts are not very confident about their pair-wise comparisons, a fuzzy linguistic scale is used instead of a numerical scale (as shown in Table.4.1). In reality, the experts don't see the fuzzy numbers as described in the Table 4.1, rather they only see the linguistic terms while doing pair-wise evaluations. In general, triangular fuzzy numbers are used to capture linguistic terms from the experts that can be transformed with fuzzy transformations without much loss of information.

In the methodology used in this chapter, the priority or weight vector is calculated by means of the fuzzy AHP. After collecting pair-wise evaluations from the experts, individual square matrices $n \times n$ are constructed for each expert M_{ij}, where $i = 1, 2, ..., n$ and n being the number of criteria or factors considered. At first, the pair-wise values are inserted into the top right-hand side triangle of the matrix (above the diagonal), leaving the left-side triangle empty (below the diagonal). Afterwards, the diagonal is filled with unity value of (1,1,1) as normally used in the TFN matrices. Finally, the left-side triangle is filled with the inverse values of the right-hand triangle, following the properties of Triangular Fuzzy Numbers (TFN)s, to complete the matrix. Note that the results obtained by applying the operator defined in 4.2.1 are TFN.

Then, to obtain the experts' collective evaluations on the criteria, expert matrices M_{ij} are merged into one square matrix M (see below). Aggregators, as presented in section 4.2.2, are used to transfer information from individual matrices to a merged matrix M.

$$\begin{bmatrix} (l_{11}, m_{11}, u_{11}) & (l_{12}, m_{12}, u_{12}) & ... & (l_{1j}, m_{1j}, u_{1j}) \\ \vdots & \vdots & \vdots & \vdots \\ (l_{k1}, m_{k1}, u_{k1}) & (l_{k2}, m_{k2}, u_{k2}) & ... & (l_{kj}, m_{kj}, u_{kj}) \\ \vdots & \vdots & \vdots & \vdots \\ (l_{i1}, m_{i1}, u_{i1}) & (l_{i2}, m_{i2}, u_{i2}) & ... & (l_{ij}, m_{ij}, u_{ij}) \end{bmatrix} \tag{4.4}$$

The next step is to calculate the synthetic extension with respect to each criterion, that will allow us to obtain a set of fuzzy values. In order to find the synthetic extension values, we followed the Eq. 4.2 proposed by [42]: For each row, a single synthetic value (S_i) is obtained which is also a TFN. Then putting together all the values of S_i, we obtain a matrix S of TFNs of the dimension of n $\times$ 1.

Finally, to obtain the estimated values for the vector of weights corresponding to each criterion, we consider the following comparison relation among TFN. Let's assume that S_i and S_j are synthetic extended triangular fuzzy numbers with elements of (l_i, m_i, u_i) and (l_j, m_j, u_j) respectively. Since S_i and S_j are convex fuzzy numbers, the degree of possibility of S_i being greater or equal to S_j i.e. V($S_i \geq S_j$) is computed considering the following expression,

$$V(S_i \geq S_j) = \begin{cases} hgt(S_i \cap S_j), & m_i < m_j \\ 1, & m_i \geq m_j \end{cases}$$

where

$$hgt(S_i \cap S_j) = \frac{(l_j - u_i)}{(m_i - u_i) - (m_j - l_j)}$$

is the highest ordinate value of the points in $S_i \cap S_j$ which can be seen in the Figure 4.1.

To obtain the estimates for the vector of weights, we need to compute all the values of V($S_i \geq S_j$) and V ($S_j \geq S_i$) where i $\neq$ j. Then for each criterion with synthetic extension S_i, we consider

$$d_{S_i} = \quad min[V(S_i \geq S_k)], \quad for \quad k = 1, 2, 3, ...n, \quad \wedge \quad k \neq i$$

The weight vector is calculated after obtaining all the values of d_{S_i} and normalising as the following.

$$W = \frac{d_{S_1}}{D}, \frac{d_{S_2}}{D}, \frac{d_{S_3}}{D}, ... \frac{d_{S_i}}{D} \tag{4.5}$$

where D is the sum of $d_{S_1}, d_{S_2}, d_{S_3}, ...d_{S_i}$. Note that, the final weight values are non-fuzzy

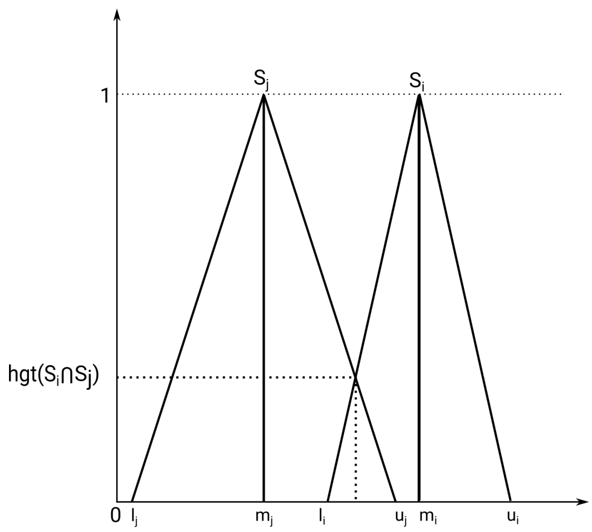


Figure 4.1: Comparative function between Triangular Fuzzy numbers

numbers [42].

Fuzzy Technique for Order Preference by Similarity to Ideal Solution (FTOPSIS)

Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) is a distance-based approach to rank alternatives in multi-criteria decision-making situations. Originally developed by Hwang and Yoon [115] for selecting the best alternative with finite number of criteria in a simple ranking system. TOPSIS selects an alternative that is simultaneously closest to the ideal solution and farthest from the negative ideal solution.

The foremost advantages of TOPSIS includes its simplicity, ease of use, and rationality. It considers ordinal values from a decision maker without accounting for their prior experience, and requires less cognitive effort with the possibility of visualisation. Additionally, with increase in criteria or attributes, the number of steps to execute the method remains unchanged. The disadvantage includes ignoring possible correlation among the attributes. Despite this disadvantage, the solution provided by TOPSIS confirms with the correct solutions found by other MCDM techniques [255].

At first, a decision matrix ($X_{m\times n}$) is constructed allowing the comparison among the alternatives (see matrix below). Assessment is generally provided by a group of experts individually (with or without any interaction between them) in ordinal scales. Then it considers certain target points, namely *ideal positive solution*(I^+) and *ideal negative solution* (I^-) that models the best and the worst alternatives respectively.

Values of the distances to ideal positive and ideal negative solutions i.e. D_i^+ and D_i^- are used to define closeness coefficient for each alternative. Finally, to rank the alternatives, relative proximity/closeness co-efficient to the ideal solution is computed

$$\begin{bmatrix} X_{11} & X_{12} & \dots & \dots & X_{1j} & \dots & \dots & X_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{i1} & X_{i2} & \dots & \dots & X_{ij} & \dots & \dots & X_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ X_{m1} & X_{m2} & \dots & \dots & X_{mj} & \dots & \dots & X_{mn} \end{bmatrix} \tag{4.6}$$

$$\begin{matrix} \{I_1^+ & , & I_2^+ & , & \dots & \dots & I_j^+ & , & \dots & I_n^+\} \\ \{I_1^- & , & I_2^- & , & \dots & \dots & I_j^- & , & \dots & I_n^-\} \end{matrix}$$

for each of them as follows:

$$CC_i = \frac{D_i^+}{D_i^+ + D_i^-} \tag{4.7}$$

Note that all the values of the relative proximity are between 0 and 1. The best alternative should be the alternative with a relative proximity closest to zero.

Fuzzy-TOPSIS method uses fuzzy assessment of alternatives, and constructs a fuzzy decision matrix. Assessment is provided on a linguistic scale by a group of experts, individually. In order to convert their linguistic judgments into fuzzy values, the scale is converted into a fuzzy scale. For example, in the existing research, some authors have adopted a symmetric 5-point triangular fuzzy scales [111, 236], which can vary between 0 and 10 and can capture uncertainty inherent in experts' opinion (See Table.4.2). Note that the linguistic scale is an example scale and it may vary

Table 4.2: Linguistic scale for rating alternatives

Linguistic variable importance	Corresponding triangular fuzzy number
Very low	$(0, 1, 3)$
Low	$(1, 3, 5)$
Medium	$(3, 5, 7)$
High	$(5, 7, 9)$
Very high	$(7, 9, 10)$

Let L_{ij} be the $m \times n$ decision matrix where L_{ij} is the fuzzy assessment of an expert j over the alternative i refers to the alternative of a specific problem (See Matrix 4.8).

Without any previous aggregation, fuzzy normalization is performed (which is different than the real number's normalisation process). For each element in the matrix L_{ij}, the fuzzy value is represented by three elements such as l_{ij}, m_{ij}, u_{ij}. Since u_{ij} values are the biggest in the possibility of all triangular fuzzy numbers, we consider $c^* = u_{ij}$ as the maximum value to normalise [236] the matrix L. After this operation, we we obtain a normalised matrix R$_{ij}$, where $R_{ij} = [r_{ij}]_{m\times n}$

$$\begin{bmatrix} L_{11} & L_{12} & \dots & \dots & L_{1j} & \dots & \dots & L_{1n} \\ L_{21} & L_{22} & \dots & \dots & L_{2j} & \dots & \dots & L_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ L_{i1} & L_{i2} & \dots & \dots & L_{ij} & \dots & \dots & L_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ L_{m1} & L_{m2} & \dots & \dots & L_{mj} & \dots & \dots & L_{mn} \end{bmatrix} \tag{4.8}$$

$$r_{ij} = (\frac{l_{ij}}{c^*}, \frac{m_{ij}}{c^*}, \frac{u_{ij}}{c^*}) \tag{4.9}$$

where $1 \leq i \leq m$ & $1 \leq j \leq n$; r_{ij} are normalised triangular fuzzy numbers in matrix R_{ij}.

This method calculates fuzzy distance between *ideal positive solution* and *ideal negative solution* for instance, euclidean distance with fuzzy sets (see Eq. .4.3). The ideal positive and ideal negative solutions are defined considering the linguistic scales. The *fuzzy ideal positive solution* (FI$^+$) and the *fuzzy ideal negative solution* (FI$^-$) are defined by considering the respective maximum and minimum values of each expert: $r_j = (r_{1j}, r_{2j}, \dots r_{mj})$, where $r_j^+ = \max\{r_{ij}\}$, where i = 1,2,...,m. Similarly $r_j^- = \min\{r_{ij}\}$, where i = 1,2,... ,m. Note that, these calculations are conducted for each expert, separately.

For Manhattan geometry, distance between an alternative, and maximum (FI$^+$) and minimum (FI$^-$) numbers are calculated by the overlap of the distances. For the triangular fuzzy numbers, the distance is calculated by sum of their mod of the differences.

$$D_{M_i}^+ = \sum_{j=1}^{n} |r_j^+ - r_{ij}| \quad and \quad D_{M_i}^- = \sum_{j=1}^{n} |r_j^- - r_{ij}| \tag{4.10}$$

where $D_{M_i}^+$ *and* $D_{M_i}^-$ are the separation measures of each alternative that are calculated by Manhattan distances between triangular fuzzy points.

Similarly, for Euclidean geometry, distance between an alternative and fuzzy ideal positive solution i.e. FI$^+$ and fuzzy ideal negative solution i.e. FI$^-$ are calculated by square root of the sum of the square of the differences between them (similar to the Eq. 4.3 described in section 4.2.1).

$$D_{E_i}^+ = \sqrt{\frac{1}{3}\sum_{j=1}^{n}[(r_j^+ - r_{ij})^2]} \quad and \quad D_{E_i}^- = \sqrt{\frac{1}{3}\sum_{j=1}^{n}[(r_j^- - r_{ij})^2]} \tag{4.11}$$

where $D_{E_i}^+$ *and* $D_{E_i}^-$ are the separation measures of each alternative that are calculated by

Euclidean distances between triangular fuzzy points. This is a common method used by GDM research dealing with fuzzy numbers [45, 69].

Afterwards, for both the methods, rankings of alternatives are obtained by ordering the closeness coefficients (see Eq.4.7) of each alternative.

Proposed methodology: A combined FAHP and FTOPSIS method for ranking

In the extant literature, emphasis has been given to combine multiple MCDM techniques to overcome their individual shortcomings. In general, MCDM methods are utilised in various selection and ranking purposes and for some cases, they have been used in GDM purpose to access or rank alternatives [286]. TOPSIS is considered an important and efficient technique to rank alternatives with preferences of experts' but less efficient in calculating subjective weights because of the *numerosity of alternatives*. However, AHP is a well suited to compare pairwise criteria to capture subjective weights from experts. Hence, a combination of AHP and TOPSIS is considered a complimentary and powerful technique to obtain preference ordering with subjective weights of criteria. Our proposed methodology goes further in this direction and uses fuzzy-AHP and fuzzy-TOPSIS to obtain a hybrid GDM approach to rank factors according to their importance - assessed by a group of experts - for specific problem frameworks. The reason for combining fuzzy AHP with fuzzy TOPSIS instead of classical AHP is that when combining multiple MCDM methods, fuzzy AHP generates better result (in the combined techniques) as compared to AHP [137, 179]. Additionally, necessary conditions for application of TFN for AHP are clear judgement of criteria (higher consistent ratio), existence of criteria for equal importance, and dominant preference(s). Under such conditions, application of TFNs induces qualitatively different priorities than classical AHP [39]. All of the conditions were fulfilled by our study. Therefore, we chose fuzzy AHP instead of the classical AHP for the methodology of this study.

Initially, we consider the literature review to collect the most important and investigated factors. Motivation for a research problem can come from a real world phenomenon, but in-depth investigation of extant literature can provide sufficient information on the research problem. After analysing the information, factors of the problem can be identified and can be categorised into groups and sub-groups with reference to relevant extant literatures. Finally, an initial problem framework can be postulated that contains most of the important factors to evaluate the research problem (see part A of Fig.4.2).

Then, we consider utilising experts' knowledge to find the weights of the categories with fuzzy-AHP method to assess their importance. In doing so, we incorporate fuzzy logic to help translate tacit and imprecise knowledge of experts into precise usable knowledge. Calculating and incorporating subjective weights of categories from fuzzy-AHP (see part B of Fig.4.2) method is considered more reliable than traditional AHP calculation [174]. Subjective weights reflect decision makers' expertise and judgement under reliable conditions which aligns well with this study's objective i.e. to gather information from experts to obtain a parsimonious model.

Finally, to get the ranking of the factors, we consider fuzzy-TOPSIS method. Moreover, fuzzy-TOPSIS is considered one of the best decision making method for selection under conditions of conflict: for criteria and alternative selection [284] and for ranking in consideration with knowledge transfer issue i.e. from qualitative to quantitative. For the proposed methodology, information collected on the criteria/alternative is conducted with fuzzy-AHP method, and there exists a possibility of using these weights to multiply with each of the decision matrix's elements (see part C of Fig.4.2). Following the extant literature of group decision-making, multiplication of non-fuzzy weights to each TFNs in a matrix, results in a matrix that contains all fuzzy numbers ranging between $[0, 1]$. In doing so, we combine subjective weights to fuzzy evaluations of alternatives, resulting in a better selection/ranking method that captures uncertainty or impreciseness but reduces loss of information in the decision making process. Therefore, we have combined FTOPSIS (ranking/sorting) with FAHP (criteria weights) to overcome weaknesses from each of the techniques.

At the end of this process, we make the initial starting framework into a parsimonious framework, which can be tested with data in the future studies. The overall detailed steps of the proposed hybrid methodology is shown in the Figure 4.2.

4.3 Conceptual background: New product development and Adoption of innovation for Innovator Group customers

New product failure phenomenon is still persisting today. Even after the consorted efforts from the firms to include customers in the NPD process, the right customers have limited participation in the overall process. We intend to understand the failure phenomenon from the IG customer's perspective, since this group is the first to adopt among customers. Also, IG customers are responsible for diffusion of information to the later customers. This conceptual background is necessary to ground our understanding in NPD, adoption process and IG customers.

New product development

Multi-disciplinary research on new product development has been proliferate for decades, especially since the introduction of the Journal of Product Innovation Management in

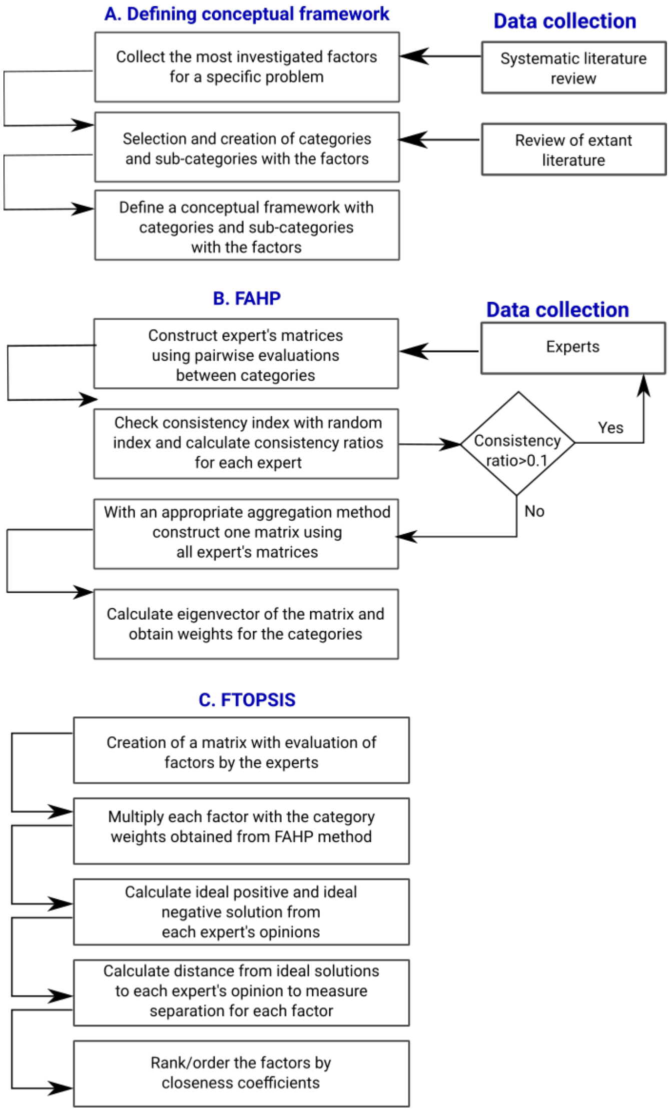


Figure 4.2: Proposed method of selecting and ranking main factors

1984. The nature of NPD process involves multi-disciplinary teams within an organisation to collaborate. Taking different perspectives into account, NPD process can be sub-divided into *Concept development or Ideation*, *Supply chain-design*, *Product design*, *Prototype testing* and *Product launch* [139].

For concept or idea development, the firms unite their internal technical and non-technical teams to brainstorm on new product ideas. However, with the exception of lead users [257, 261], other customers are rarely involved in this stage because they have been found less effective [43]. Lead users are recruited by firms to collaborate in screening [159] or pyramiding methods [165, 261] which depends on surveys that suffers from self-selection bias. Therefore, firms face challenges for identifying valuable lead users for radical new products for concept development phase [90, 165]. The *supply chain design* stage is production oriented where only technical professionals are involved. Similarly, the *Product design* stage is a technical and production heavy stage where depending upon the product, a firm's R&D team works in collaboration with other internal teams for product specific configurations.

Prototype testing is a latter stage in the overall NPD process where select customers are invited to try new products before its launch, for example film-critics and randomly chosen movie-goers for movie screening [38]; trial offers at retail stores [233]; virtual prototype testing for products [57]. On the other hand, by real-usage and experience, customer participants provide feedback which could be useful for firms to make amendments to their yet-to-be launched products. The last stage of the NPD process is *Product launch* where new products become accessible in the market to wider population. This is a critical stage for building customer awareness and recognition of the product's presence via marketing campaigns and activities such as mass media advertisement, influencer selection, sales team preparation, and brand community activities on social media.

For such a complex process, NPD research has placed higher importance on organisational factors (cross-functional team, team-communication, team-coordination, managerial perception), process factors (technological proficiency, marketing proficiency), strategy factors (speed-to-market) [76, 101], and less on customer related factors (market research, consumer preferences, customer selection). Customer-centric marketing has gained momentum since lead user's success [257]. For innovation driven firms, internal-teams and customers can be complimentary resources [185] because they bring different aspects of knowledge to product performance. Customer's need/preference recognition and alignment with offered products is crucial to reduce new product failure. However it is a challenging task to synchronise products with the changing customer preferences and needs. Additionally, not all customers are equally knowledgeable or valuable to the NPD process. Therefore, understanding adoption decision-making of the most valuable customers in new products (i.e. IG customers) is an important and non-trivial problem.

Adoption of innovation by IG customers

Diffusion of innovation process is the spread of an innovation or a novel idea, product, or practice in a market [203] which disseminates through four key elements, i.e. an innovation, communication channels, time of adoption, and social system [204]. Adoption time and adopters [203, 204] in a social system constitute the time and the space components respectively for the diffusion equation.

In the marketplace, information spreads through human interactions. Especially for product specific diffusion, it spreads when consumers start adopting (buying) new products or sharing information with other people (through word-of-mouth). Therefore, adoption of a product by customers accelerate its success rate and dis-adoption slows the rate. At the core of NPD process lies individual customers and their adoption decisions. Chronologically, the first customers to adopt new innovative products before the majority of people are known as Innovators and Early adopters (together they are known as Innovator group customers) [161]. These customers are knowledgable, influencers, and experts in specific domains, for example tech-enthusiasts or food enthusiasts are innovators/early adopters for specific product domains but may not be experts in both. Lead users are another group of customers who invent products to satisfy their specific needs in the event of unavailability of such products in the market [257]. However, involving lead users (or customers in general) in the developmental stage of NPD has no significant improvement [43] as compared to the firms' experts.

Considering the extant literature on Wisdom-of-crowd, few select experts provide better judgements than a crowd or a best expert [164]. As IG customers are the expert-consumers, their knowledge could be utilized to generate new product ideas, or to assess prototypes. By distribution of information through word-of-mouth, status signalling or high-consumption indicators, these customers help a product's diffusion in the market.

4.4 Real case implementation: Selection and ranking of main influencing factors on IG customers

4.4.1 Data collection

For the purpose of selecting and ranking the main influencing factors on IG customers, we have conducted two rigorous data collection process: a systematic literature review to collect most investigated factors, and surveys addressed to industry experts to find our their opinion on the importance of these factors (see section B and C of Fig.4.2). By doing so, we gathered information from both the theory and the practice.

Systematic literature review

A systematic literature review was conducted to collect and synthesize the adoption decision-making factors for the IG customers. The selected articles reviewed for this study were analysed using content analysis because of its objective coding scheme that truncates data and makes it comparative to other classifications and analyses [166, 222]. In addition, systematic literature review is a widely accepted method for literature reviews [181].

We searched SCOPUS (Science-Direct) database since it is the largest full-text multidisciplinary academic database, in parallel to Google Scholar, Web of Science, and Business Source Premier. The primary search terms were a combination of “early adopters”, “innovators”, “customers”, and “retail”. The principal reason for adding retail in the search was to include all consumer products and services, and to exclude industrial products.

Additionally, we limited our discussion to studies that measure customer adoption in NPD and new-product performance, making customer adoption our unit of analysis. We selected articles about customer adoption of new products (rather than firm/industry NPD strategy adoption) between 1988 to October 2017 because we wanted to capture NPD research evolution after the launch of Journal of Product Innovation Management (1984). Due to the growing consensus among NPD researchers about all consumer products and services are to be considered as *products* [101, 138, 139], we did not exclude services from our search. In order to gain comprehensive insights, we included multidisciplinary academic articles from “decision sciences”, “computer science”, “business, management, and accounting”, “economics, econometrics, and finance”, and “social science”. Finally, we restricted our search to published peer-reviewed academic articles in support of our aim to include quality, proven, and value adding knowledge materials.

After employing the search words in titles, abstracts or keywords in the databases, an initial findings of 276 articles were obtained. After discarding articles for missing names, different level of analysis, a total of 112 articles were identified for full-article reading. Thereafter, seven new articles were included from three editorial articles, increasing the total selection to 116 (by removing editorial articles). After comprehensive reading, 72 relevant articles (see Table 3.7) were selected from a total of 283 initial articles, with a 25% acceptance rate.

A total of 103 investigated factors were collectively studied in the 72 articles (see Table 3.7) and some of the articles have studied more than one of these factors. According to the nature of the factors, we categorised them into four major categories: *individual traits*, *product traits*, *environment*, and *relationship between product and individual traits*.

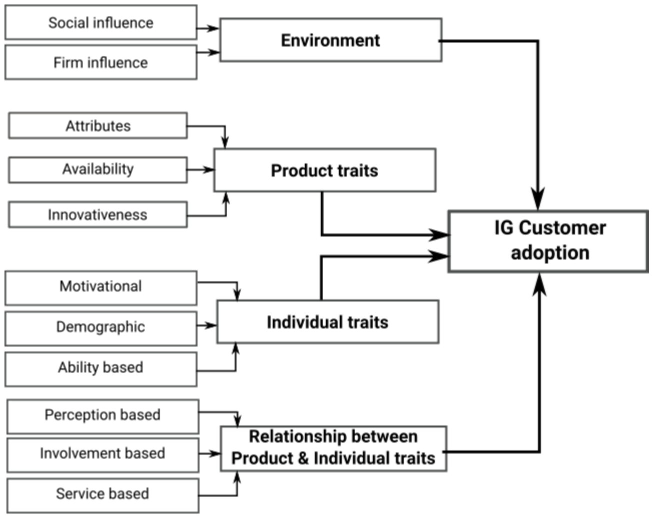


Figure 4.3: Conceptual framework of influencing factors on IG customer

The relationship between these categories helped us form the conceptual framework for the study. Product traits and individual traits factors interact among themselves under the influence of social and (brand) firm influence. The interactions between the new products and the customers form perception, involvement and service based influences that ultimately results in customer's adoption/non-adoption decision. Previous research has investigated factors separately and has not emphasised on the importance of each factor/category for the decision making process. We focus on finding relative importance of each adoption factor in order to comprehend the decision making of IG customers better than the piecemeal way.

Table 4.3: Summary of influencing factors on IG customer decision-making

Category	Unique factors	% Share of factors	Articles	% Share of articles
Environment	13	13%	34	10%
Product traits	21	20%	44	13%
Individual traits	37	36%	124	37%
Relationship between individual & product traits	32	31%	129	39%
Total	103	100%	331	100%

Individual traits. Based on the socio-psychological theories, the individual trait category is formed by factors that operate at the individual level. Personal innovativeness is one of the key distinguishing characteristics of IG customers, and it depends on a customer's degree of innovativeness [6, 156, 188, 241]. Other key individual traits include domain specific innovativeness, risk-taking attitude, knowledge sharing attitude, self-motivation, do-it-yourself attitude, information-seeking attitude, hedonistic attitude, expertise, and inquisitiveness. Demographic factors such as income, age, education, gender and life event also influence the adoption decisions for these customers. A combination of these individual traits portrays IG customers as opinion leaders and experts in specific areas (domains), who willingly share their knowledge with social communities and exhibit a strong inquisitiveness towards new products.

Product traits. Research in marketing has helped us categorize factors that belong to product trait category. Adoption literature emphasizes product characteristics as a major component for the adoption of innovation [175, 205] and primarily emphasizes product attributes, product advantage, and additional new product features. The use of prototypes, trial offers, availability of products at retail stores (both online and offline), availability of choices, and quantity of products are among the key factors that encourage IG customers to try new products.

Environment. Researchers grounded in sociological perspectives have examined the social aspects of customers' adoption process. The knowledge from this research has helped us categorise the factors under this category. In general, the factors belonging to environment category operate externally to individual customers and are socially embedded. Multiple sociological theories have explained the implicit social impact on individuals with elaborated mechanisms for accumulating social capital [20, 122, 130, 154]. Mass-media/brand advertisements, social norms, social images, subjective norms, values, cultural norms, social approval, and biases are some of the important environmental factors for IG customers. Some macro-level factors like regulations and commerce treaties may indirectly impact individual customers but they go beyond the scope of this study.

Relationship between product and individual traits (RPIT). From multi-disciplinary perspectives, factors grounded on perception, involvement and satisfaction were included in the RPIT category. Perception is a multi-level construct that is based on an individual's experiences, personal characteristics and social influences. Innovation of new products is subjectively judged by customers and the degree of involvement is conditional on the perception from previous experience [283]. Involvement may be higher in customers for their preferred firms/brands [5, 31]. Similarly, perceived benefits or advantages, price (sensitivity) and comparative value are perception based which differs from what firms/manufacturers envision. Trust is a pre-requisite for exploring innovation and newness, and it plays a big role in forming early willingness to purchase new products [16, 167]. Satisfaction is formed by actual product usage and expectation is based on past satisfaction level [153]. Overall, brand experience is based on interaction with products at various touch points.

Expert knowledge

The core group of experts for the study consisted of Marketing Directors and Senior Brand Managers - who have extensive and varied experience in new-product launches. Previous literature in GDM has suggested that if the expertise/knowledge of experts is higher, then a smaller group of experts are sufficient for decision making problems, as compared to a larger group of experts with less knowledge [249]. Hence we decided to gather a small and highly knowledgable group of experts who have extensive experience in new product launches. To gain access to these experts, we approached ESADE's "Group Trade-marketing Retail forum" that convenes bimonthly where we presented our project. All nine attendees agreed to participate in our study. At the same time, we prepared and sent an online survey to 102 marketing managers whom we approached via LinkedIn. These selected experts have on average of 10 years experience in new product launches in Fast Moving Consumer Goods (FMCG) sector.

We realised that description of the factors were needed at the beginning of the survey. For the fuzzy-TOPSIS section, we presented 75 factors for experts' final evaluation. Moreover we realised that the experts may have different confidence regarding each factor's importance while answering the survey with the linguistic scale (see Table 4.2). To accommodate hesitancy in experts' opinion, we provided more than one option for each factor. For e.g. if an expert had doubt whether a factor's importance was very high or high for adoption decision, then s/he could choose both the options. In the fuzzy-AHP section, experts were asked to do only six comparisons making the task easier to perform.

At the end of the process, we have managed to collect 7 completed questionnaire from the ESADE forum participants and 9 from the online survey. In total, we collected information from 16 experts.

4.4.2 Fuzzy AHP to assess category weights

The hierarchical structure of the proposed decision model of the study shows the levels in the decision problem (see Fig. 4.3). Defining IG customer's adoption decision-making framework fulfils the complexity of a multi-criteria and multi-person decision making problem, despite being originated in a marketing discipline.

Meeting our expectations, fuzzy-AHP technique was easy to comprehend and execute for the experts without having prior experience in participating in the fuzzy-AHP method. The necessary conditions for application of fuzzy-AHP (TFN) of clear criteria, existence of criteria for equal importance, and dominant preference(s) are fulfilled in our study [39]. Therefore, under these conditions, application of TFNs may induce qualitatively different priorities than classical AHP, and justifies our choice for Fuzzy AHP. After the collection of information from the experts, we converted their responses to triangular fuzzy numbers and created individual matrices. The following section describes the calculation steps in detail:

At first, we collected information from 16 experts and created individual matrices. We have used fuzzy linguistic scale with TFN (as explained in section 4.2.1) because in the relative scale, the modal values represent the approximated intended judgements of the experts. We chose this particular linguistic scale Table 4.1 as a simplifying ansatz and a common choice among researchers in the MCDM field. Additionally, no experts showed hesitancy during the evaluation process making the input values of matrix easier to work with. Then, we checked consistency index of each expert with an internal logic. We came across two experts whose answers failed to pass the logical test. We contacted them and requested to reconsider their answers and they returned logically consisted answer sets.

Next, we aggregated all individual matrices into one matrix M by the arithmetic mean method and we realised that geometric mean renders similar result i.e. difference in weights were inconsiderable. We decided to use arithmetic mean for the final matrix calculation (see matrix in Table 4.4). In our study, each expert is assumed to have similar expertise in the consumer goods industry and they didn't interact among themselves while participating, as they belonged to separate firms. Therefore, the aggregator falls under the *independent* category.

Table 4.4: Combined FAHP matrix (M) with all experts' evaluations

$$\begin{bmatrix} (1.0, 1.0, 1.0) & (2.796, 4.061, 5.217) & (3.266, 5.021, 6.783) & (1.362, 2.502, 3.655) \\ (1.706, 2.487, 3.351) & (1.0, 1.0, 1.0) & (3.016, 4.771, 6.471) & (1.420, 2.684, 3.963) \\ (0.620, 0.921, 1.350) & (0.627, 0.935, 1.405) & (1.0, 1.0, 1.0) & (0.765, 1.412, 2.080) \\ (2.282, 3.210, 4.254) & (1.857, 2.668, 3.629) & (3.491, 4.896, 6.346) & (1.0, 1.0, 1.0)] \end{bmatrix}$$

The calculation of priority vector for matrix M was conducted following *synthetic extension value* technique proposed by Chang [42] (see section 4.2.1). The synthetic extension was conducted on each element of the matrix that comprised of TFNs i.e. in order to find synthetic extensions (see Eq. 4.2 & section 4.2.1). After the synthetic extension analysis, we obtained six values of degree of similarity which were also TFNs. We have utilised the *min* function to select the smallest TFN (see Eq. 4.2.2) in this chapter.

Then, for each category, we calculated d_{s_i} and these values corresponds to the weights of each category. At this stage, the values were not fuzzy numbers and the weights were not normalised. Therefore, we normalised to get the final weights for the four categories as shown in Table.4.5.

Table 4.5: AHP weight calculation with Arithmetic mean

Category	Weights (Arithmetic mean method)
Individual trait	0.327
Product trait	0.295
Environment	0.067
Relationship between Product and Individual	0.311

The final weights of the categories indicate that Individual factors and Relationship formed between products and customers are the two most important categories. These combined factors influence the adoption decision making of IG customers more than either Environmental or Product Trait factors.

4.4.3 Fuzzy TOPSIS for ranking influencing factors of IG customers

The large number of factors collected from the systematic literature review created an initial problem of over-representation and a lack of parsimony. Therefore, ranking according to their importance was required to gain requisite knowledge on IG customers.

A linguistic scale was deployed to collect information from the experts and it was observed that most of the experts did not show any uncertainty while answering on the importance of 75 factors. However, two experts showed hesitancy in one of their answers out of 75 (2 in 1200 or 0.16% of all answers). For calculation, we converted their answers to their respective lower values. Similarly, when experts failed to provide any answer, we converted it to "N" for the calculation. A judgement matrix was created where names of all factors were placed on the row-heads; experts were placed on column-heads and the values collected from experts were filled inside the matrix (see Table 4.6 and Table 4.9). Afterwards, the collected information on linguistic scale for each factor was converted to their corresponding TFNs (see Table 4.2) and filled into a fuzzy matrix L.

Next, the fuzzy matrix L was normalised for the purpose of making all the fuzzy numbers range between $[0, 1]$. Following the best practices of MCDM and GDM literatures [45], we normalised according to the formula shown in equation 4.9 and obtained the normalised matrix R. Then, we multiplied the weights to each element of matrix R with the corresponding category weight (see Table 4.5) similar to the method suggested by Sun *et al.* [236]. The weights were incorporated into the decision matrix to get each factor's relative importance.

The next step was to find fuzzy ideal positive solution (FIPS) and fuzzy ideal negative solution (FINS) for each expert (see section 4.2.2). This step was performed by analysing all the values provided by each expert and selecting the maximum and the minimum value from them.

Afterwards, euclidean distance between each element and the FIPS and FINS were calculated for each expert and denoted as d^{+}_{ji} and d^{-}_{ji} respectively. In completion of the process, we calculated the combined distances for each factor i.e. $D^{+}_{e_i}$ and $D^{-}_{e_i}$ values that were obtained by applying Equation 4.11. These values are known as separation measures and these measures i.e. $D^{+}_{m_i}$ and $D^{-}_{m_i}$ were calculated for "Manhattan distance" as well (see Eq. 4.10).

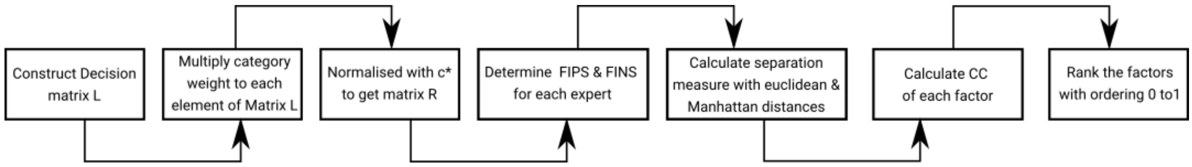


Figure 4.4: Visual representation of steps used for Fuzzy-TOPSIS technique

Then, closeness coefficient for each factor was calculated by following the equation 4.7. The preference order of each factor is arranged from 0 to 1 with proximity to zero indicates a higher ranking. The detailed results from both the distances are shown in Table 5.5 and 4.8. Hence, by utilising fuzzy-TOPSIS method, we calculated the ranking for each influencing factor, and the entire process of the calculation is shown in the Figure 4.4.

Table 4.6: Decision input in linguistic scale from the Marketing experts

Variables	Experts															
	E_1	E_2	E_3	E_4	E_5	E_6	E_7	E_8	E_9	E_{10}	E_{11}	E_{12}	E_{13}	E_{14}	E_{15}	E_{16}
Availability of product choice	V	V	E	E	M	E	E	S	S	M	V	M	V	S	V	E
Innovativeness of product	E	V	V	M	E	V	M	E	M	E	E	E	E	E	V	E
Financing	M	N	N	N	N	M	N	N	M	V	M	S	S	N	M	E
Relative advantage	E	M	E	V	E	E	V	M	V	M	E	V	M	V	V	E
Trialability	V	S	E	M	V	E	E	V	N	E	E	M	V	V	E	E
Quantity	M	M	V	M	V	V	M	S	N	S	N	M	S	N	M	E
Speed of use	M	M	V	M	M	M	S	E	S	E	M	V	V	V	M	E
Technical specifications	M	M	V	V	M	S	M	S	S	V	M	V	S	M	M	E
Store distance	M	V	V	V	S	M	V	M	S	M	V	V	V	S	V	E
Functionality	V	M	E	E	V	E	E	E	M	V	E	V	V	V	V	V
Observability	M	V	V	E	M	V	E	V	S	V	E	V	E	M	V	V
Technology generation	V	M	V	N	M	M	M	V	S	V	E	M	V	V	V	V
Product type	V	M	V	E	N	V	V	E	M	M	E	M	M	E	M	V
Product size	V	V	M	E	M	V	M	M	N	S	S	M	M	M	V	V
Visual appeal	E	V	E	E	E	E	E	V	S	E	V	V	V	E	E	V
Personalized offering	V	S	E	S	M	M	S	V	V	V	M	S	V	V	M	V
Incentives	V	S	M	E	S	V	E	S	S	M	V	S	M	S	M	V
Prototype	M	N	N	N	N	V	N	M	M	M	S	S	V	S	M	V
Innovativeness of technology	E	N	N	N	N	S	S	E	V	V	V	V	V	V	M	V

Age	S	S	V	E	V	E	S	E	S	M	E	M	V	V	V	V
Income	M	V	M	E	E	E	V	S	V	V	E	V	V	V	E	V
Education	N	M	S	M	S	V	S	M	S	E	M	V	M	V	E	V
Gender	M	M	S	M	V	E	M	S	S	M	N	N	N	N	S	V
Life event	M	S	M	V	M	V	S	V	M	E	E	N	M	N	S	V
Personal Innovativeness	V	V	E	E	M	V	V	V	V	V	E	E	V	E	V	V
Information seeking attitude	M	V	V	M	M	V	E	M	M	M	V	V	M	V	V	V
Knowledge sharing attitude	V	V	V	E	S	V	E	M	V	V	S	E	V	V	E	V
Do-it-yourself attitude	N	M	N	S	N	V	S	V	S	S	M	N	M	M	M	V
Hedonistic attitude	E	V	V	M	E	E	E	V	S	V	M	M	V	V	M	V
Risk-taking attitude	M	M	M	E	N	V	M	S	V	V	V	V	V	V	E	E
Variety seeking attitude	V	E	V	V	E	E	V	S	M	E	V	E	V	M	E	E
Novelty seeking attitude	V	E	E	V	M	E	V	S	M	E	E	E	V	V	E	E
Convenience seeking attitude	M	E	V	M	M	E	E	V	S	S	V	M	M	M	M	E
Loyalty	N	V	V	E	M	E	M	S	M	M	V	S	M	S	M	E
Absorptive capacity	V	S	E	V	N	V	M	E	S	M	S	V	M	V	M	E
Strong individuality	M	M	E	N	N	M	M	E	M	M	N	V	M	V	V	E
Inquisitiveness	S	E	V	S	S	V	S	M	N	S	M	M	M	M	M	V
Social Image	M	V	E	E	V	M	S	V	V	E	V	E	M	V	E	V
Expertise	M	M	V	M	S	M	S	M	V	M	V	V	V	V	V	V
Status signalling	N	M	E	E	V	V	S	V	V	E	V	V	M	M	V	V
Learning	S	M	V	V	N	M	S	S	M	M	M	M	M	V	M	V
Subjective norms	S	S	V	V	N	M	M	M	S	S	M	M	V	M	V	V
Media dependency	V	V	E	E	S	S	N	S	M	E	M	M	M	S	V	V
Impulsiveness	M	E	E	E	M	E	E	M	S	V	M	E	M	M	V	V
Aesthetic value	M	V	V	V	M	M	V	V	M	E	V	M	M	V	V	V
Materialism	N	V	E	M	S	S	S	M	M	V	S	V	M	V	M	V
Self-awareness	M	S	M	E	N	M	S	V	V	S	M	V	V	V	V	V
Empathetic behaviour	M	V	V	M	N	S	S	M	V	V	M	V	M	M	E	V
Resistance to mass media	V	S	M	N	N	M	N	E	M	N	M	S	M	M	M	V
Values	M	M	M	V	S	V	M	M	V	M	M	M	M	V	M	V
Self-monitoring	V	V	V	V	N	V	M	M	S	M	M	M	M	M	E	E
Social and Cultural norm	M	V	E	E	S	E	M	E	V	E	M	E	V	S	V	E
Network externality	V	V	V	V	M	V	M	M	S	E	E	V	V	V	E	V
Mass media influence	E	V	V	E	E	E	E	N	M	V	M	M	S	M	E	V
No. of initial adopters	V	M	V	M	S	V	S	M	S	M	S	V	V	S	M	V
Industry standard	E	V	M	V	V	V	V	S	N	S	N	M	M	S	M	V
Urbanization	S	S	M	E	M	V	V	E	M	M	S	V	M	V	M	V
Social approval	M	V	V	M	V	E	M	S	V	E	M	V	E	E	M	V
Perceived usefulness	V	E	E	V	M	E	E	E	V	M	E	V	E	E	V	V
Perceived ease of use	V	E	V	E	V	E	E	E	S	V	E	S	E	V	E	V
Perceived Brand image	E	E	E	E	E	E	E	M	V	E	E	V	V	V	E	V
Previous experience	M	V	V	M	E	E	E	S	S	S	M	V	V	E	V	V
Perceived risk	M	S	M	V	S	M	M	S	V	S	M	V	M	M	V	V
Perceived compatibility	V	S	M	E	M	E	M	M	M	V	M	M	M	S	V	V
Involvement	V	E	E	V	M	V	E	V	V	E	M	E	M	V	V	V
Expectations	E	E	E	E	M	E	E	V	M	E	M	V	M	E	E	V

Satisfaction	V	E	E	E	M	V	V	M	V	V	M	V	M	E	E	V
Perceived complexity of products	S	M	M	E	S	V	M	S	M	M	M	M	V	V	M	V
Perceived (guidance) Information	M	V	V	E	S	M	V	M	S	M	S	S	M	V	V	V
Privacy concerns	V	N	M	N	N	M	S	M	S	M	N	S	M	V	V	V
Adoption duration time	V	M	V	N	N	M	S	M	S	M	S	M	M	M	V	V
Search time	M	V	V	N	V	V	V	M	M	V	M	S	M	M	V	V
Evaluations	V	N	M	E	S	V	V	M	M	E	M	S	M	M	E	V
Perceived brand trust	E	E	E	E	E	V	E	V	V	E	E	S	V	V	E	V
Security concerns	N	N	N	E	M	V	M	M	V	M	S	N	M	E	V	V

4.4.4 Results

The final results from the hybrid methodology of fuzzy-AHP and fuzzy-TOPSIS shows that Relationship between Product and Individual traits dominates the top rankings among the adoption factors. Perception about brand image and brand trust, usefulness of product, relative advantage of product and expectation from new product are crucial for IG customers while making adoption decisions. Product's visual appeal, innovativeness and attributes determine the level of acceptance by the IG customers who are domain specific experts and can judge the products minutely. Personal innovativeness, novelty seeking attitude and variety seeking attitude drive IG customers to seek new products among the individual traits. This indicates that IG customers' perception, involvement, and interaction with products creates a strong influence during their adoption of new products.

The result from our proposed methodology reveals that the top 30 factors are common to both euclidean and Manhattan distances (see Table 5.5, 4.8) but they differ in their ordering. This confirms the validity of the variables which are ranked higher in their importance to the decision-making of the Innovator group customers. One of the critical findings of this study is the emergence of RPIT as an influential category which is more impactful than either product or environmental category (see Table 4.3). RPIT is a new category to our knowledge and has not been proposed as a separate category in any previous study. For complete list of the factors and their ranking, please refer Appendix Table 4.10 and 4.11.

4.5 General discussion

Selecting the most important factors for IG customers' adoption decision-making is an important step towards understanding the new product failure phenomenon in consumer goods sector. With an intricate knowledge on the influencing factors, firms can design, develop and promote their products according to the customers' preferences.

Table 4.7: Ranking of IG customer's adoption factors with FTOPSIS (Euclidean distance)

Category	Variable_name	Closeness coefficient	Rank (Preference order)
RPIT	Perceived Brand image	0.0991	1
RPIT	Perceived brand trust	0.1337	2
Product Trait	Visual appeal	0.1340	3
RPIT	Perceived usefulness	0.1496	4
Product Trait	Innovativeness of product	0.1521	5
RPIT	Expectations	0.1705	6
Individual Trait	Personal Innovativeness	0.1794	7
Product Trait	Attributes of product	0.1819	8
Individual Trait	Novelty seeking attitude	0.1865	9
RPIT	Perceived ease of use	0.1869	10
Individual Trait	Variety seeking	0.2025	11
RPIT	Involvement	0.2160	12
Product Trait	Relative advantage	0.2182	13
RPIT	Satisfaction while comparing products	0.2343	14
Individual Trait	Income	0.2343	15
Individual Trait	Social Image	0.2528	16
Product Trait	Trialability	0.2552	17
Environment	Social norm	0.2575	18
Individual Trait	Knowledge sharing	0.2688	19
Product Trait	Observability	0.2690	20
Environment	Network externality	0.2849	21
Individual Trait	Pleasure seeking attitude	0.2873	22
Individual Trait	Impulsiveness	0.2921	23
Environment	Social approval	0.3054	24
Environment	Mass media influence	0.3078	25
Product Trait	Availability of product choice	0.3264	26
Individual Trait	Aesthetic value	0.3349	27
Individual Trait	Status signalling	0.3371	28
Product Trait	Product type	0.3397	29
RPIT	Previous experience	0.3425	30

The combination of extant literature's investigated factors and extraction of FMCG experts' knowledge with fuzzy set theory can become the foundation for a rigour required for this kind of research.

Contrary to the previously held believe that *Product Trait* and *Individual Trait* categories contain the most crucial factors for IG customers' adoption decisions, the results show that relationship formed between product and individuals is the most critical category (five out of the top ten factors belong to RPIT). Since RPIT is critical for IG customers, focusing on building relationships between new products and IG customers may accelerate the rate of adoption for new products. This finding alone can motivate managers to emphasize on key pre-launch activities such as trial offers, feedback sessions, idea generation for new (related) category, brand community activities with discussions to name a few.

The final ranking indicates that *Perceived brand image*, *visual appeal of products*, and *perceived brand trust* are the top three influencing factors for adoption of IG customers. Also

Table 4.8: Ranking of IG adoption factors with FTOPSIS (Manhattan distance)

Category	Variable_name	Closeness coefficient	Rank (Preference order)
RPIT	Perceived Brand image	0.0969	1
Product Trait	Visual appeal	0.1313	2
RPIT	Perceived brand trust	0.1313	3
RPIT	Perceived usefulness/Benefits	0.1469	4
Product Trait	Innovativeness of product/service	0.1500	5
RPIT	Expectations	0.1688	6
Individual Trait	Personal Innovativeness	0.1750	7
Product Trait	Attributes of product/Functionality	0.1781	8
RPIT	Perceived ease of use	0.1844	9
Individual Trait	Novelty seeking	0.1844	10
Individual Trait	Variety seeking	0.2000	11
RPIT	Involvement	0.2125	12
Product Trait	Relative advantage/Usability	0.2156	13
RPIT	Satisfaction while comparing products	0.2313	14
Individual Trait	Income	0.2313	15
Individual Trait	Social Image	0.2500	16
Product Trait	Trialability	0.2531	17
Environmental	Social and Cultural norm	0.2563	18
Individual Trait	Knowledge sharing	0.2656	19
Product Trait	Observability	0.2656	20
Environmental	Network externality	0.2813	21
Individual Trait	Pleasure seeking or Hedonistic	0.2844	22
Individual Trait	Impulsiveness	0.2906	23
Environmental	Social approval	0.3031	24
Environmental	Mass media influence	0.3063	25
Product Trait	Availability of product choice	0.3250	26
Individual Trait	Aesthetic value	0.3313	27
Individual Trait	Status signalling	0.3344	28
Product Trait	Product type	0.3375	29
RPIT	Previous experience	0.3406	30

perception of brands, products' usefulness, expectation, and relative advantage in customer's eyes play significant role in the overall adoption of new products. This indicates that firms need to build their brand image and brand trust; to make new products that are perceived as easy-to-use; to create opportunities for customer-product interactions with new products i.e. trial offers, in-store testing etc. On the other hand, customers can expect more from these new products and derive satisfaction as they try and use before purchasing, while simultaneously comparing with other products. The perception that customers create/build with the brands/products can help them decide on their future purchases of new products. This indicates that for "not-new-to-world"/incremental products, stronger bond formation with firms/brands may propel new products' adoption among the IG customers.

Nonetheless, the largest category impacting these early adopters, with maximum number of factors, is the *Individual Trait* category (ten out of top thirty factors). This indicates that attitudinal and demographic characteristics of IG customers play a greater role in whether the customer decides to purchase a new product or not. Personal innovativeness, novelty seeking attitude, variety seeking attitude, income, pleasure seeking attitude, impulsiveness, and aesthetics heavily influence IG customers' decision on adoption. For high-priced and/or radical products, status signalling by using the new products also motivates IG customers to adopt new products and influence the late adopters.

Visual appeal of new products attracts the attention of IG customers among all the product related attributes. Innovativeness of products play an important role in these customers' decision making as new products not only attract visually but also entice them with their newness. Attributes of products which are new and presents a different value to customers comes third in the *Product Trait* category of variables. Relative advantage over other products and the chance to try the new products before buying them also helps IG customers making up their minds. Availability of product at the point of sale after product's launch reinforces IG customer's desire to seek and access the new products. Also observing the new products being used and benefits derived by others have effect on customers though we are not sure of the measured direction (whether positive or negative). As IG customers are the first to buy, the explanation we can muster is that IG customers' derive satisfaction by purchasing earlier and sharing information on products among their social networks. Hence, satisfaction from previous sharing may motivate them to diffuse product information again.

4.5.1 Theoretical implications

Our major contribution of the study is highlighting the ranking of IG customer's adoption decision-making factors and their importance in the NPD process, in particular for three phases: *idea generation*, *prototype testing* and *launch* phases. For idea generation, customer participation has proven to be efficient for select customers (Lead users or IG) and understanding the underlying needs/preferences helps create new products. From our ranking of adoption factors for IG customers, customer participation activity in NPD can be planned with more focus on exploiting perceptual and individual trait factors as a starting guideline to generate useful information. In doing so, we contribute to customer participation literature.

We contribute to prototype testing phase by highlighting the most critical of product related factors such as visual appeal, innovativeness of product, new attributes, relative advantage and chance to trial before launch. We also contribute to launch phase by highlighting key factors such as perceived brand image, brand trust, perceived usefulness and innovativeness of products that most affect IG customer's adoption. Therefore, we contribute to adoption of innovation literature by providing the ranking of most influencing factors for IG customers.

Second, we have showed that a hybrid group decision making methods' can be used to a real world problem where the methods are adapted to any specific problem i.e. not retrofitting the methods to the problem. Our contribution lies in the adaptability of fuzzy- TOPSIS and AHP method to solve a multi-criteria group decision-making problem with the knowledge of the industry (marketing) experts. We provide a different approach i.e. fuzzy logic based group decision making methods to identify IG customers' adoption factors with industry experts' knowledge. We have demonstrated that operation research methods can be effectively used in the marketing domain while capturing human expertise that is embedded in uncertainty and tacit knowledge.

Additionally, we have demonstrated that for group decision making scenarios, where experts don't interact among themselves (in our study, experts belonged to separate firms), hybrid techniques can be successfully applied for selection and ranking. It ensures that result can be obtained without risking forced consensus or lowering the quality of decision even with divergent opinions [9]. Hence, we contribute to the group decision making literature.

4.5.2 Managerial implications

In our study, we aim to provide paths for managers to take micro-level actions [270] grounded in theory-based conceptual framework to manage NPD performance. Our study guides managers in planning and executing concerted actions towards attracting IG customers. Managers can design their new products on the basis of i.e. how IG customers perceive usefulness, innovativeness, attributes and ease of use, and how IG customers get influenced by visual appeal of products, trial offers, variety in products and hedonistic aspect of products. These factors together can provide a guideline for managers to design their products to be attractive, fun, with variety and easy to use. In addition, customers' participation activity in the NPD process can be purposefully targeted towards understanding perception about products and can help designing the new products according to the important individual traits and perceptual factors. These factors can guide NPD managers on strategic decision for developing prototypes that are developed keeping customer's adoption behaviour in mind.

IG customers' adoption decision making is also influenced by how they perceive brands (positively or negatively) and that affects the likelihood of their purchase from the brand. We suggest that firms need to build trust and brand image with new attributes, innovation and perceived usefulness of products, We also suggest firms to design marketing campaigns and brand communications to attract these domain specific expert customers (i.e. IG) who like to explore new products through variety and innovativeness. NPD Managers can also provide opportunities to IG customers to try new products and get their feedback on perceived usefulness, ease of use, expectation, satisfaction, functionality, and hedonistic/utilitarian aspects. The collected information can be insightful to the managers for learning about new products' actual reception in advance and can help them take necessary steps to improve prototypes or communication strategy before launching new products in the market.

For post-launch success, IG customers are proven to be opinion leaders and influencers who disseminate information on product and signal low risk with their high usage of the new products. These customers signal social status through tacit knowledge to late adopters. Managers should recognise IG customers' role as expert influencers and incorporate communication strategies to attract them to new products. Notwithstanding which way IG customers influence - imitation or risk mitigation - managers should utilise these customers' information sharing capability, word-of-mouth, and high status in the social network to disseminate positive product information. However, managers should be aware of the interplay between social network and mass media communication for new product diffusion. For example, it is observed that increased advertisement reduces online-word-of-mouth [79]. Therefore, managers should take a balanced approach to employ multiple tools to induce diffusion of new products with the IG customers.

4.5.3 Limitations and Future research

Despite our best efforts, the study has some limitations. First, the number of experts selected for the hybrid FAHP-FTOPSIS method was small. Though the diversity in location, nationality and experience mitigates some of the shortcomings but a larger number of experts could have presented better results. Second, all the experts belonged to FMCG sector in the study which may have overrepresented the sector's idiosyncrasies. However, FMCG firms operate in a highly competitive environment and the managers may have more knowledge on new product launches than other less competitive sectors. Additionally, comparative analysis of alternative combination of techniques is beyond the scope of this study.

For future researchers, application of MCDM, GDM and several other operational research methods into marketing research could be an interesting way forward, following our initiative. Since, there is no unique way to solve a complex multi-level organisational problem - where human expertise is imperative in solving but is often accompanied by ambiguity and uncertainty - fuzzy logic based group decision-making methods can be utilised for solving future organisational problems in marketing and strategy. Based on our result, real world applications for new product development - ideation, prototype testing and launch - can be conducted with firm level data. Additionally, in the future, validation of our framework can be conducted with customer data. Furthermore, future researchers can apply a combination of GDM techniques where expert knowledge is required without direct interactions among them. Some possible application cases include, ranking of top restaurants by diverse food experts, selecting awards for films by film critics, open source community consensus on new standard library and ranking of educational institutions to name a few.

4.6 Conclusion

Consumer goods sector's high rate of new product failure indicates that there is a gap in understanding about the phenomenon among academics and practitioners. We have proposed a framework grounded on customer adoption perspective i.e. from IG customer's adoption decision making to address the issue. Additionally, we have selected and ranked the most important factors affecting IG customers' with a hybrid fuzzy group decision-making technique that utilises knowledge extracted from industry experts and extant literature. Results reveal that perceived brand image, perceived brand trust, visual appeal, perceived usefulness and innovativeness of product affect the IG customer's the most. With this knowledge, managers can plan, design and launch new products with this insight on customer's decision behaviour. We have also proposed to involve IG customers at idea generation, prototype and launch phases of NPD to attract them to new products, to create products that these customers are likely to purchase, and to nudge them to disseminate product information in the market, respectively. Our finding shows future directions to researchers to further investigate new product failure phenomenon from customer's adoption perspective with real data and with fuzzy logic based methods.

Appendix

Table 4.9: Linguistic scale conversion to codes

Linguistic variable importance	Codes for conversion	Corresponding triangular fuzzy number
Very low	N	(0, 1, 3)
Low	S	(1, 3, 5)
Medium	M	(3, 5, 7)
High	E	(5, 7, 9)
Very high	V	(7, 9, 10)

Table 4.10: Detailed Ranking of factors with FTOPSIS-Euclidean distance

Category	Variable_name	Closeness Coefficient	Rank
RPIT	Perceived Brand image	0.0991	1
RPIT	Perceived brand trust	0.1337	2
Product Trait	Visual appeal	0.1340	3
RPIT	Perceived usefulness	0.1496	4
Product Trait	Innovativeness of product	0.1521	5
RPIT	Expectations	0.1705	6
Individual Trait	Personal Innovativeness	0.1794	7
Product Trait	Attributes of product	0.1819	8
Individual Trait	Novelty seeking attitude	0.1865	9
RPIT	Perceived ease of use	0.1869	10
Individual Trait	Variety seeking attitude	0.2025	11
RPIT	Involvement	0.2160	12
Product Trait	Relative advantage	0.2182	13
RPIT	Satisfaction while comparing products	0.2343	14
Individual Trait	Income	0.2343	15
Individual Trait	Social Image	0.2528	16
Product Trait	Trialability	0.2552	17
Environment	Social norm	0.2575	18
Individual Trait	Knowledge sharing	0.2688	19
Product Trait	Observability	0.2690	20
Environment	Network externality	0.2849	21
Individual Trait	Pleasure-seeking attitude	0.2873	22
Individual Trait	Impulsiveness	0.2921	23
Environment	Social approval	0.3054	24
Environment	Mass media influence	0.3078	25
Product Trait	Availability of product choice	0.3264	26
Individual Trait	Aesthetic value	0.3349	27
Individual Trait	Status signalling	0.3371	28
Product Trait	Product type	0.3397	29
RPIT	Previous experience	0.3425	30
Individual Trait	Information seeking attitude	0.3532	31
Individual Trait	Risk-taking attitude	0.3551	32
Individual Trait	Age	0.3609	33

Individual Trait	Convenience seeking attitude	0.3789	34
Product Trait	Speed of use	0.3949	35
Product Trait	Store distance/Reputation/Service	0.4082	36
Individual Trait	Expertise	0.4240	37
RPIT	Evaluations	0.4290	38
RPIT	Perceived compatibility	0.4291	39
Environment	Urbanization	0.4292	40
RPIT	Search time	0.4398	41
Product Trait	Technology generation	0.4399	42
Product Trait	Personalized offering	0.4449	43
Individual Trait	Values	0.4606	44
Individual Trait	Self-awareness	0.4606	45
Individual Trait	Empathetic behaviour	0.4606	46
Individual Trait	Absorptive capacity	0.4631	47
RPIT	Perceived (guidance) Information	0.4635	48
Individual Trait	Loyalty	0.4655	49
Individual Trait	Media dependency	0.4661	50
Individual Trait	Strong individuality	0.4788	51
Product Trait	Product size	0.4793	52
RPIT	Perceived complexity of products	0.4817	53
Individual Trait	Education	0.5000	54
Product Trait	Technical specifications	0.5000	55
Product Trait	Incentives	0.5028	56
Product Trait	Innovativeness of technology	0.5108	57
Environment	Industry standard	0.5134	58
RPIT	Perceived risk	0.5157	59
Environment	No. of initial adopters	0.5159	60
Individual Trait	Life event	0.5161	61
Individual Trait	Materialism	0.5161	62
RPIT	Security concerns	0.5473	63
Individual Trait	Subjective norms	0.5500	64
Individual Trait	Inquisitiveness	0.5711	65
Individual Trait	Self-monitoring	0.5839	66
RPIT	Adoption duration time	0.5842	67
Product Trait	Quantity	0.6024	68
Individual Trait	Learning	0.6026	69

Individual Trait	Resistance to mass media	0.6370	70
RPIT	Privacy concerns	0.6526	71
Individual Trait	Gender	0.6738	72
Individual Trait	Do-it-yourself attitude	0.6894	73
Product Trait	Prototype	0.7053	74
Product Trait	Financing	0.7580	75

Table 4.11: Detailed Ranking of factors with FTOPSIS-Manhattan distance

Category	Variable_name	Closeness Coefficient	Rank
RPIT	Perceived Brand image	0.0969	1
Product Trait	Visual appeal	0.1313	2
RPIT	Perceived brand trust	0.1313	3
RPIT	Perceived usefulness	0.1469	4
Product Trait	Innovativeness of product	0.1500	5
RPIT	Expectations	0.1688	6
Individual Trait	Personal Innovativeness	0.1750	7
Product Trait	Attributes of product	0.1781	8
RPIT	Perceived ease of use	0.1844	9
Individual Trait	Novelty seeking	0.1844	10
Individual Trait	Variety seeking	0.2000	11
RPIT	Involvement	0.2125	12
Product Trait	Relative advantage	0.2156	13
RPIT	Satisfaction while comparing products	0.2313	14
Individual Trait	Income	0.2313	15
Individual Trait	Social Image	0.2500	16
Product Trait	Trialability	0.2531	17
Environmental	Social and Cultural norm	0.2563	18
Individual Trait	Knowledge sharing	0.2656	19
Product Trait	Observability	0.2656	20
Environmental	Network externality	0.2813	21
Individual Trait	Pleasure seeking or Hedonistic	0.2844	22
Individual Trait	Impulsiveness	0.2906	23
Environmental	Social approval	0.3031	24
Environmental	Mass media influence	0.3063	25
Product Trait	Availability of product choice	0.3250	26

Individual Trait	Aesthetic value	0.3313	27
Individual Trait	Status signalling	0.3344	28
Product Trait	Product type	0.3375	29
RPIT	Previous experience	0.3406	30
Individual Trait	Information seeking	0.3500	31
Individual Trait	Risk-taking	0.3531	32
Individual Trait	Age	0.3594	33
Individual Trait	Convenience seeking	0.3781	34
Product Trait	Speed of use	0.3938	35
Product Trait	Store distance/Reputation/Service	0.4063	36
Individual Trait	Expertise	0.4219	37
Environmental	Urbanization	0.4281	38
RPIT	Evaluations	0.4281	39
RPIT	Perceived compatibility	0.4281	40
RPIT	Search time	0.4375	41
Product Trait	Technology generation	0.4375	42
Product Trait	Personalized offering	0.4438	43
Individual Trait	Self-awareness	0.4594	44
Individual Trait	Empathetic behaviour	0.4594	45
Individual Trait	Values	0.4594	46
Individual Trait	Absorptive capacity	0.4625	47
RPIT	Perceived (guidance) Information	0.4625	48
Individual Trait	Media dependency	0.4656	49
Individual Trait	Loyalty	0.4656	50
Product Trait	Product size	0.4781	51
Individual Trait	Strong individuality	0.4781	52
RPIT	Perceived complexity of products	0.4813	53
Individual Trait	Education	0.5000	54
Product Trait	Technical specifications	0.5000	55
Product Trait	Incentives	0.5031	56
Product Trait	Innovativeness of technology	0.5094	57
Environmental	Industry standard	0.5125	58
Environmental	No. of initial adopters	0.5156	59
RPIT	Perceived risk	0.5156	60
Individual Trait	Materialism	0.5156	61
Individual Trait	Life event	0.5156	62

RPIT	Security concerns	0.5469	63
Individual Trait	Subjective norms	0.5500	64
Individual Trait	Inquisitiveness	0.5719	65
RPIT	Adoption duration time	0.5844	66
Individual Trait	Self-monitoring	0.5844	67
Individual Trait	Learning	0.6031	68
Product Trait	Quantity	0.6031	69
Individual Trait	Resistance to mass media	0.6375	70
RPIT	Privacy concerns	0.6531	71
Individual Trait	Gender	0.6750	72
Individual Trait	Do-it-yourself	0.6906	73
Product Trait	Prototype	0.7063	74
Product Trait	Financing	0.7594	75

Chapter 5

Understanding and predicting Innovator group customers in consumer goods industry: An Artificial Intelligence approach

Abstract

People make a plethora of simple to complex decisions everyday, including consumption related decisions. For consumer goods manufacturers and retailers, understanding customers' decision making remains unclear, especially with a high-rate of new-product failures. Additionally, for customer cocreation purposes, managers need to select a crowd that can provide useful solutions. Recognising these two intertwined problems of new product innovation, authors help identify and predict future Innovator Group (IG) customers from a real world data from a Spanish consumer goods firm. This study employs several supervised machine learning algorithms including ensemble methods to identify the innovator customers from their transactional, behavioural and demographic data collected between 2015-2018. The results indicate that combining experts' knowledge in determining feature weights, with Artificial Intelligence techniques can help managers identify IG customers with higher accuracy, for targeting and approaching them for cocreation purposes.

5.1 Introduction

In today's hyper-connected world, marketing managers have urgency to identify prospective customers. Moreover for the new innovative products, the search is even more urgent because most of the new products in consumer goods are susceptible to failure - 60% of new products seize to exist within the first 3 years of their respective launches [228].

It is clear that New Product Development (NPD) research is important to academics and managers alike, however it is less clear how to reduce the product failure rate to improve the success rate. Investigation by prior research on determinants of new product success and failure has remained inconclusive - with lower explanatory power of the factors which in turn have declined over the last decade [76]. Organisational theorists have investigated NPD process through team, project, firm, and industry level extensively but have not elucidated it satisfactorily. And while customer participation and cocreation in NPD has accelerated customer's active participation in the NPD process [43], the high rate of new product failure still persists. Cocreation and crowdsourcing research employ various methods to identify and utilise customers' knowledge for creating [257] or improving [160, 257] the new products. Several studies including a meta-analysis on cocreation in NPD have shown that these customers are only effective either at the ideation or at the pre-launch stages of NPD [43]. This is logical as any new product's technical design stage is better handled by the experts at the firms. However, the participation should have helped firms disentangling the failure phenomenon, and the effect is not universally same. We argue that the identification of "right customers" may be the source of the problem. If firms can include those knowledgable customers who can provide user's perspective and ideas to create/improve new products, and help transfer the sticky knowledge on preferences [258], then the new products may represent products that are more appealing to the customers, hence become more successful.

Identification of potential new customers is a hard problem that managers are trying to solve for decades. Customers are complex and often change their purchase decisions [36] including their adoption related decisions. However, with improved computational capacity, Big-Data on customer transactions has seen increase in volume, variety and velocity that in turn, has increased data mining complexity [268]. Big-data captures actual purchase and behavioural information that is simultaneously richer, more granular and less biased than other forms of data. Among the many existing techniques, Artificial Intelligence (AI) techniques are considered better suited to handle such complex, and large amount of data with higher dimensionality [268]. These AI techniques can process voluminous data on real-time with higher precision in forecasting [152]. AI algorithms have proven to be effective for predictive purposes [268] but they have not been widely employed in NPD and marketing research.

Because of the complexity and importance of NPD for firms, academic community needs a revised theoretical lens i.e. *customer cocreation with adoption perspective* to better disentangle the failure phenomenon [71, 76]. By understanding customers' adoption decision-making for new products, firms can identify customers to form a special crowd (from a large population) who are knowledgable, capable and perhaps can be motivated to cocreate with firms. In order to achieve this goal, identification of Innovator Group customers [161, 203] is the main approach adopted in this study. We argue that by identifying and predicting IG customers, *who purchase new products earlier than the early majority, late majority and laggards* [203], managers can utilise this special crowd's sticky knowledge and ideas to create better products. Additionally, NPD managers may reduce new product failures by also targetting IG customers for new product purchase, and nudging them to diffuse information on products to other customers. The IG customers can be identified from their real purchase and behavioural data which are less biased, observational and cheaper as compared to other traditional methods (interviews, focus group, pyramiding or surveys) that rely on customers' memory and perceptions.

In this study, we employ a combination of extant literature and data mining techniques. The characteristics for identification of IG customers are selected from a pool of tested/investigated variables (extant literature) which are weighted according to their importance by the industry experts. Such a measure ensures that the characteristics are generic to the IG customers but not specific to any individual or any product category. In this way, we address the marketing research question on targeting prospective customers from the past data where the same customers shouldn't be part of the training data [229]. We argue that prospective IG customers need not to be part of any dataset but they may share the common identification features with the group. Therefore, customers can still be identified and predicted from the characteristics of other IG customers. On the topic of imbalanced data with asymmetric cost problem [229], our method of selecting the features of IG customers, rather than accepting low-response data or relying on data with probability of leaving potential customers, is robust in principle. The logic is simple: we utilise the Big-Data to identify customers from a set of verified features (by prior research and industry experts), and do not rely solely on the data structure. In this way, we combine the best of academic knowledge with the prowess of Big-Data and AI techniques to test our models.

The methodology selected for this study is primarily supervised machine learning ML based on a number of algorithms including k-Neighbors, Multi-layer Perceptron, ensembles (AdaBoosting, Random forest and Bagging) to classify or identify IG customers from the rest. We have explored the extant literature on classification-algorithms for our specific problem and found that a number of Random Forest based ensembles are better suited because of the nature of the decision tree classification. Also, training with AI algorithms when conducted with labelled data - identified by marketing-domain experts - often yields superior customer insights [78]. In consideration to all of the above discussion, we have formulated two research questions for this study:

- How to identify the Innovator group customer from the Big-Data with machine learning techniques?
- Which is an optimal algorithm/method for Innovator Group's identification?

We tackle our research questions, by proposing a framework that academics and managers can employ to design and identify IG customers. However, there are a number of challenges in achieving this goal. First, classifying Innovator group customers requires the knowledge of all possible variables that affects their adoption decision-making. Second, from the data perspective, the covariate space for the variables is high-dimension in nature, and there would be insufficient data for any firm (practically) to collect and utilise all the variables simultaneously. Therefore, we need to identify the most important variables that can be collected and used for learning the optimal way to classify IG customers. This would also ensure that data's granularity at the customer level is not lost by aggregation for the purpose of gaining better insights.

5.1.1 Our approach

We use a novel combination of methods to build theory: (a) Extensive synthesis of extant literature provides some of the critical characteristics to identify IG customers. (b) Explorative data analysis reveals and provides real information on customer purchase behaviour for new products. (c) Machine learning with ensemble algorithm provides large-scale analysis and corroborates and extends extant literature constructs with more precise predictions. (d) Comparative methods provides a clear distinction between different algorithms for obtaining a better solution for the classification problem.

We contribute to *new product innovation*. First, we select the most important factors to identify IG customers. These characteristics are weighted according to their degree of influence on an IG customer and can be applied to early adopters of other sector. In contrast, past research has presented fragmented factors of identification and often missed the importance of these factors in customer's decision making process.

Second, we contribute to research method for *customer identification problem in crowdsourcing*. This method is economical in time and resources as compared to past research that employed mass screening, pyramiding or surveys that consumes more time and resources. Additionally, the screening only works with large online communities who have self-selected to different forums. Thus, a key contribution of our study is finding a new practical way to identify IG customers with less resources and higher probability. Third, we contribute to the *technology management research* where we take a balanced approach to theory-driven and data-driven perspective to address a new product innovation problem. We address the lack of human insights into machine learning methods constructs and novel theoretical relationships. In doing so, we include human insights from experts and IG, cocreation and crowdsourcing literatures in advancing the theory on early adopters' refined characteristics. As important, we contribute to novel theory building way: combining extant literature, human experts and machine learning. This combination enables theory building with machine learning that often has failed in explaining the *why's* of a phenomenon. The overall contribution of the study is to develop scalable and practical decision support capability for new product innovation and marketing managers.

We also address some of the key research deficits that management literature has when dealing with machine learning applications. We not only employ relevant features and predict Innovator group customers, we also present descriptive analysis with evaluation of a number of methods (on their performance scores).

The paper proceeds in the Section 2 where we review the literature on New product development and cocreation/crowdsourcing in the process. In Section 3 we describe the natural experiment that provided the data for training and validation of the machine learning methods. In this method, we also describe the data, the specific membership linked information designed by the manufacturer and the features selected for training the machine. In the Section 4, we present preliminary results from the methods, including the prediction scores and comparison among them. We also present statistical analysis of the algorithms performances. In the last section, we conclude with implications for academia and practitioners with some limitations of our study.

5.2 Research background

5.2.1 Customer adoption decision-making: Focus on new-product innovation and cocreation with the crowd

New product failure is a complex phenomenon that is affected by both macro (government, economic situation, natural resources, organisational) and micro-factors (customers, employees, managers, local). If practitioners can gain access to customer preferences, needs and choice structures, and other factors that affect their decision-making, then products could be made fitting to various customer segments. In time, segments may increasingly give away to personalised offers with improved understanding on individual customer's decision-making factors. Altogether, when firms gain better knowledge on customer adoption-factors, products could be made accordingly and then, customer adoption may increase for the new products.

Researchers have studied customer adoption decisions for decades to find out the key factors, and characteristics of certain group of customers (Innovator Group and Lead Users) who adopt new products first. Although, there is a rich literature in this area, the diverse perspectives and areas of research make the research hard to comprehend. For this reason, a systematic literature review was conducted by the authors to gain an in-depth knowledge on the subject. Nonetheless, systematic identification of early adopters or innovators or lead users has been difficult and resource intensive [261].

Research on lead-users or user-innovators have employed qualitative interview and case study methods [212, 242, 257]. On the other hand, research on identification of lead users involves quantitative methods such as (mass) screening of potential innovators - with self-selection bias [160, 261] or netnography of screening from online communities [14, 17] or pyramiding with referrals [261]. These methods have resulted in selection of lead users primarily based on opinion leadership and ignoring other important characteristics suggested by Luethje [159]. Lead users can help generate ideas, solve underlying problems and create the products or services themselves. However, their representation in the customer base is extremely low. As compared to lead users, the IG is more representative in the population (16% for consumer goods) and can influence product launch and post-launch diffusion as successfully as the lead users [107]. For the Innovator Group customers, focus group studies/interviews, case studies [160, 257] and surveys for conjoint analysis [100, 214] are a common practice in marketing which is cost intensive. However, the information collected from these methods is mostly on past behavioural data on a certain product or a single product category. Because of the limited ways to approach customers have been cumbersome, traditional methods have resulted in less desired outcomes for the practitioners [134]. With data privacy regulations in place in Europe, accessing customers for experiments/interviews is becoming increasingly difficult. Management science needs better and unbiased ways to find these valuable customers with the help of big-data and machine learning techniques that are well suited for larger features and voluminous data for prediction purposes.

The debate has not been settled yet for determining the separating threshold for lead users (LU), emergent nature consumers and Innovator group customers as it depends on product categories, and the numbers can range between 10-38% of all users [159, 160, 178]. The rarity of lead users who actually innovate on their own has made the detection of this group even more difficult [205]. On the creativity and adoption scale, the innovators and early adopters come next to the LUs but they are more representative in a population than LUs.

However, the IG customers are also product dependent or domain specific experts similar to LUs. Though some recent studies have taken an approximate percentage (3.6%) for these early adopters for practical purposes [97], we argue that the true representation of these customers can be traced with the influencing factors in the general population that affect them without any crude approximation. Therefore, our algorithms will identify the IG customers from all users with weighted factors and give them score from a discrete range of 0 or 1, where 1 being IG customer and 0 being non-IG customers. We consider any customer who is an innovator group for a product, can also be a non-innovator for another product or in another time period. By doing so, we try to capture the chameleon nature of humans (changing with time and circumstances) rather can invalidating their shifting (realistic) behaviour to a fixed and inflexible pre-deterministic category.

5.2.2 Machine Learning as a tool

AI techniques are considered better suited to handle large amounts of data with higher dimensions and complexity [4, 152]. They can process voluminous data on real-time with higher precision to forecast/predict. The techniques that are capable to work on voluminous customer data primarily include, machine learning (ML) - supervised, unsupervised and reinforced learning. The reasons why AI techniques are well suited for the customer transactional data can be analysed from either technical or managerial perspective. From the technical perspective, AI techniques are anticipated to perform well in process optimisation [121], automation of administrative tasks [136], error detection and accurate predictions [15]. Studies show that generation of valuable insights from AI enabled analysis becomes input resources for creating new products, services and processes [184]. This is made possible because of AI techniques' adaptable algorithmic designs and training with the input knowledge (labelled data) by domain experts including marketing, finance, operations etc. From the managerial perspective, the speed of execution and the diligent processing of the data in real-time, can exploit the strength of AI techniques to its full potential [199, 265] for e.g. finding intricate correlational insights on customer behaviour and customer decision making from the datasets. Also creating clever algorithms that are unmatched in accuracy, precision or speed is also an advantage for using AI techniques. Hence, the unique insights generated from this process can act as input-resources for future innovations in products/services, customer relationship management. Therefore, the potential of AI based machine learning techniques is worth exploring for customer behavioural data for identification problems.

Under supervised machine learning, various popular methods include k-nearest neighbour (knn), artificial neural networks (ANN), support vector machine (SVM), decision trees, naive bayes etc. The objective of supervised machine learning, at its core, is to build systems that perform better with training experience and generate good predictions with unseen data. In a simple way, among different learning paradigms, "inductive learning with examples is the most widely used paradigm" [67]. Researchers have found that it is easier to train an intelligent system with examples by showing the desired input-output operations [126]. In general, a supervised learning setting comprises of a pair (X,y) where X is an input vector and y is a scalar output. A loss function is defined as the cost of predicting y* when the actual answer is y. The loss function f belongs to F family of functions which is often parametrized by a weight w. The aim of the learning system is to minimize the loss function $l(f^{w}(x))$ when averaged over the examples. Additionally, an empirical risk function ($E^{n}f$) measures the performance over the training sets, which according to the statistical learning theory is sufficient when the chosen family is somewhat restrictive [21]. The bias-variance trade off indicate the balance between under-fitting and over-fitting in a model. The learning system should identify the rich structure in a dataset but restrains itself from over-fitting the spurious patterns.

Although artificial neural network works within the supervised ML and deep learning techniques, they are a set of processing elements that are interconnected in networks similar to the neurons in the human brain [279]. Artificial neuron networks have the same functionality as the biological neurons: transferring information. The neurons can be single or multi-numbered that determines their functionality. For example feed-forward neurons (single-directional) transfers information from input layer to hidden layer to outer layer; recurrent neurons (bi-directional) can transfer to and from layers; and convolutional neurons (in loops) imitate the neural networks of animal visual cortex [279].

On the other hand, unsupervised machine learning methods differs from its supervised counterpart on the account of data. Only the input variables are available for these methods where the output/target variables are often unknown or undefined. The primary goal of the unsupervised ML is to uncover the hidden patterns from the data. Some of the popular methods include cluster analysis, deep learning and dimension reduction techniques.

Semi-supervised machine learning method falls in between the supervised and unsupervised machine learning. This method often deals with data that has the target variable with incomplete information and it tries to complete the information by continuous learning. This approach is extremely useful in reality because collecting complete set of data is expensive and unfeasible at times. However, firms collect some data for their operations which are incomplete but may offer insights when utilised properly with learning techniques.

On the other hand, Reinforced learning is a rapidly evolving research field in computer science where the basic learning takes place with an active learning method. For each task, the algorithm is either rewarded or penalised for its choice and learns from this experience to get optimal solutions. This method may resemble Markov Chain's method and Reinforced learning has gained popularity in the automated systems, self-driving cars, multi-agent environment in particular. Inspired by behavioural economics with rewards and penalty, the system learns with each task.

In recent years, much effort in the AI and ML research has been devoted to improve on the existing methods. The result of these efforts has advanced *ensemble*, *deep neural network*, and *probabilistic graphical models* techniques. As the name suggests, ensemble methods are collection of a number of algorithms where the deficiencies of individual algorithm are mitigated by the combined form. Some of the common techniques to achieve this includes bootstrap, boosting aggregating or bagging and stacking. In the bootstrap technique, a sample data set that is bootstrapped a large number times is used to train the predictors. For boosting, classifiers are trained sequentially on the basis of their performance. AdaBoosting is a popular example of boosting where the classifier starts training with the original dataset but subsequentially works on inaccurate instances and works on more difficult tasks while working on copies of the original dataset. Similarly, Gradient Boosting method is another ensemble that has gained popularity because of its high predictions among ML practitioners and it uses *decision tree* as its the baseline learner. Similarly, *Random Forest* also uses Decision tree as its base learner.

All of these methods are based on some *decision-making problem* which can be broadly categorized into *classification*, *discrimination*, *ranking*, and *prediction* [67]. Classification can be defined as assignment of finite set of elements into some pre-defined groups where the characteristics among the groups may not overlap. For e.g. Fisher's Latent discriminant analysis (LDA) functions under this definition. However, with human behavioural data, traditional classification methods face difficulty in mimicking the real life. Therefore, fuzzy set theory - which deals with uncertainty, hesitancy and overlapping information - performs more effectively in such scenarios than the traditional classification methods.

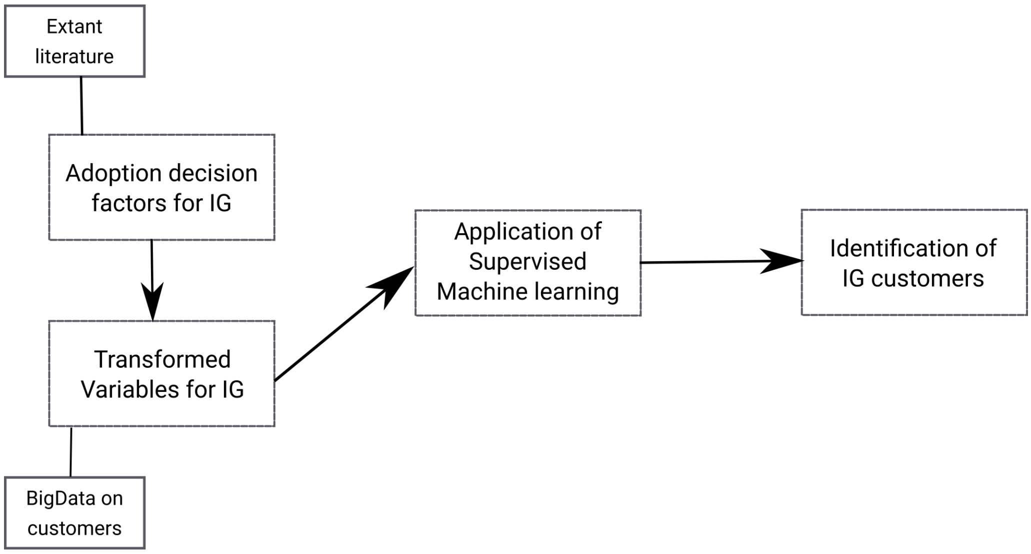


Figure 5.1: Conceptual framework for identifying Innovator group customers

5.3 Supervised Innovator group classification Model

In this study, the core focus is to identify or classify innovator group customers from an observed data spanning multiple years. The longitudinal data provides a wider window of time where customers can behave differently in separate time periods. This aspect of the model is unique because it captures the changing nature of humans in general, and early adopters in particular, who can show variation in their behaviour within or outside of their domain expertise. In our proposed model, we combine extant literature's investigated knowledge, the new product experts' ranked variables with importance, and then these variables form the feature-space for the supervised machine learning algorithms. Finally, with optimisation and statistical analysis, the model selects the best algorithm(s) for the identification of innovator group customers.

As shown in Figure 5.1, the key challenge is to find the most important variables for identifying IG customers and to find data for each of them. In the real world, collecting all the variables remain impractical, expensive and mired with biases. Hence, with structured unbiased data, we transformed the variables to represent the feature space for the classification problem. The result of conducting a systematic literature review gave us 103 investigated factors that affects the IG customer. By extracting knowledge from new product launch experts, we obtained the respective weights for these factors. In our knowledge, we are the first to provide a meta-analytic view to the early adopters collective factors and to rank the factors in their importance. Additionally, we train the algorithm to learn the underlying relationships among the variables and predict the Innovator group customers from structured data i.e. labelled according to customers' purchase history.

In the Innovator group classification model, the estimator has data D: where D is the dataset for inferring a learning model. The feature variables for the data are X_i, ...X_n that are the observed values for the system, for example 629552 customers of the Spanish firm. The target variable y_i, ..., y_n is the target value for each customer. Y is a composite variables created for identifying first level Innovator group customers from the recency, frequency and quantity of purchase in the first $1 - 30$ days of each new product. To classify the customers, training data (X,y) is provided as an input and output data, where X and y are split into training and test data from the beginning. X is the input vector X_j and y is an discrete variable, integers ranging between 0 and 1. X comprises of 28 brands, 230 varieties, 78 families, 22 formats, 2 sexes, 9 area codes, 6 product period-ranges, 5 membership variables, novelty score, attractive product score, natural product, retail and facial values, and demographic variables such as age, family members, number of children. In total, for each customer, X vector has 410 features. In comparison, y variable is a discrete variable indicating 2 degrees of IG customer ($0 =$ Non-IG, and $1=$ IG).

Data description

The real-world data we have studied in this paper is from a Spanish consumer goods manufacturer. This setting is appropriate for many reasons. First, the Spanish brand is one of the most innovative firms in the world, in the consumer goods sector. The firm launches a lot of new products on a frequent basis - approximately 150 new products in a year. Additionally, the innovative firm has a dedicated mechanism to track new product's sales i.e. every product has a bar-code underneath and customers can scan the code with their phones to send it to the firm. In this way, the firm collects information on products directly from the customers without violating any data privacy regulations and manages to gather crucial information on customer's purchase behaviour.

The firm's app works as a loyalty programme where customers get themselves registered, for scanning promotional barcodes and sharing with the firm. Every time a customer shares a code, s/he receives some coupons/discounts. However, in order be a member, customers also share some demographic details such as age, sex, address, number of children and family members. For each scan, customers receive "pins" or membership points on their account, and these pins can be exchanged to purchase any product in the future. Since, the firm is not a retailer, it is an innovative way to track product purchase through mobile devices. The data we received was collected by this mechanism and it contains customer's purchase information including date, product, quantity, demographic details filled out while registration and pin related information.

In the dataset, there are 629552 unique customers who used the app actively for the period 2015-2018: four years of customer transactional data. The firm has provided us with some labels (features) that they use internally in the organisation. We had elaborate discussions on the meaning of these features with the team to understand the data properly. Also, the definition of a new product is in accordance to their industry and includes both moderate and radical innovations. Hence, with each new SKU id in the data, we consider that to be a new product for the purpose of this study.

5.3.1 Data issues

In machine learning literature, there are a number of data challenges including covariate-shift, aggregating target variable, concept shift and class imbalanced data. Addressing these issues, we explain how our data is selected. In our study, the predictive variables and target variable belong to the same real-dataset. This implies that the relationship (functional form) derived by the training data is also applicable to the test data. Hence, we avoid encountering both covariate shift and concept shift in our data because the distribution of the predictive variables in both training and test data remains same. We do not sample from different data collected over different time or method. This robustness of data gives our study the requisite confidence in avoiding the common data problems as mentioned earlier.

Second, the data is generated by the customers themselves when they send the scanned barcode to the firm's app. Hence, the data contains both the predictive variables and the target variable for each customer. Unlike most marketing data that requires aggregation of target variable and loss of information in the process, our data is trained and predicted at the customer level. Lastly, the imbalanced data representation is a particularly difficult problem. Our goal is to predict Innovator group customers who are around 2.5-5% in any population. This is true for our data that represents the population. Therefore, we take care of this challenge by balancing the data with the techniques such as random under-sampling, random over-sampling [131, 132]. After analysing our real-world data, we employ pre-processing method to overcome the data imbalance problem, in particular with K-fold cross validation technique and SMOTE method (Synthetic Minority Oversampling Technique) for classification of innovator group customers [80]. The reason for employing SMOTE was to make the training-data class balanced because in general, the number of IG customers in a population remains disproportionately lower than the number of non-IG customers [97, 203]. SMOTE is considered one of the most influential data sampling algorithms which rebalances an imbalanced data by synthetic addition (interpolation) rather than replication of the minority class instances [80, 85]. Additionally, SMOTE is an easy to use and accessible oversampling method for researchers with open source software packages.

5.3.2 Data pre-processing

From a systematic literature review, 103 investigated influencing factors were collected and the authors developed a framework where these variables were categorized into various socio-psychological, product and environmental categories. Among these selected variables, 30 most important variables were selected after being weighted by new product industry experts. Upon receiving data from the Spanish firm, the authors mapped real variables to the selected variables from the framework to prepare the feature-space for machine learning. Second, among the real variables, context-dependent relevant variables were selected by the experts and the researchers who have extensive experience in consumer goods sector. This step is equivalent to *feature sub-set selection*. After analysing both sets of variables (features), authors also created new features by combining some variables. This process is known as *feature co-dependency creation* in machine learning literature.

The stepwise process for classifying the innovator group customers, in accordance to the conceptual framework, is the following:

- **Data preparation:** Preparing data primarily includes cleaning unnecessary information, errors, and treatment of missing values. We used sci-kit learn and python language for all the coding purposes as python and Numpy arrays are easy to scale for machine learning algorithms.

- **Attribute selection/Feature engineering:** Mapping variables from the real data to the variables selected from the literature. We considered weights from the previous study in the process to select the most important variables for the study.

- **Feature selection:** Labelling data with assigned attributes and filter out non-relevant variables from the database. Visualisation of data with Matplotlib and Seaborne packages helped indicating underlying distribution of each selected feature.

- **Training:** Dividing the entire database into validation and training data sets, then randomly dividing the training data into train and test sets in 70:30 proportion. After training each dataset with algorithms, we tested with the unseen test-data to provide validity to the (accuracy) predictions. Visualisation of test and train accuracy for each algorithm is performed.

- **Validation** Utilising the untouched test-data to predict IG customers, based on the results from the training. We also compared the performance of algorithms' prediction accuracy.

Real-world data are often not suitable for machine learning algorithms as they contain missing values and errors which are unsuitable for processing. Some of the primary steps required to clean the data includes treatment of missing values and error correction i.e. typo/space/other erroneous characters. This step is an essential task that needs to be done before conducting any statistical analysis on the variables (features). The missing value of the data in our study were of different kinds. The most common missing values were in the form where variables had no entry for a customer. When we import the CSV file into pandas data-frame, most of the variables were in “object“ data type and few were in integer64 format. According to their nature, we transformed these data-types into “numeric“ or “datetime“ format. With this step, the missing value cells were filled with “NaN” or “NaT“ respectively. These transformation were done for the numeric variables only. We did not treated the individual NaN/NaT values at this stage because we use “Imputer” algorithm in the later stage to treat all the variables simultaneously.

For the categorical variables such as Brand, Family, Variety, Format, Area and Sex, there were text related errors. Once we cleaned all the textual errors, by checking each variable individually, we transformed the categorical variables to dummy variables. This step is otherwise known as “One-hot-encoding”. By doing so, the categorical variables could retain information in numerical values without losing any important information. After all of these cleaning process, we obtained 28 brand variables, 78 Family variables, 230 variety variables, and 33 format variables. For the Sex variable, we calculated the probability of female to male in the existing data and then filled the missing values randomly with female or male in the same probability. Then we transformed the variable into dummy variables. Similarly, we treated the missing values for “NielsenArea” variable with the probability of the existing data. In this way, we guessed the probability of a customer’s residence zone in an unbiased manner. We also transformed the variable into 8 dummy variables.

5.3.3 Feature engineering

Feature selection process or otherwise known as *feature engineering* is an important step in determining the final outcome/target variable. This step ensures that researchers have selected the most relevant features for their training dataset.

In our study, we retain the demographic details of each customer and created some new variables from the data. For example, we created “CustomerAge” from “Date of birth” and “last login date” variables as age is an important factor but it was not explicitly present in the data. Since, customers were active for different lengths of time, we mapped their Age to the last date of their activity for a better evaluation. Applying the same principle. we created “MembershipPeriod” variable from customer’s registration date and last login date. Also, we extrapolated last login date to the last validation date where the values were missing. In doing so, we captured customer's actual purchase activities.

We also calculated the number of products that each customer bought from a combination of unique Validation date and "SKUID". Since customers may buy multiple products (with different SKU ids) or same product in one or many days, we combined the two variables to filter all the products bought by the customers. We named the variable as "NumProds". We also created "ProductAge" from the date of launch and the actual date when a customer bought a product. There were 613 different products in the database. The next variable that we created from the data was "Product range". This variable informed us the day on which a particular SKU was bought by a customer, in the post-launch phase of a product. We created this variable from the "ProductAge" variable. The information is vital for understanding who the early adopters are and their behaviour. According to Roger's theory, we segregated the range of days into nine periods: $1-3$ days, $4-7$ days, $8-30$ days, $1-2$ months, $3-6$ months, $7-12$ months, $13-24$ months, $25-36$ months and greater than 36 months. Since, most of the consumer goods products either fail or truncated sales, we decided to segregate all products above 3 years to be bundled under one range. We calculated a variable named as "$ProductperFam$" which indicates how many products a person has bought per family member (his/her family). The variable indicates products purchased per person more accurately and informs us about whether the (number of) products are bought for an individual or on behalf of a household.

We also created some important variables from the data. "NaturalVar" was one of the variable that represented all the products that were labelled natural/non-flavoured. We also created "NoveltyScore" from the number of new products ($1-7$ days range) and the total products. "AttractiveProd" was created from the redeemed points that the customer used to buy a product. This indicates the attractiveness of the product for the customer as compared to the product gamut. The variable captures intrinsic qualities of a product that seemed attractive for a customer.

We created our single target variable (whether IG or not) "$Class\_cust$" by analysing whether the customer bought any new products in the time periods of $1-3$ days, $4-7$ days, and $8-3$ days. The IG customers are coded as 1 and Non-IG are coded as 0 for easy interpretation. In addition to that, we also considered the variable "$Products/Family$" and "NumofProds" for creating three sub-groups for the IG customers. In the literature, RFM ((Recency Frequency Monetary) defines a customer as a viable customer [77]. Even though our goal in this study is to define the IG customers with most of their important variable, for determining the group in the algorithms, we created the base target variable with recency, frequency and quantity of new products (or RFQ) variables. The reason for not including monetary value was from lack of data on each purchase. However, the threshold for quantity was determined by analysing the distribution of "NumofProds" of people whose purchased new products in the first 3 days. The threshold was assigned at 10 as this was the mean number of products, found in the data analysis. Similarly, the threshold for frequency was determined from the distribution of the data. The following table shows the subgroups and their RFQ values for clarity. The reason for choosing the thresholds from earliest customers was because of the clear nature of IG customers and their willingness to purchase early. Also their overall purchase remains higher than the population's mean (4.34).

5.3.4 Feature selection

From the previous study, we found the most important factors for innovator group customers (see Table 5.5). Starting with the 30 variables, our task was to map them to the real-world data. In particular, perception based variables were hard to map, since the formative items of such variables were collected by direct answers from participants. In the real data, these variables become unobservable variables. However, with deliberation and referring to the literature, we could map some of the variables with logical reasoning. However, environmental or social factors are unobservable in the data as it takes place outside of the purview where the data is created. Below is the table (Table 5.1) that shows the transformation of such variables in detail.

5.3.5 Algorithm selection

Training of machines with algorithms is the principal constituent for machine learning. Studies show that when a training is conducted with labelled data, identified by (marketing) domain experts, results yield superior insights [78]. The foundation of our study lies on the extant literature of the NPD, customer cocreation and decision-making. Therefore, we included variables from the extant literature (validated by experts) and another set of variables that were transformed/selected from the real-world customer transactions. In doing so, we expect the result would yield a rich set of fitting variables or feature variables that are suitable for training. There exist many supervised machine learning algorithms that learn from input-to-output operations to perform better predictions. Algorithms have improved on their performance over the years with increased attention from practitioners and availability of BigData [287]. Nonetheless, the algorithms which are considered best have two main characteristics: simple to understand and easy to train the data.

Table 5.1: Transformation of variables from extant literature to/from real-data

Source	Variable_name	Transformed variable
Literature	Perceived Brand image	Brands from which purchased
Literature	Perceived brand trust	Max number of purchases from Brand
Literature	Visual appeal	Format type
Literature	Perceived usefulness	Variety of product
Literature	Innovativeness of product	Number of products with flavours
Literature	Expectations	Frequency after new product purchase
Literature	Personal Innovativeness	Number of new products purchased
Literature	Attributes of product	Variety of product
Literature	Novelty seeking attitude	Number of new products purchased
Literature	Perceived ease of use	Size of product
Literature	Variety seeking	Different varieties purchased
Literature	Involvement	Duration of active membership
Data	Facial value of redeemed products	Relative advantage
Literature	Satisfaction while comparing products	Different product family purchased
Literature	Income	Area of residence
Data	Sex of customer	Sex
Data	Date of birth	Age of customers
Data	Purchase date	Product age when bought
Data	Registration date	Active membership period
Data	Number of children	Size of customer's Family
Data	Number of Products purchased	Product for each Family
Literature	Pleasure seeking attitude	Number of flavoured products
Literature	Impulsiveness	Number of products per purchase
Data	Purchase history	Purchase periods for (new) products
Literature	Aesthetic value	Attractiveness of products
Literature	Product type	Different SKUs

For classification problems, many algorithms can train labelled data, learn the underlying relationship function and classify the target variable effectively. Some of the popular algorithms are *Perceptron*, *Decision Tree*, *Support Vector Machine*, *k Nearest Neighbour (kNN)*, *Random Forest*, *Naive Bayes* and *Artificial Neural Network (ANN)*. Although the most efficient algorithms largely depend on the nature of the data, algorithms also get affected by (number of) feature space and goal of the project. The underlying learning methods also contribute to the effectiveness of the algorithms. Decision tree, bagged trees, boosted trees and boosted stumps learn from each tree in the forest and depends on the number of splits for each decision. Perceptron, neural network, Multi-layer Perceptron learn the underlying non-linear function among the feature variables in a dataset and approximate for either regression or classification estimator. The hidden layers between input and output layers transform the learned linear and non-linear functions without explicit information on this stage. This is known as the *Blackbox* in machine learning. However, there has been a conscious effort from the ML community to improve explainability, transparency and interpretability [207].

The machines can learn these functions without making it human readable and hence extracts more information and lacks causal explanations. In any experimental setting (natural or field), omitted variables are integral part of the set-up and they contribute to the standard errors. To mitigate the omitted variable's short coming, by including as many variables available, ML helps in capturing most of the information in its prediction. We are mindful that even with a large number of feature variables, some variables may escape the observation data due to their indirect influencing nature. For example, social network or direct advertisement may have inspired a customer to purchase certain new products that can not be observed from the observational data.

Performance measurements among algorithms are difficult to obtain as there are a good number of algorithms that generate equally satisfying measurement metrics. Studies show that no single algorithm is more efficient than the other but an *ensemble* of algorithms can mitigate some of the individual algorithm's deficiency. Hence, we decided to start with kNN, MLP, Random forest and then use ensemble methods such as AdaBoosting, Bagging with decision trees as base estimator to compare the performance task of classifying the IG customers. In general, comparative analysis remains difficult and inconclusive as theory of machine learning is constantly evolving. The underlying assumption is that with improvements, measurement metric's comparison may also improve with time. While deciding on performance score, we encountered a problem similar to financial fraud problem. Our research aim is somewhat similar to the fraud detection because the number of identification unit (frauds) is rare in the population but their importance is high for the financial firms. Hence, we adopted similar metric rules for gauging our algorithms' performance with scores such as precision, classification report, confusion matrix, F1 score, roc-auc with GridSearch cross validation.

5.3.6 Descriptive analysis

The descriptive analysis of the features reveal interesting information about the IG and Non-IG groups. As shown in Fig. 5.2, the overall range of products purchased for each family is between 0.5 to 10, whereas the product range for Non-IG family lies between 0.5 to 4. This indicates that the overall purchase number is higher for IG customers, even when there total number in the population is significantly lower than the Non-IG customers. Managers can look deeper into each groups in order to plan specific marketing interventions. Please note that the figure is created with suppressing the outliers for better statistical interpretation between the groups.

By looking at the number of pins (every time customer scans and shares for loyalty points) reveals interesting information on the two groups (see Fig. 5.3). The number of pins generated by the IG group customers is higher than the Non-IG group's. This indicates a higher interest of the IG customers in the loyalty programme and in the innovative way to engage with the firm. Managers can even consider the specific sub-group of IG customers (4th quartile) for engagement in new product communication as these customers are the most actively participating in the loyalty programme. Additionally, 75% of the IG customers had 3 or more pins per family member, which is higher than the number of products per family by the IG. This indicates that they purchase products with higher pin-value. In contrast, the Non-IG customers are less engaged in the loyalty programme or pins, and have slightly more number of pins than the number of products purchased.

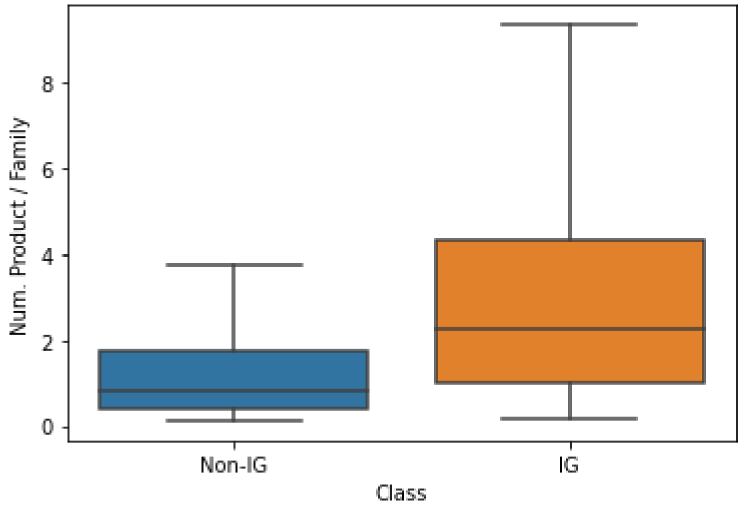


Figure 5.2: IG vs Non-IG customers: Number of products per Family member

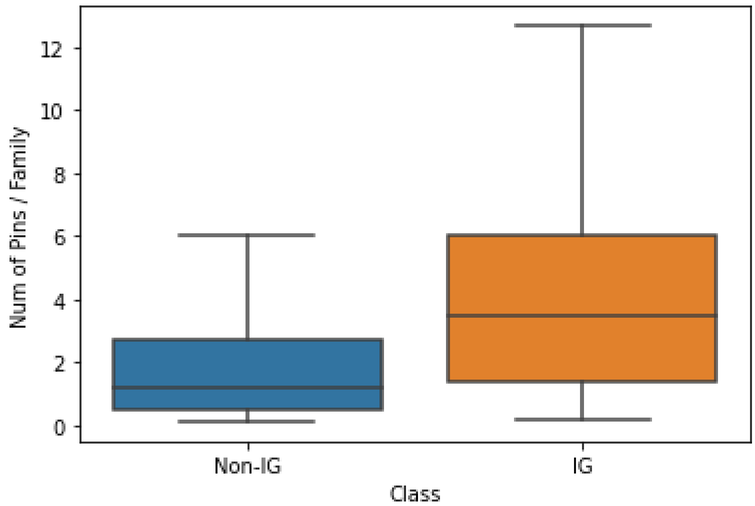


Figure 5.3: IG vs Non-IG customers: Number of pins generated per Family member

With the analysis of customer's age with IG customer and Non-IG customer, some interesting information are revealed. IG customers who bought the most products per family seem to be middle aged (42 years) and Non-IG who bought most number of products for their family are also middle aged (44 years). Non-IG customers' age ranges more than the IG customers which gradually decreases with age where as IG customers age ranges shorter, and the decline of product purchase is sharper than their Non-IG counterparts (Fig. 5.4).

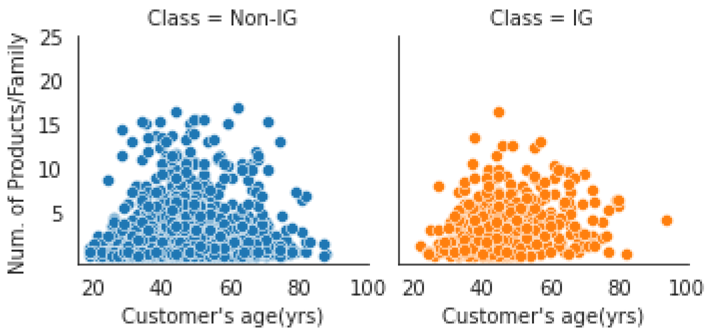


Figure 5.4: IG vs Non-IG customers: Number of products and Customer age relationship

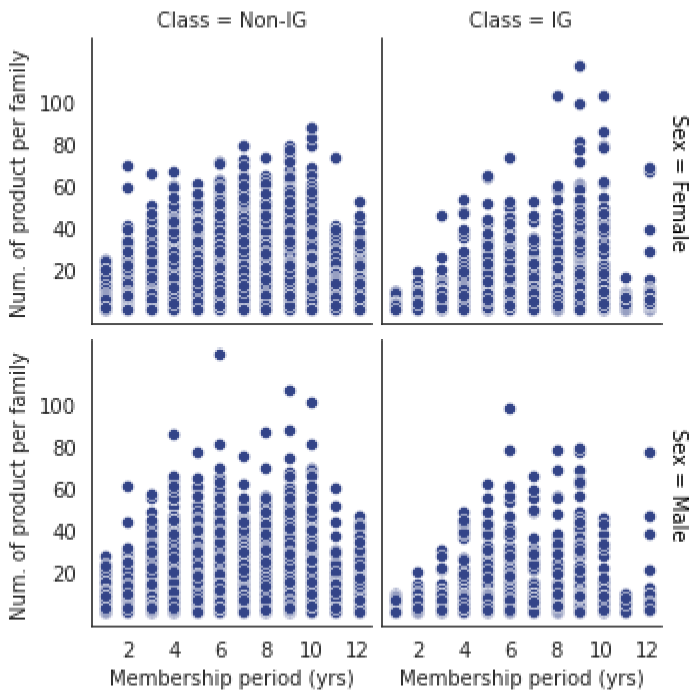


Figure 5.5: IG vs Non-IG customers: Number of products and Membership relationship

Similarly, the membership periods for both IG and Non-IG customers seem to be similar (see Fig. 5.5) with more consistent purchase by the Non-IG group. However, the number of products bought by IG women customers reaches peak in their 9^{th} year (see Fig. 5.5). In general, women retain membership with the firm longer than their male counterparts, for both the groups. There is a rebound of interest in the 12^{th} year by IG (both men and women) and Non-IG women.

5.4 Method

This section introduces a generic description on the existing ML techniques that are widely used by the practitioners and academia [70, 243, 280] and we use these methods in this study. We discuss the measures we take for each method to accommodate their idiosyncrasies such as feature scaling, hyperparameter tuning, etc.

From the discussed supervised machine learning algorithms, we trained and tested our data with Multi Layer perceptron (MLP), k Nearest Neighbour (kNN), Random Forest, AdaBoosting, and Bagging algorithms. In particular, we chose the ensemble algorithms that have been demonstrated to be more efficient for classification purpose than any single algorithm. Their strength lies with the ability to mitigate individual algorithm's weaknesses and build a stronger learning method. For each of the algorithm, we also conducted "GridSearchCV" which incorporated multi-fold Cross Validation (CV) while meticulously searching for the best hyper parameter values. Additionally, grid search selects the best estimator with the best combination of parameters for the model, and then we train and test our data with the best estimator. For reproducibility of our study, we have included a single random number as the seed and split our data in the ratio of 70 to 30 (train and test). For the training and testing of different algorithms, we employed *Sci-kit learn* software for the purpose with Python language [189]. We also used SMOTE algorithm to balance the data with a compatible software with Sci-kit learn [148]. For pre-processing of data to prepare for machine learning, we used in-house Python codes for the purpose.

Multi-layer Perceptron (MLP) classification

MLP classifier is a neural network based supervised learning algorithm that trains with neural networks, along with one or more non-linear hidden layers. Given a feature space $X = X_i \ldots X_n$ and a target variable y, MLP uses a non-linear activation function (sigmoid or ReLu function) to approximate the underlying function from the data. For classification, MLP implements back-propagation learning for its training. Additionally, the model optimizes a cross$-$entropy loss function for classification. Since back-propagation learning is based on some form of gradient descent on each node for weights, it starts from the y $_i$ and calculates from outer to inner layers. Given the training examples X_i, y_i with one neuron and one hidden layer, MLP learns the function:

$$f(x) = W_2 g(W_1^T x + b_1) + b_2 \tag{5.1}$$

where $g : R- > R$ is the activation function, W_1,W_2 are weights of input and hidden layers respectively, and b_1, b_2 are bias added to the hidden and outer layers respectively. Also W_1 E $\mathbf{R}^m$ and W_1, b_1, b_2 E $\mathbf{R}$ are model parameters. In general, the activation function (g) is hyperbolic tan and it passes through a logistic function $f(x)$ for binary classification:

$$g(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \tag{5.2}$$

$$g(z) = \frac{1}{1 + e^{-z}} \tag{5.3}$$

The classifier is particularly sensitive to feature scaling. Hence, we have scaled all our features with Standard scaler and then balance the data with Synthetic minority oversampling technique (SMOTE) algorithm for addressing IG customer under representation problem. MLP also requires tuning of a number of hyperparameters such as number of hidden neurons, layers and iterations. Therefore, we employed GridSearchCV for the hyperparameter tuning along with cross-validation of the training data.

K Nearest Neighbour (knn) classifier

The kNN is one of the most used methods for classification(pattern recognition) because of its simplicity and performance. It is an instance based learning that simply stores the instances of the training data rather than building a model. The kNN rule is based on nearest neighbour principle and it classifies each unlabelled data by majority label among a predefined number of training samples (k) closest to a new point and predict the label from the samples. The performance heavily depends on the distance metric used to identify the nearest neighbours. As the name suggests, users need to define the number of samples beforehand. In order to reduce the arbitrary nature of k, many studies have suggested how to decide the hyperparameter [88]. However, k is data dependent. Thus, we preferred an exhaustive searching by GridSearcgCV algorithm to find the best solution. When k value is higher, effects of noise in also reduced [192].

kNN method used euclidean/Mahalanobis distance for the new data classification and for this reason, features need to be scaled prior to the calculation. For comparison purpose, we also conducted regression on the data with Logistic Regression as it can classify the dependent variable if it is in discrete (not continuous) or binary form.

Decision tree classifier

Decision tree classification is a non-parametric learning method that falls under supervised machine learning category. The overall aim of decision tree classifier is to predict a target variable (discrete values) by learning simple decision rules from the data: features and splits. In the classification decision trees, there are three nodes: root/parent, internal node/branch and leaf/end node. The learning starts from the top with a parent node and then moves downwards until it reaches a class decision at a leaf.

Inductive algorithms that learn at each node of a decision tree are ID3, C4.5, C5 and Classification and Regression trees (CART), where CART has become the most popular algorithm among the listed. Learning happens with information gain at each node where the split at a node depends on the previous node's decision and selects the information with a conditional probability. The quality of such decisions are computed with an impurity function (loss function) such as Gini, Entropy or Misclassification.

$$IG(T,a) = H(T) - H(T|a) \tag{5.4}$$

$$H(T) = I_E(p_1, p_2, ...p_i,) = -\sum_{i=1}^{j} p_i \; log_2 p_i \tag{5.5}$$

$$H(T|a) = \sum_{i=1}^{j} -Pr(i|a) \; log_2 p_i \; Pr(i|a) \tag{5.6}$$

where H(T) is the entropy of the root, H(T | a) is the sum of entropy of the children nodes.

The most important and time consuming part of DT learning is to find the best split points with most informative feature (largest information gain). The deeper the trees, the more complex are the learnings. Hence, the models are better fitted. The simplicity of decision tree helps understanding a model with Visualisation tools that can help explainability for otherwise Blackbox aspect of machine learning. However, the learning method has some drawbacks. Learning an optimal decision is usually achieved by creating greedy algorithms that decide on local minima and may not scale for global optimal trees. Decision trees can be easily affected by imbalanced data with one class dominating the others and get unstable with small variation in the data. All of these problems can be mitigated by training multiple trees in an ensemble such as Random Forest, boosting and bagging. Hence we have trained with multiple trees (n $\geq$ 300) and balanced the data with an oversampling method.

Random Forest ensemble

Random forest classifier is a meta-estimator that fits a number of decision tree classifiers on various sub-samples of the dataset. Decision Tree is considered as building block for many ensemble algorithms because it is a non-parametric supervised learning method that learns simple decision rules by inferring from the data features. Each decision tree in the ensemble is built from a sample drawn with replacement from a training dataset. The best split is obtained by including all features or a random subset of size *max-features*.

For prediction accuracy, the Random Forest estimator utilises averaging each classifier's probabilistic prediction. The randomness in forest helps controlling over-fitting of an estimator. The sub-sample size is controlled by the parameter "max−samples" when boosting is employed. In the absence of boosting, the classifier includes the entire dataset for building the trees.

AdaBoosting ensemble

AdaBoosting or adapted boosting is a meta-estimator that uses decision trees for learning. The basic principle of AdaBoosting is to fit a sequence of weak learners repeatedly on modified versions of the data and to obtain a final prediction as a weighted majority vote of all predictions. The strength of AdaBoosting is focusing on harder-to-classify instances as the learnings progress with weak learners. At the start, all samples have the same weight $w_i = 1/N$. For subsequent iterations, sample weights are modified individually and the algorithm learns with reweighed data. At a given step, the training examples that are incorrectly predicted are given more weight than those that predicted correctly. Hence, as the training progresses, the incorrect examples are prioritised with increased weight and learners focus on these most difficult-to-classify examples. This mechanism helps later stage weak learners to work on examples that previous learners predicted incorrectly.

Bagging ensemble

Bagging (**b**bootstrap **agg**regat**ing**) is an ensemble of meta-algorithms that can help improve stability and accuracy of machine learning algorithms. At the core, bagging is a model averaging approach for regression and voting for classifiers. Although bagging is most often used with decision tree learning algorithm, it can work with other learning methods. It helps in reducing overfitting by removing variance in some datasets.

5.5 Results

We have conducted ML classification for identifying the early adopters from the observational data. We employed five algorithms for the corresponding models to train, test and cross validate. In the process, we divided our data in 70 : 30 proportion between training and testing respectively. Hence, we reserved 30% of our randomly selected data for testing (69417) which was unseen by training algorithm. We also performed grid search for hyper parameter tuning for each of the algorithm and it resulted in one *best estimator* with the best possible combination of hyper parameters. However, our random sample size was quite large with $231,390$ unique customer level data. After applying the SMOTE algorithm for balancing the data, with increased sample size, the grid search became computationally expensive. Additionally, for fair comparison purposes, we cross-validated the data for 5 times for each algorithm, since validation after the training is performed to check over-fitting of the model. We couldn't perform $5x2$ CVrule for classifiers because of the increased computational expenses.

For classification purposes, we chose scores such as precision, recall, F1 and confusion matrix for better understanding of the performances. Precision measures the ability of a classifier not to label as positive when it is negative. Recall measures the ability of a classifier to find all positive samples. F1 is the balanced (evenly) weighted between precision and recall that is the harmonic mean of the two. The results of the performance measures are shown in the Tables below.

Table 5.2: Comparative analysis of predictive accuracy among algorithms

Algorithm / Performance	**AdaBoosting**	**MLP**	**Random Forest**	**Bagging**	**k-Neighbor**
Accuracy (Overall)	0.98	0.99	0.92	0.93	0.72
F1 (IG)	0.86	0.96	0.14	0.18	0.83
Precision (IG)	0.95	0.94	0.55	0.84	0.11
Recall (IG)	0.78	0.97	0.08	0.1	0.38

The Table for classification report (see Table 5.2 provides some important insights on the performance measures for evaluation among the algorithms for classifying between innovator group and the other customers (non-innovator group). Precision in Eq. 5.7, Recall in Eq. 5.8, F1 in Eq. 5.9 and Accuracy in Eq. 5.10 indicate the detailed performance.

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive} \tag{5.7}$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative} \tag{5.8}$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{5.9}$$

$$Accuracy = \frac{TruePositive + TrueNegative}{TruePositive + TrueNegative + FalsePositive + FalseNegative} \tag{5.10}$$

Since IG customers are relatively low in any population and hard to identify [203]. Precision is important to identify IG customer and recall is crucial to identify the mis-labelled IG as Non-IG. Therefore, F1 score is a better metric to capture both precision and recall. We have included all of these metrics to provide the overall understanding of the result. Explanability of a ML model is beneficial for human understanding and interpretation, especially for managers who would like to take actions from the models' results. Hence, domain knowledge for interpreting results is a crucial way to the reduce Blackbox. From the above results, managers can decide whether losing some IG customers is more expensive than adding some Non-IG customers in the selection. This can guide their decision to choosing between AdaBoosting (95% i.e. higher precision) and MLP (97% i.e. higher recall).

Overall MLP and AdaBoosting perform best among all the algorithms, in terms of overall accuracy, precision and recall. This indicates the inherent strength for neural network learning for MLP and ensemble with decision tree for AdaBoosting respectively. The predictions for MLP and AdaBoosting algorithms as shown in Table above are 0.99 and 0.98, which means they are able to select both innovator group and non-innovator group correctly for 99% and 98% of times in the testing data. kNN has identified Non-IG with more precision than the IG group which may indicate that the best estimator with 6 neighbours couldn't locate all IG customers in the test data. Similarly, Random Forest algorithm performed well in the precision for Non-IG customer identification but it lacked identifying the IG customers with lower precision (0.55). This may be a result of over hyperparameter tuning with 180 combinations for the GridSearchCV. On the other hand, Bagging algorithm performed well above both Random Forest and kNN with better precision ratio of 0.84.

Table 5.3: Confusion matrix comparison among algorithms

Algorithms		**AdaBoosting**		**MLP**		**Random Forest**		**Bagging**		**kNeighbor**	
		Prediction		**Prediction**		**Prediction**		**Prediction**		**Prediction**	
		Non-IG	**IG**	**Non-IG**	**IG**	**Non-IG**	**IG**	**Non-IG**	**IG**	**Non-IG**	**IG**
Actual	**Non-IG**	63721	212	63598	335	63571	362	63823	110	47727	16206
	IG	1186	4298	165	5319	5043	441	4923	561	3386	2098

5.5.1 Statistical Analysis

In machine learning literature, the debate over whether the performance between algorithms can be statistically tested is contested. When drawn from the same test data, due to k-fold validations in ML, the samples are not independent after the training and testing. Additionally, the feature variables may not adhere to normality principle which is a pre-requisite for a parametric statistical testing. Although non-parametric testing is less powerful than its parametric counterpart, it can show the statistical difference between the models in terms of homogeneity or the proportion of errors. Additionally, the cost of computation for a large dataset is high and in our case, we couldn't perform multiple runs with GridSearchCV and cross-validation for the algorithms. Hence, considering our case where we did only 5-fold cross-validations for each algorithm, we have chosen McNemar non-parametric testing to analyse the differences [64]. The null hypotheses of marginal homogeneity states the marginal probability of each outcome are the same [169]. the Hereby, we formulate corresponding hypotheses as the following:

- H_{1a}: Marginal probability of proportion of errors for MLP and Bagging classifier are same.
- H_{1b}: Marginal probability of proportion of errors for MLP and AdaBoosting classifier are same.
- H_{1c}: Marginal probability of proportion of errors for MLP and Random Forest classifier are same.
- H_{1d}: Marginal probability of proportion of errors for MLP and KNeghbour classifier are same.
- H_{1e}: Marginal probability of proportion of errors for Bagging and AdaBoosting classifier are same.
- H_{1f}: Marginal probability of proportion of errors for Bagging and Random Forest classifier are same.
- H_{1g}: Marginal probability of proportion of errors for Bagging and kNeighbor classifier are same.
- H_{1h}: Marginal probability of proportion of errors for AdaBoosting and kNeighbor classifier are same.
- H_{1i}: Marginal probability of proportion of errors for AdaBoosting and Random Forest classifier are same.

- H_{1ja}: Marginal probability of proportion of errors for Random Forest and kNeighbor classifier are same.
- H_{2a}: Marginal probability of proportion of errors for MLP and Logistic regression classifier are same.
- H_{2b}: Marginal probability of proportion of errors for AdaBoosting and Logistic regression classifier are same.
- H_{2c}: Marginal probability of proportion of errors for Bagging and Logistic regression classifier are same.
- H_{2d}: Marginal probability of proportion of errors for Random Forest and Logistic regression classifier are same.
- H_{2e}: Marginal probability of proportion of errors for kNeighbor and Logistic regression classifier are same.

The results from the Mcnemar's tests shown in the Table 5.4 informs that there exist statistical difference between the errors of each algorithm. Additionally, logistic regression that was used as the benchmark classifier to contrast ML classifiers, found to be different for each of the comparisons. This indicates that the performances among the ML algorithms are different even when they were pooled from the same Big-Data. Additionally, MLP and AdaBoosting algorithms perform better than the logistic regression in the selection of the IG customers.

5.6 General discussion

We are addressing two key questions in this study. The first question addresses a new way to classify Innovator group customers from BigData, in a natural experiment setting. The second question addresses the performance of machine learning methods as compared to multiple good algorithms, including some traditional methods.

Our first research question pertains, "How to identify Innovator Group customers from the Big-Data?" Broadly, we propose a framework by combining theoretical constructs with constructs derived from the real-data. The combined predictive variables capture IG customers' behavioural information that manifests in the transactional data. The relationship among the variables contains traits that are particular to IG customers. Our primary theoretical insight is built upon the collective knowledge from the previous

Table 5.4: Comparative analysis with Mcnemar's test

Algorithms	Null hypothesis	Statistic	p-value	Results
MLP vs Bagging	H1a	330	.000	Rejected
MLP vs AdaBoosting	H1b	306	.000	Rejected
MLP vs Random Forest	H1c	327	.000	Rejected
MLP vs kNeighbor	H1d	290	.000	Rejected
Bagging vs AdaBoosting	H1e	212	.000	Rejected
Bagging vs Random Forest	H1f	284	.000	Rejected
Bagging vs kNeighbor	H1g	1911	.000	Rejected
AdaBoosting vs kNeighbor	H1h	586	.000	Rejected
AdaBoosting vs Random Forest	H1i	249	.000	Rejected
Random Forest vs kNeighbor	H1j	2074	.000	Rejected
MLP vs Logistic Regression	H2a	191	.000	Rejected
AdaBoosting vs Logistic Regression	H2b	1125	.002	Rejected
Bagging vs Logistic Regression	H2c	1179	.000	Rejected
Random Forest vs Logistic Regression	H2d	1182	.000	Rejected
kNeighbor vs Logistic Regression	H2e	924	.000	Rejected

research and we extend the knowledge by adding weights to the variables. We have employed machine learning algorithms to predict the innovator group customers with reasonable precision, by employing the best practices of cross validation, GridSearchCV for hyperparameters and balancing the dataset. With this result, we can show how the investigated factors actually help identify IG customers in real life decision-making where the new customers share the same traits with the other IG customers.

Our second research question probes, "Which is the optimal algorithm or method for Innovator Group's identification/classification?" To answer this question, we compared different machine learning algorithm's performance for prediction using different metrics. Since, the aim of the algorithms was to classify IG customers, the specific performance scores were appropriate for the comparative purpose.

Our research makes two broad contributions to the literature. First, from new product innovation perspective, we prescribe a quantitative way to classify and identify early adopters from the BigData, without assorting to field experiments. In this way, customers are in their natural environment, unobserved and act without any explicit manipulation, Moreover, firms can understand these customer's decision making unobtrusively without explicitly intervening. This approach also helps firms in their decision to plan/design interventions, with optimal effectiveness for the new products by involving IG customers for cocreation or for diffusion of information.

Second, from a methodological perspective, we present a framework that managers and researchers can use to evaluate the effectiveness of early adopter's selection of new products. Although our findings on relative performance of the estimators are context dependent, the innovator group framework can be useful in general contexts where irrespective of product category, firms can identify the customers from the Big-Data.

Additionally, explainability is an interesting and relevant question for ML research. We have incorporated domain knowledge of Marketing from literature and expert knowledge to reduce the Blackbox syndrome from the ML models. This is in line with Roscher et al.'s suggestion of "domain knowledge" in contributing to "explainability" that in turn helps in interpretability and transparency in ML [207]. However, the full explanation of the ML models is not entirely possible and the right step has been taken in this study to incorporate domain knowledge in the feature selection, algorithm knowledge and hyperparameter selection in this study [207].

5.6.1 Implications for managers

By showing that a combination of expert-knowledge and supervised ML algorithms can identify IG customers better than other classification methods, the study shows that there is potential for managers and academics to develop new models for predictive purposes. The research problem of identifying and predicting IG customers from the past customer behaviour has many implications for managers. First, when managers design targeting policies or utilise marketing resources to attract IG customers, they have better idea about these customers. Identifying the common characteristics for each category of products can also help in creating marketing campaigns for acquisition and/or retention of customers with similar traits. Managers can also design new products according to these early adopters' preferences by distilling information from their past products' trends. Additionally, by incentivise or creating loyalty programmes, the IG customers can be involved in diffusion of new product information to the population. Brand ambassadors, brand community leaders or expert customer forums can be created to nurture and nudge these customers in spreading information. For managers, understanding early adopters' needs and decision making behaviour can help them to create desirable/compatible new products for these customers. In doing so, they may tackle the high-rate of new-product failure phenomenon.

With identification of the right group of customers, product managers can have multiple options to design their ideation strategy or marketing promotion strategy. The knowledge on their specific IG groups can help them further creating innovative products and targeting to this group. Managers can also direct their targeting and evaluation efforts towards IG group for new products, with expected better feedback. Also by identifying the potential early adopters, managers can nudge these customers with/without incentives to spread the information on new products. Other implications can be derived in the area of personalised product offerings for these customers. Involving IG customers in the new-product development process - ideation and/or launch phase - could lead to more customer-centric products. Altogether, managers can benefit by involving IG customers in various ways to reduce new-product failure and in the process, improve their firm's profitability.

Additionally, aligning with manager's goal for the project on IG customer - cocreating new products, testing prototypes, diffusion of information - selection of appropriate ML models can be decided. For e.g. for cocreation purpose, the higher F1 is crucial to identify IG customers with the models with higher F1 value (closer to 1.0). This score indicates that most IG customers can be located, and then they can be approached for the cocreation engagement. For the post-launch information sharing purpose, a higher recall score may be desirable as new product's information reaching IG and some Non-IG customers may not have any adverse effect for the marketing campaigns.

5.6.2 Implications for academia

For academia, the study presents a way to integrate (supervised) machine learning algorithms with real customer data in developing a framework to identify and predict IG customers. Even though statistical models perform better with a fixed set of assumptions, with increased complexity especially with human behavioural data, supervised machine learning algorithms perform much better in handling the vast amount of information, correlate/covariate relationships between factors that are hard to account for in the traditional techniques. Additionally, the vagueness and changing nature of the customer can also be extracted by a natural experiment setting.

The study shows how to identify Innovator group customers from a real dataset by adapting a number of supervised machine learning algorithms. We also aimed to create a data-driven decision-support system from learning and implementing a real-world project. Although synthetic data has been used extensively before in the research, working on real behavioural data showcases the practical problems that needs to be solved before addressing any particular classification problem. Additionally, a combination of human expertise in feature selection and weight calculation along with selection of suitable ML algorithms makes the task more robust than either machine or human efforts separately.

Managers in consumer goods industry and/or other industry who are grappling with high new-product failure will find value from this study. We showed a way to combine human and machine proficiency to achieve a difficult task of identifying Innovator group customers from the real-data. Since, humans change their behaviour constantly, we have taken this aspect into consideration while developing an intelligent system with real data that can learn, identify and predict future IG customers. Besides the implementation method for ML, the results also have implications for the practitioners.

5.7 Future research and limitations

We have combined a number of techniques and algorithms to conduct the research to identify early adopters from anonymised structured data. However, there are some limitations to our approach. First, the data was from a single manufacturer in Spain where the registered users were self-motivated to participate in the membership programme. In many ways, these customers are similar to the survey/interview participants who willingly participate in the surveys. Second, we have combined different machine learning algorithms to identify IG customers, which was a complex method to obtain. Managers may chose to employ unsupervised machine learning techniques to distil the information from the Big-data, not guided by the existing research. Additionally, there exists no single algorithm or ensemble that can outperform all other existing algorithms as demonstrated by the previous research. Hence, the results we showed is a demonstrated way to create customer profile from the Big-data. We suggest future researchers may find better ways to obtain general classification to be applied to many different organisational goals and strategies. Third, the findings we report are derived from the natural experiment setting with a single manufacturer in Spain. It may be true that the innovative data collection method used by the firm may not be possible for other manufacturers. In particular, the natural experiment with new product launches was designed for retaining customers. Because of this organisation's goals, we could use the setting for identification and classification of IG customers and this may not occur for other settings.

Lastly, the proportion of IG is very low in the entire dataset, which is in accordance to Roger's model [205]. Therefore, low representation of IG customers may affect the distribution of the target variable, and this may affect the performance of various ML methods used for training. We applied stratified sampling to mitigate the effect, in addition to that we also used SMOTE algorithm to balance the data. Therefore, managers need to pay attention to the data imbalance issues before drawing any conclusions.

For future researchers, we suggest some directions to take forward the research on cocreation with customers and NPD research in general. First, after identification of IG customers, research can investigate the overall ROI with IG and lead users, ENC and market maven. Some studies have looked into cocreation knowledge from lead users and others [163], however, a comparative analysis of the ROI - cost and benefit - can show which select group performs best for the entire NPD process - idea generation to post-launch diffusion. Second, future researchers can look into unsupervised ML techniques to find faster, better ways to identify select customers from their usage Big-data. A further comparative analysis can reveal how different ML techniques can augment or diminish the selection process. Research can also explore in which stage, the innovator group is most valuable to the NPD process. Additionally, the crowdsourcing researchers can delve deep into the role of artificial intelligence in the NPD process and find out how AI affects the final new products. A comparative analysis on the effectiveness of AI at selecting the right customers, AI filtering the right ideas or AI in dissemination of information on new products. Lastly, combining social media data, when legally available, can also be added to the database to include perceptual factors for the IG customers. Though it is considered hard to obtain the RPIT factors unobtrusively, future research can incorporate novel techniques to mitigate this shortcoming.

Appendix

Table 5.5: Ranking of IG customer's adoption factors with FTOPSIS (Euclidean distance)

Category	Variable_name	Closeness coefficient	Rank (Preference order)
RPIT	Perceived Brand image	0.0991	1
RPIT	Perceived brand trust	0.1337	2
Product Trait	Visual appeal	0.1340	3
RPIT	Perceived usefulness	0.1496	4
Product Trait	Innovativeness of product	0.1521	5
RPIT	Expectations	0.1705	6
Individual Trait	Personal Innovativeness	0.1794	7
Product Trait	Attributes of product	0.1819	8
Individual Trait	Novelty seeking attitude	0.1865	9
RPIT	Perceived ease of use	0.1869	10
Individual Trait	Variety seeking	0.2025	11
RPIT	Involvement	0.2160	12
Product Trait	Relative advantage	0.2182	13
RPIT	Satisfaction while comparing products	0.2343	14
Individual Trait	Income	0.2343	15
Individual Trait	Social Image	0.2528	16
Product Trait	Trialability	0.2552	17
Environment	Social norm	0.2575	18
Individual Trait	Knowledge sharing	0.2688	19
Product Trait	Observability	0.2690	20
Environment	Network externality	0.2849	21
Individual Trait	Pleasure seeking attitude	0.2873	22
Individual Trait	Impulsiveness	0.2921	23
Environment	Social approval	0.3054	24
Environment	Mass media influence	0.3078	25
Product Trait	Availability of product choice	0.3264	26
Individual Trait	Aesthetic value	0.3349	27
Individual Trait	Status signalling	0.3371	28
Product Trait	Product type	0.3397	29
RPIT	Previous experience	0.3425	30

Chapter 6

Conclusion

This chapter presents an integrated discussion of the Theoretical contributions, managerial implications, limitations and future research directions for the chapters 3, 4 & 5.

6.1 Theoretical contributions

The PhD thesis has addressed the challenges and opportunities that new product innovation is facing in the recent times, where new products have increasingly become collaborative projects with the customers. However, managing and creating sustainable ecosystem for new product innovation has become a challenge for managers, also presenting opportunities to improve the NPD process with creative solutions. On one hand this PhD thesis has contributed to the field of cocreation and crowdsourcing. On the other hand, bringing forth big-data and AI techniques to address the challenges, this PhD thesis has contributed to the field of technology management for innovation, which has just begun to explore AI.

6.1.1 Theoretical contribution to the field of cocreation

Increasingly, evidence suggests that for new product innovation, the locus of innovation has shifted from organisations to customers [97, 168]. To substantiate the shift, customer participation research has exploded since early 2000 [195, 196, 254]. The result of customer's involvement at various stages of NPD has contributed to organisations for increasing their market and social welfare [83, 102]. On one hand, customers are improving the ideas for new products, better than traditional market research or internal employees [274] by motivations originating from philanthropy, volunteerism, self-efficacy, financial gain and social status [81, 225].

On the other hand, finding customers who can participate and create good quality, feasible and realisable ideas has posed challenge to the firms. lead users, emergent nature consumers and innovator group customers have been recognised as key individuals who can generate new ideas and disseminate information on the new products to later adopters [11, 12, 97, 107, 203]. Some research have conducted netnography, pyramiding and screening techniques with surveys to capture potential innovators. However, despite their contribution and usage in new product innovation process, identification of these customers remain highly unstructured. Surprisingly, very few researchers have paid attention to the core of cocreation process i.e. selection of right customers. Hence, the research gap is relevant not only for academic researchers but also for the managers responsible for the idea generation or pre-launch stages in the NPD process.

We have addressed the problem by approaching it from adoption decision-making perspective. Scholars argue that lead users are better at creating completely new products themselves [159, 160, 257, 260], but they are hard to locate. Additionally, they are years ahead of the regular customers in terms of preference and need and often don't represent the population's taste and preferences in the same time-period [159, 257]. On the other hand, emergent nature consumers are those who imagine and recreate new usage of the existing products, after the original products are available in the market. However, so far the identification is done with surveys and depends on self-declaration and self-selection methods [11, 97, 107, 160, 257].

Hence we proposed to identify innovator group customers who are easier to locate with their purchase behaviour as the first step. Also innovator group falls between the lead users and the majority in terms of domain knowledge, preference and compatibility. Although scholars have studied these customers (early adopters) for decades, they have not been thoroughly categorized for their specific characteristics. Since these customers are knowledgable, early adopters of new products and can influence later adopters, understanding these customers can help identify them from the population. And they can be engaged with the new product innovation process rather systematically than before.

The reasoning for the approach was to understand what motivates innovator group customers to adopt new products. Although previous research studied innovators (early adopters) for a number of products, technology and settings (national, culture, age). However, a meta-analytic approach in combining all the factors that affects the customer's adoption decisions has not been conducted. We contribute to the cocreation literature by proposing a starting framework where the adoption of innovator group customer is organized according to four categories. This helps in advancing knowledge on innovator group with their most investigated factors of influence.

In chapter 3 and 4, we created a knowledge database that provides a list of cumulative factors affecting the innovator group customers. We not only gathered most of the key factors from a systematic literature review, but also accessed these factors' importance. We took a different new approach than meta-analysis, and employed industry experts for validating the factors. In this way we aimed to extract knowledge from the industry experts who have been launching new products for consumer goods industry (average number of years is 10). In doing so, we addressed the problem of finding out weight of each factor, without collecting large scale customer's surveys. We combined expertise from the managers and utilising fuzzy logic, extracted tacit and imprecise knowledge into clearer information to rank the most important factors for identification of innovator groups. The customers can be engaged in either exploration or exploitation purposes more effectively and the firms can also benefit from the customer preference knowledge extracted for the core concepts or for the prototype selection, with a manageable number of ideas.

The final ranking reveals the most crucial factors of IG customers' adoption on new products tend to be perceptual, visual and innovation driven. The findings has implication for academia in NPD and cocreation research areas. For the NPD research, the major contribution of the PhD thesis is highlighting the ranking of IG customer's adoption decision-making factors and their importance, in particular for three phases: *idea generation*, *prototype testing* and *launch* phases. For idea generation, customer participation has proven to be efficient for select customers (lead users, innovators, emergent user consumers), and understanding their underlying needs and preferences can help create new products. From our ranking of adoption factors for IG customers, customer participation activity in cocreation and NPD can be planned with more focus on exploiting perceptual and individual trait factors, as a starting guideline to create new features that are perceived better by the customers.

From the study in Chapter 4, we also contribute to the NPD research, in particular to the prototype testing phase by highlighting the most critical of product related factors such as, visual appeal, innovativeness of product, new attributes, relative advantage and chance to trial before launch. We also contribute to the launch phase by highlighting key factors such as perceived brand image, brand trust, perceived usefulness and innovativeness of products that most affect IG customer's adoption. Hence, not only we gathered most of the relevant factors that affect customer's adoption decision, but also by ranking the factors, we contribute the understanding the effect of these factors on the IG customers.

6.1.2 Theoretical contribution to the field of crowdsourcing

Gradually scholars have showed that customers as a crowd can contribute to the idea generation, and prototype testing [1, 109, 217]. The contribution of collaborative new ideas have resulted in interesting new products that firms have launched in the market. Additionally, integrating customers and their ideas during the NPD process with the internal team - who focuses on intra-team collaboration, team cohesion,task - has created lower than expected innovation outcomes [246]. Also, the goal of crowdsourcing is to find some individuals who have the best solutions to the broadcasted problem [22]. There remains a gap in systematic method to identify a select crowd that has higher chance to find solutions in their local neighbourhood than a crowd of dispersed search locals [1].

Customer classification is not a new problem that academia and managers have faced before. Previous studies have addressed the issue by utilising techniques such as *screening* (with surveys) and *netnography* (with online community), pyramiding, and crowdsourcing. These techniques help filter the customers who are proactive in participating in either surveys or in online communities. On the other hand, specific customer user groups such as lead users and emergent nature consumers have been studied extensively and their characteristics have been identified by quantitative (survey questionnaire) methodology.

However, for managers, who have access to large customer databases that comprises granular user purchase history, social media information and transactional data, are under-utilising the rich information with application of old techniques. By introducing supervised machine learning technique's predictive powers, we have shown that the selection of right customer can be approached by some effective ML techniques.

6.1.3 Theoretical contribution to the field of group decision making

In the chapter 4, we showed that a hybrid group decision making methods' can be used to a real world problem where the methods are adapted to that specific problem i.e. not retrofitting the methods to the problem. Our contribution lies in the adoption of FTOPSIS and FAHP method to solve a marketing problem which is framed in a multi-criteria group decision-making context. We provide a different approach i.e. fuzzy logic based group decision making methods to identify IG customers' adoption factors with industry experts' knowledge. We also showed that operation research methods can be effectively used in the new product innovation domain while capturing human expertise that is embedded in uncertainty and tacit knowledge.

Additionally, we have demonstrated that for group decision making scenarios, where experts don't interact among themselves (in our study, experts belonged to separate firms), fuzzy hybrid techniques can be successfully applied for selection and ranking. It ensures that result can be obtained without risking forced consensus or lowering the quality of decision even with divergent opinions [9]. Hence, we contribute to the group decision making literature.

In the chapter 4, we also show that with industry experts, ranking and selection can be achieved for a number of important topics. We showed that for finding out importance of the adoption factors for the innovator group, the new product launch managers could provide their tacit and imprecise knowledge in terms of the choices. Then we could translate them with fuzzy decision making techniques. In doing so, first we indicate that a diverse group of experts from various organisations can come to a convergence on solutions without interacting with each other. Second, we also showed that managers like customers have vast amount of tacit and operational experience knowledge that they may not explicitly access themselves. With the fuzzy AHP for comparative analysis and with the fuzzy TOPSIS, with each expert's own measurement range, we successfully found out the ranking with weights for the adoption factors. This is one of the many applications that managers' knowledge can provide. For example, in future, similar to crowdsourcing for customers, there can be inter-organisational *crowdsourcing with managers* for accessing, ranking and evaluating a number of important issues in business and society. The aim will be to gather as diverse fields of knowledge as possible but the members will be experts in their own domains who without disclosing proprietary knowledge can help finding solutions to some difficult problems.

6.1.4 Theoretical contribution to the field of technology management with AI

This thesis takes a new perspective on knowledge aggregation from experts and customer analytics and adds new knowledge in identifying innovator group customers from the structured data. We also propose a new select (innovator group) crowd for the cocreation in the NPD process. Moreover, customer behaviour has dramatically changed with the advent of technologies such as internet, mobile phone, personal computers, GPS, Internet-of-Things, to name a few. The innovative firms who innovate and try to sell new products to customers have great difficulty in knowing the changing customers' needs and preferences. Therefore, the origin of customer participation research arose to fulfil the gap in knowledge for the firms: *only if customers can tell them what they really want and need*. The puzzle of customer's need recognition or preference knowledge has employed various methods including conjoint analysis and machine learning for finding relationships and patterns that even the customers are themselves unaware of.

Since cocreation and crowdsourcing literatures are facing challenges with regard to manage the crowd's ideas, selecting and evaluating the conceptualisation and pre-launch phases with the internal teams, we address the core of the problem: *finding the right customers*. In Chapter 3, we collected the most investigated factors (103) on innovator group customers and proposed a framework that categorized the factors and linked to the NPD process. In chapter 4, we sought industry experts' knowledge to rank these factors to gain deeper insights. In chapter 5, after finding out the weight structure of the factors, we employed this information in a real case. From a longitudinal data of four years on customer's purchase, we mapped these factors to each customer and predicted their probability of belonging to the Innovator group.

In the Chapter 5 of this thesis, we propose a framework based on Artificial Intelligence techniques to identify and predict future innovator group customers from the customer data. The methodology combines a number of algorithms to train, validate and predict for the classification of these early adopters. The rigour of the method suggests a combination of AI algorithms that can be applied by businesses who have some form of data collection on their customers. Since no single algorithm can be considered performing best among different algorithms in AI, we have compared the prediction accuracy to have an informed perspective about the efficacy of the algorithms. We also test the statistical analysis of the models' validity and checked whether they are statistically different from each other. Results show that even if they are pooled from the same data, cross-validated a number of times, the predictions are not same (they have different proportions of error). Theoretically, we also suggest the efficacy of the framework and its performance with existing methods. We analyse comparative performances for various ML techniques and explore the challenges and possible application of real-world data on firm strategies.

6.2 Managerial implications

The findings from chapters 3, 4, and 5 of the PhD thesis has several implications for the managers. In summary, for cocreation with customers, identifying the innovator group customers for conceptualisation and pre-launch product selection in a cost-effective and time-efficient way has value for the managers.

Chapter 3 shows that by understanding adoption decision-making factors of innovator group customers may help managers evaluate customer preference better and accordingly include those features in the new products. The findings may also help understanding the complex adoption decisions that influence these customers. Since, personal characteristics are not the only relevant information, rather interactions with product equally affect IG customers. This information is crucial for cocreation with these customers as they can be the perfect select-crowd, who have domain knowledge and get affected by the relationships they form with the products. For prototype testing in particular, IG customers can be more effective in selecting the products that will be liked by rest of the customers. Hence, understanding IG customers can help managers in many ways: *preference disentanglement for new products*, *prototype selection* and *information dissemination* to wider customer base. In doing so, we help focusing on selecting a relevant crowd (innovator group customers) in a systematic manner that managers can select from their own database. Therefore, the intricate information we found in this study can help managers plan for more successful new product creation with matched customer preferences and well-disseminated product information among the population.

In this PhD thesis, the overall aim is to provide paths for managers to take micro-level actions [270] grounded in theory-based conceptual framework to manage NPD performance. Our study guides managers in planning and executing concerted actions towards attracting IG customers and also approaching them for cocreation process. Managers can design their new products on the basis of i.e. how IG customers perceive usefulness, innovativeness, attributes and ease of use, and how IG customers get influenced by visual appeal of new products, trial offers, variety in products and hedonistic aspect of the products. These factors together can provide a guideline for managers to design their new products to be attractive, fun, wider variety and easy to use. In addition, customers' participation in prototype or pre-launch activity in the NPD process can be purposefully targeted towards understanding perception about products and can help managers design the new products according. These factors can also guide NPD managers on strategic decision for developing prototypes with the customers, by giving them opportunities to try new products and get their feedback on perceived usefulness, ease of use, expectation, satisfaction, functionality, and hedonistic or utilitarian aspects of the product. Altogether, managers can create new products with the IG customer's feedback, choice patterns and participation in post-launch process that may become more successful than without the knowledge on these customer's adoption process.

IG customers' adoption decision making is also influenced by how they perceive brands (positively or negatively) and that affects the likelihood of their purchase from the brand. We suggest that firms need to build trust and brand image with new attributes, innovation and perceived usefulness of products, We also suggest firms to design marketing campaigns and brand communications to attract these domain specific expert customers (i.e. IG) who like to explore new products through variety and innovativeness. The collected information can be insightful to the managers for learning about new products' actual reception in advance and can help them take necessary steps to improve prototypes or communication strategy before launching new products in the market.

For post-launch success, IG customers are proven to be opinion leaders and influencers who disseminate information on product and signal low risk with their high usage of the new products. These customers signal social status through tacit knowledge to late adopters. Managers should recognise IG customers' role as expert influencers and incorporate communication strategies to involve them. Notwithstanding which way IG customers influence - imitation or risk mitigation - managers should utilise these customers' information sharing capability, word-of-mouth, and high status in the social network to disseminate positive product information. However, managers should be aware of the interplay between social network and mass media communication for new product diffusion. For example, it is observed that increased advertisement reduces online-word-of-mouth [79]. Therefore, managers should take a balanced approach to employ multiple tools to induce diffusion of new products with the IG customers.

From the chapter 4, we showed a way to harness knowledge from experts who have diverse backgrounds and experience in their fields. For a large organisation, top managers can solve some inter-functional problems by a) utilising the implicit knowledge the firm's managers have without even mobilising them together for brain-storming. This can act as a first step in finding solutions without forced consensus based on norm or group thinking. b) the active applications of expert's knowledge can be a repository from where AI algorithms can learn the underlying relationships and paths. By doing so, organisations can conserve experts' knowledge and find out new ways to solve organisational problems.

From the chapter 5, we showed how ML algorithms can be utilised for identifying and predicting IG customers. In doing so, we suggest managers can benefit from selecting the appropriate models that align with their specific strategies. For e.g., a firm that aims to find IG customers may chose a model with a higher F1 score. Whereas when a firm's goal is to spread new product's information to wider population, then with a moderately higher precision (80% to 90%) may work well for their purpose, as precision includes some mislabelled Non-IG customers. Therefore, the managers can decide which ML models suit their specific strategies, and they can decide on the implementation of information from those models.

Altogether, for the generalizability of the findings, this thesis can help managers to manage a better designed NPD process. With some digitalization in place for collecting customer behavioural or transactional and demographic information, any consumer oriented firm can apply our framework with supervised machine learning techniques for identifying their own IG customers. With the most crucial factors identified, managers can design their new products with the important factors identified and can pre-test with a small group of customers. The overall information on the influencing factors and on selection of IG customers for co-creation can help managers in consumer goods and other sectors, where the end-customers are also individual consumers who go through adoption decision making process (excluding B2B customers) for new products.

6.3 Limitation and future research

Despite our best efforts, the study has some limitations. In chapter 3, our study is based on 72 studies that we included in the systematic literature review after careful selection process. We have proposed a *conceptual framework* not a *theoretical framework* to understand the failure phenomenon of new-product failures from customer adoption perspective in consumer goods sector. Therefore, our research may lack generalizability, especially in the industrial goods market.

Next, we did not include any peer reviewed research before 1987 because we decided to focus on the changing landscape of adoption of innovation over the last three decades, and especially on current practices in academia and industry. Nonetheless, we may have missed some important variables prior to 1988 or post 2017. However, these shortcomings may have been mitigated by following the best practices of academic rigour in the systematic literature review [180, 245]. We considered only published and peer-reviewed articles from multiple disciplines. The information lost may be in line with the accepted standard in the literature review process that ensures only the best knowledge sources are included [237]. Furthermore, our framework offers a theoretical linking of IG customer feedback to NPD process, and it does not invalidate other arguments that IG customers are drastically different to majority customers. Generalizing IG customer preferences for the entire population may not work for all products. Nonetheless, there is no empirical evidence supporting the claim, even in high-technology sectors [49, 205].

Another caveat to our framework is that we did not consider the cost of collecting data on IG customers. We assume that in the Big Data era, most firms have systems to collect customer usage and buying-behaviour information. However, a study on big data by McKinsey [24] showed that discrepancy in data collection and storage exists among geographical regions. Hence, managers should be aware of the effort and cost of collecting data for using the proposed framework across countries. Also, our conceptual framework links social contagion mechanism for the diffusion of new products for NPD performance that hinges on the premise that IG customer can influence other customers.

Additionally, the framework assumes IG customers would positively respond to actions by firms by providing feedback. However, it does not explain scenarios where they are non-responsive or non-reactive to these efforts by the firms.

In the chapter 4, the number of experts selected for the hybrid FAHP-FTOPSIS method was small. Though the diversity in location, nationality and experience mitigates some of the shortcomings. Nonetheless a larger sample of experts could have presented better results. However, extant literature in GDM has shown that when the level of expertise is higher, a small number of experts are effective as a group for decision making purposes [249]. Next, all the experts belonged to consumer good sector in the study which may have overrepresented the sector's idiosyncrasies. However, consumer good firms operate in a highly competitive environment and the managers may have more knowledge on new product launches than other less competitive sectors. Additionally, comparative analysis of alternative combination of techniques is beyond the scope of this study. We have combined a number of techniques and algorithms to conduct the research to identify early adopters from anonymised structured data. However, there are some limitations to our approach. First, the data was from a single manufacturer in Spain where the registered users were self-motivated to participate in the membership programme. In many ways, these customers are similar to the survey/interview participant who willingly participate in the surveys.

In the chapter 5, we combined different machine learning algorithms by applying it to a real data to create a composite dependent variable with recency, frequency and quantity variables to define the IG customers from the Non-IG customers. Managers may chose to employ unsupervised machine learning techniques to distil the information from data, not guided by existing research. Additionally, there exists no single algorithm or ensemble that can outperform all other existing algorithms as demonstrated by some previous research. Hence, the results we showed is a demonstrated way to create customer profile from the data. We suggest future researchers may find better ways to obtain general classification to be applied to many different organisational goals and strategies.

In the chapter 5, the findings we report are derived from the natural experiment setting with a single manufacturer in Spain. It may be true that the innovative data collection method used by the firm may not be possible for other manufacturers. In particular, the natural experiment with new product launches was designed for retaining the customers. Because of this organisation's goals, we could use the setting for identification and classification of IG customers and this may not occur for other settings.

Lastly, the proportion of IG is very low in the entire dataset, which is in accordance to Roger's model [205]. Therefore, the low representation of IG customers may affect the distribution of the target variable, and this may affect the performance of various ML methods used for training. We applied stratified sampling to mitigate the effect, in addition to that we also used SMOTE algorithm to balance the class data. Therefore, managers need to pay attention to the data imbalance issues before drawing any conclusions, especially while designing a dynamic intelligent systems.

6.3.1 Future research directions

For future researchers, we have several directions in research areas of cocreation, crowdsourcing, group decision making and applied AI.

On the basis of the expectations and limitations of our conceptual framework in chapter 3, we suggest several directions for future research in the NPD performance and customer-adoption research domains. First, innovation and marketing literature has begun to study the inclusion of customers, including the lead users, into the NPD process with good outcomes [43, 124, 182, 195]. This on-going line of research may be continued to explore the effect of including IG customers at either ideation or pre-launch stages of NPD process to generate customer-centric ideas and improve prototypes respectively. Based on the current work, we expect that IG customers' participation has a greater effect on NPD performance than the average customers' participation. This proposition can be examined in a future empirical study. Future research could also look at IG customer participation at specific stages. For example, speed-to-market moderated by IG customer participation at the ideation stage may increase NPD (financial) performance more than participation by regular customers.

Second, the aggregated framework posits the possible impact on NPD performance when considering feedback from IG customers on consumer goods. It will be interesting to explore the difference in performance between consumer goods and industrial goods when IG customers are involved. Furthermore, not many studies have examined IG customers' negative social contagion effect on diffusion of new products. Current research focuses on positive word-of-mouth and neglects the importance of negative effects when IG customers impede NPD performance. To what extent does negative word-of-mouth by IG customers disrupt diffusion of a new product? Future researchers could also explore the effect of lack of IG customers' opinion leadership for new products. It would be interesting to investigate both of these mechanisms' impact on the NPD performance.

In chapter 4, the application of multi-criteria decision making, group decision making, and several other operational research methods into the management research could be an interesting way forward, following our initiative. Since, there is no unique way to solve a complex multi-level organisational problem - where human expertise is imperative in solving but is often accompanied by ambiguity and uncertainty - fuzzy logic based group decision-making methods can be utilised for solving future organisational problems in marketing and strategy. Based on our result, real world applications for new product development - ideation, prototype testing and launch - can be conducted with data collected from the firm and experts. Additionally, in the future, validation of our framework can be conducted with customer data.

Furthermore, future researchers can apply a combination of group decision making techniques where expert knowledge is required without direct interactions among them. Some possible application cases include, ranking of top restaurants by diverse food experts, selecting awards by film critics, open source community consensus on new standard library and ranking of educational institutions to name a few.

In chapter 5, we applied supervised machine learning methods to identify innovator group customers from the structured data. In the future studies, researchers could take some of the following paths. First, with unsupervised machine learning, researchers can explore the intrinsic correlations among the customer's behaviour and firm interventions. Studies can look into effectiveness of innovator group's cocreation and compare with lead users, ENC, market mavens and regular customers. Research can also explore in which stage, the innovator group is most valuable to the NPD process. Additionally, the crowdsourcing researchers can delve deep into the role of artificial intelligence in the NPD process and find out how AI affects the final new products. A comparative analysis on the effectiveness of AI at selecting the right customers, AI filtering the right ideas or AI in dissemination of information on new products. In doing so, cocreation and crowdsourcing research fields with AI applications may make the research more enriched than before.

References

[1] Allan Afuah and Christopher L. Tucci. Crowdsourcing As a Solution to Distant Search. *Academy of Management Review*, 37(3):355–375, 2012.

[2] Ritu Agarwal and Jayesh Prasad. A Conceptual and Operational Definition of Personal Innovativeness in the Domain of Information Technology. *Information Systems Research*, 9(2):204–215, 1998.

[3] GEORGE A. Akerlof. The Market for Lemons: Quality uncertainty and the market mechanism. In PETER Diamond and MICHAEL Rothschild, editors, *Uncertainty in Economics*, pages 235–251. Academic Press, January 1978. ISBN 978-0-12-214850-7.

[4] Omar Y. Al-Jarrah, Paul D. Yoo, Sami Muhaidat, George K. Karagiannidis, and Kamal Taha. Efficient Machine Learning for Big Data: A Review. *Big Data Research*, 2(3):87–93, 2015.

[5] Ian Alam. Removing the fuzziness from the fuzzy front-end of service innovations through customer interactions. *Industrial Marketing Management*, 35(4):468–480, 2006.

[6] Joaquín Aldás-Manzano, Carlos Lassala-Navarré, Carla Ruiz-Mafé, and Silvia Sanz-Blas. The role Of consumer innovativeness and perceived risk in online banking usage. *International Journal of Bank Marketing*, 27(1):53, 2009.

[7] B.J. Allen, Deepa Chandrasekaran, and Suman Basuroy. Design Crowdsourcing:

The Impact on New Product Performance of Sourcing Design Solutions from the "Crowd". *Journal of Marketing*, 82(2):106–123, 2018.

[8] Teresa M. Amabile. Creativity, Artificial Intelligence, and a World of Surprises. *Academy of Management Discoveries*, 6(3):351–354, 2019.

[9] Allen C. Amason. Distinguising the effects of functional and dysfunctional conflict on strategic decision making: Resolving a paradox for top management teams. *Academy of Management Journal*, 39(1):123–148, 1996.

[10] Kendall W. Artz, Patricia M. Norman, Donald E. Hatfield, and Laura B. Cardinal. A Longitudinal Study of the Impact of R&D, Patents, and Product Innovation on Firm Performance. *Journal of Product Innovation Management*, 27(5):725–740, 2010.

[11] Stuart J. Barnes and Andrew D. Pressey. Cyber-mavens and online flow experiences: Evidence from virtual worlds. *Technological Forecasting and Social Change*, 111:285–296, October 2016.

[12] F.M. Bass. A new product growth model for consumer durables. *Management Science*, 15(1):215–227, 1969.

[13] Russell Belk. You are what you can access: Sharing and collaborative consumption online. *Journal of Business Research*, 67(8):1595–1600, 2014.

[14] Frank-Martin Belz and Wenke Baumbach. Netnography as a Method of Lead User Identification. *Creativity and Innovation Management*, 19(3):304–313, 2010.

[15] S. Benson Edwin Raj and A. Annie Portia. Analysis on credit card fraud detection methods. In *2011 International Conference on Computer, Communication and Electrical Technology (ICCCET)*, pages 152–156, March 2011.

[16] Anol Bhattacharjee. Individual Trust in Online Firms: Scale Development and Initial Test. *Journal of Management Information Systems:*, 19(1):211–241, 2014.

[17] Volker Bilgram, Alexander Brem, and Kai-Ingo Voigt. User-centric innovations in new product development — systematic identification of lead users harnessing interactive and collaborative online-tools. *International Journal of Innovation Management*, 12(03):419–458, 2008.

[18] Ivo Blohm, Jan Marco Leimeister, and Helmut Krcmar. Crowdsourcing: How to Benefit from (Too) Many Great Ideas. *MIS Quarterly Executive*, 12(4):199–211, 2013.

[19] Marcel Bogers, Allan Afuah, and Bettina Bastian. Users as Innovators: A Review, Critique, and Future Research Directions. *Journal of Management*, 36(4):857–875, 2010.

[20] Stephen P. Borgatti, Ajay Mehra, Daniel J. Brass, and Giuseppe Labianca. Network Analysis in the Social Sciences. *Science*, 323(5916):892–895, 2009.

[21] Léon Bottou. Large-Scale Machine Learning with Stochastic Gradient Descent. In Yves Lechevallier and Gilbert Saporta, editors, *Proceedings of COMPSTAT'2010*, pages 177–186, Heidelberg, 2010. Physica-Verlag HD. ISBN 978-3-7908-2604-3.

[22] Kevin J. Boudreau, Nicola Lacetera, and Karim R. Lakhani. Incentives and Problem Uncertainty in Innovation Contests: An Empirical Analysis. *Management Science*, 57(5):843–863, 2011.

[23] Roderick J. Brodie, Ana Ilic, Biljana Juric, and Linda Hollebeek. Consumer engagement in a virtual brand community: An exploratory analysis. *Journal of Business Research*, 66(1):105–114, 2013.

[24] Brad Brown, Michael Chui, and James Manyika. Are you ready for the era of 'big data'. *McKinsey Quarterly*, 4(1):24–35, 2011.

[25] Shona L. Brown and Kathleen M. Eisenhardt. Product Development: Past Research, Present Findings, and Future Directions. *The Academy of Management Review*, 20(2):343, 1995.

[26] Shona L. Brown and Kathleen M. Eisenhardt. The Art of Continuous Change: Linking Complexity Theory and Time-Paced Evolution in Relentlessly Shifting Organizations. *Administrative Science Quarterly*, 42(1):1–34, 1997.

[27] Erik Brynjolfsson and Andrew McAfee. The Business of Artificial Intelligence. *Harvard Business Review*, (July):20, 2017.

[28] Alan J. Bush and E. Stephen Grant. Analyzing the Content of Marketing Journals to Assess Trends in Sales Force Research: 1980–1992. *Journal of Personal Selling & Sales Management*, 14(3):57–68, 1994.

[29] Gülçin Büyüközkan, Orhan Feyzioğlu, and Erdal Nebol. Selection of the strategic alliance partner in logistics value chain. *International Journal of Production Economics*, 113(1):148–158, 2008.

[30] Tomasa Calvo, Anna Kolesarova, Magda Komornikova, and Radko Mesiar. Aggregation Operators: Properties, Classes and Construction Methods. In *Aggregation Operators. Studies in Fuzziness and Soft Computing*, volume 97, pages 3–107. Physica, Heidelberg, 2002. ISBN 978-3-7908-1787-4.

[31] Alexandra J. Campbell and Robert G. Cooper. Do Customer Partnerships Improve New Product Success Rates? *Industrial Marketing Management*, 28(5):507–519, 1999.

[32] Rona Campbell, Pandora Pound, Catherine Pope, Nicky Britten, Roisin Pill, Mifanwy Morgan, and Jenny Donovan. Evaluating meta-ethnography: A synthesis of qualitative research on lay exp...: :. *Social Science & Medicine*, 56(4):671–685, 2003.

[33] Pinar Cankurtaran, Fred Langerak, and Abbie Griffin. Consequences of New Product Development Speed: A Meta-Analysis. *Journal of Product Innovation Management*, 30(3):465–486, 2013.

[34] Mònica Casabayó and Núria Agell. El gran potencial de la lógica difusa. *Harvard Deusto business review*, (234):50–57, 2014.

[35] Monica Casabayo, Nuria Agell, and Juan Carlos Aguado. Using AI Techniques in the Grocery Industry: Identifying the Customers Most Likely to Defect. *Int. Rev. of Retail, Distribution and Consumer Research*, 14(3):295–308, 2004.

[36] Monica Casabayo, Nuria Agell, and German Sanchez. Improved market segmentation by fuzzifying crisp clusters: A case study of the energy market in Spain. *Expert Systems With Applications*, 42(3):1637–1643, 2015.

[37] Junyi Chai, James N. K. Liu, and Eric W. T. Ngai. Application of decision-making techniques in supplier selection: A systematic review of literature. *Expert Systems with Applications*, 40(10):3872–3885, 2013.

[38] Anindita Chakravarty, Yong Liu, and Tridib Mazumdar. The Differential Effects of Online Word-of-Mouth and Critics' Reviews on Pre-release Movie Evaluation. *Journal of Interactive Marketing*, 24(3):185–197, 2010.

[39] Hing Kai Chan, Xuting Sun, and Sai-Ho Chung. When should fuzzy analytic hierarchy process be used instead of analytic hierarchy process? *Decision Support Systems*, 125:113114, 2019.

[40] Kenny K. Chan and Shekhar Misra. Characteristics of the Opinion Leader: A New Dimension. *Journal of Advertising*, 19(3):53–60, 1990.

[41] Kimmy Wa Chan, Chi Kin (Bennett) Yim, and Simon S.K. Lam. Is Customer Participation in Value Creation a Double-Edged Sword? Evidence from Professional Financial Services across Cultures. *Journal of Marketing*, 74(3):48–64, 2010.

[42] Da-Yong Chang. Applications of the extent analysis method on fuzzy AHP. *European Journal of Operational Research*, 95(3):649–655, 1996.

[43] Woojung Chang and Steven A. Taylor. The Effectiveness of Customer Participation in New Product Development: A Meta-Analysis. *Journal of Marketing*, 80(1):47–64, 2016.

[44] Gilles Chemla and Katrin Tinn. Learning Through Crowdfunding. *Management Science*, 66(5):1783–1801, 2019.

[45] Chen-Tung Chen. Extensions of the TOPSIS for group decision-making under fuzzy environment. *Fuzzy Sets and Systems*, 114(1):1–9, 2000.

[46] Chen-Tung Chen, Ching-Torng Lin, and Sue-Fn Huang. A fuzzy approach for supplier evaluation and selection in supply chain management. *International Journal of Production Economics*, 102(2):289–301, 2006.

[47] Yupeng Chen, Raghuram Iyengar, and Garud Iyengar. Modeling Multimodal Continuous Heterogeneity in Conjoint Analysis—A Sparse Learning Approach. *Marketing Science*, 36(1):140–156, 2017.

[48] Henry William Chesbrough. *Open Innovation: The New Imperative for Creating and Profiting from Technology*. Harvard Business Press, 2003. ISBN 978-1-57851-837-1.

[49] Vittorio Chiesa and Federico Frattini. Commercializing Technological Innovation: Learning from Failures in High-Tech Markets*. *Journal of Product Innovation Management*, 28(4):437–454, 2011.

[50] Kwai-Sang Chin, Dong-ling Xu, Jian-Bo Yang, and James Ping-Kit Lam. Group-based ER–AHP system for product project screening. *Expert Systems with Applications*, 35(4):1909–1929, 2008.

[51] R. G. Cooper and E. J. Kleinschmidt. New products: What separates winners from losers? *Journal of Product Innovation Management*, 4(3):169–184, 1987.

[52] R.G. Cooper and E.J. Kleinschmidt. Success factors in product innovation. *Industrial Marketing Management*, 16(3):215–223, 1987.

[53] Robert G. Cooper. New Products—What Separates the Winners from the Losers and What Drives Success. In Kenneth B. Kahn, editor, *The PDMA Handbook of New Product Development*, pages 3–34. John Wiley & Sons, New Jersey, USA, 3rd edition, 2013.

[54] Robert G. Cooper, Scott J. Edgett, and Elko J. Kleinschmidt. Benchmarking Best NPD Practices—I. *Research-Technology Management*, 47(1):31–43, January 2004.

[55] Anna S. Cui and Fang Wu. Utilizing customer knowledge in innovation: Antecedents and impact of customer involvement on new product performance. *Journal of the Academy of Marketing Science*, 44(4):516–538, 2016.

[56] R.M. Cyert and J.G. March. *A Behavioral Theory of the Firm*. Englewood Cliffs, N.J. : Prentice Hall, cop. 1980, 1963.

[57] Ely Dahan and V. Srinivasan. The Predictive Power of Internet-Based Product Concept Testing Using Visual Depiction and Animation. *Journal of Product Innovation Management*, 17(2):99–109, 2000.

[58] Erik Dane and Michael G. Pratt. Exploring Intuition and its Role in Managerial Decision Making. *Academy of Management Review*, 32(1):33–54, January 2007.

[59] TH Davenport and Julia Kirby. Just How Smart Are Smart Machines? *MIT Sloan Management Review*, 57(3), 2016.

[60] Veronika Denes-Raj and Seymour Epstein. Conflict between intuitive and rational processing: When people behave against their better judgment. *Journal of Personality and Social Psychology*, 66(5):819–829, 1994.

[61] Afef Denguir-Rekik, Jacky Montmain, and Gilles Mauris. A possibilistic-valued multi-criteria decision-making support for marketing activities in e-commerce: Feedback Based Diagnosis System. *European Journal of Operational Research*, 195 (3):876–888, 2009.

[62] James Derbyshire and Emanuele Giovannetti. Understanding the failure to understand New Product Development failures: Mitigating the uncertainty associated with innovating new products by combining scenario planning and forecasting. *Technological Forecasting and Social Change*, 125:334–344, 2017.

[63] C. Anthony Di Benedetto. Identifying the key success factors in new product launch. *Journal of Product Innovation Management*, 16(6):530–544, 1999.

[64] Thomas G. Dietterich. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. *Neural Computation*, 10(7):1895–1923, 1998.

[65] Anhai Doan, Raghu Ramakrishnan, and Alon Y. Halevy. Crowdsourcing systems on the World-Wide Web. *Communications of the ACM*, 54(4):86–96, 2011.

[66] Thomas Donaldson and Lee E. Preston. The Stakeholder Theory of the Corporation: Concepts, Evidence, and Implications. *Academy of Management Review*, 20(1): 65–91, 1995.

[67] Michael Doumpos and Constantin Zopounidis. *Multicriteria Decision Aid Classification Methods*. Kluwer Academic Publishers, 2002.

[68] Cornelia Droge, Michael A. Stanko, and Wesley A. Pollitte. Lead Users and Early Adopters on the Web: The Role of New Technology Product Blogs*. *Journal of Product Innovation Management*, 27(1):66–82, 2010.

[69] Didier Dubois and Henri Prade. *Fundamentals of Fuzzy Sets*. Springer Science & Business Media, December 2012. ISBN 978-1-4615-4429-6.

[70] Daria Dzyabura and John R. Hauser. Active Machine Learning for Consideration Heuristics. *Marketing Science*, 30(5):801–819, 2011.

[71] Martin Eisend. Have We Progressed Marketing Knowledge? A Meta-Meta-Analysis of Effect Sizes in Marketing Research. *Journal of Marketing*, 79(3):23–40, 2015.

[72] James F. Engel, David T. Kollat, and Roger D. Blackwell. *Consumer Behavior*. Hinsdale, Ill. : Dryden, cop. 1973, 1973. ISBN 978-0-03-086371-4.

[73] Tülin Erdem, Joffre Swait, and Jordan Louviere. The impact of brand credibility on consumer price sensitivity. *International Journal of Research in Marketing*, 19(1): 1–19, 2002.

[74] Holger Ernst, Wayne D. Hoyer, Manfred Krafft, and Katrin Krieger. Customer relationship management and company performance—the mediating role of new product performance. *Journal of the Academy of Marketing Science*, 39(2):290–306, 2011.

[75] Enrique Estellés-Arolas and Fernando González-Ladrón-de-Guevara. Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2):189–200, 2012.

[76] Heiner Evanschitzky, Martin Eisend, Roger J. Calantone, and Yuanyuan Jiang. Success Factors of Product Innovation: An Updated Meta-Analysis. *Journal of Product Innovation Management*, 29(S1):21–37, 2012.

[77] Peter S. Fader, Bruce G.S. Hardie, and Ka Lok Lee. RFM and CLV: Using Iso-Value Curves for Customer Base Analysis. *Journal of Marketing Research*, 42(4):415–430, 2005.

[78] Shaokun Fan, Raymond Y. K. Lau, and J. Leon Zhao. Demystifying Big Data Analytics for Business Intelligence Through the Lens of Marketing Mix. *Big Data Research*, 2(1):28–32, 2015.

[79] Jie Feng and Purushottam Papatla. Advertising: Stimulant or Suppressant of Online Word of Mouth? *Journal of Interactive Marketing*, 25(2):75–84, 2011.

[80] Alberto Fernandez, Salvador Garcia, Francisco Herrera, and Nitesh V. Chawla. SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-year Anniversary. *Journal of Artificial Intelligence Research*, 61:863–905, 2018.

[81] Karsten Frey, Christian Lüthje, and Simon Haag. Whom Should Firms Attract to Open Innovation Platforms? The Role of Knowledge Diversity and Motivation. *Long Range Planning*, 44(5):397–420, October 2011.

[82] Christoph Fuchs and Martin Schreier. Customer Empowerment in New Product Development*. *Journal of Product Innovation Management*, 28(1):17–32, 2011.

[83] Alfonso Gambardella, Christina Raasch, and Eric von Hippel. The User Innovation Paradigm: Impacts on Markets and Welfare. *Management Science*, 63(5): 1450–1468, 2016.

[84] Garcia and Calantone. A critical look at technological innovation typology and innovativeness terminology: A literature review. *Journal of Product Innovation Management*, 19(2):110–132, 2002.

[85] Salvador García, Julián Luengo, and Francisco Herrera. Tutorial on practical tips of the most influential data preprocessing algorithms in data mining. *Knowledge-Based Systems*, 98:1–29, 2016.

[86] Hubert Gatignon and Jean-Marc Xuereb. Strategic Orientation of the Firm and New Product Performance. *Journal of Marketing Research*, 34(1):77–90, 1997.

[87] Antonio Ghezzi, Donata Gabelloni, Antonella Martini, and Angelo Natalicchio. Crowdsourcing: A Review and Suggestions for Future Research. *International Journal of Management Reviews*, 20(2):343–363, 2018.

[88] Anil K. Ghosh. On optimum choice of k in nearest neighbor classification. *Computational Statistics & Data Analysis*, 50(11):3113–3123, 2006.

[89] Michael Gibbert, Marius Leibold, and Gilbert Probst. Five Styles of Customer Knowledge Management, and How Smart Companies Use Them To Create Value. *European Management Journal*, 20(5):459–469, 2002.

[90] Ronald Goldsmith and Leisa Reinecke Flynn. Identifying innovators in consumer Product markets. *European Journal of Marketing*, 26(12):42–55, 1992.

[91] Ronald E. Goldsmith, Jon B. Freiden, and Jacqueline K. Eastman. The generality/specificity issue in consumer innovativeness research. *Technovation*, 15(10): 601–612, 1995.

[92] Google. Google Trends for AI vs other technologies. https://trends.google.com/trends/explore?date=today%205-y&q=%2Fm%2F0mkz,Blockchain,Internet%20of%20things,%2Fm%2F01kl97, 2021.

[93] Paul E. Green and V. Srinivasan. Conjoint Analysis in Marketing: New Developments with Implications for Research and Practice. *Journal of Marketing*, 54(4):3–19, 1990.

[94] Dhruv Grewal, Michael Levy, and V. Kumar. Customer Experience Management in Retailing: An Organizing Framework. *Journal of Retailing*, 85(1):1–14, 2009.

[95] Dhruv Grewal, Anne Roggeveen, and Jens Nordfaelt. The future of retailing. *Journal of Retailing*, 93(1):1–6, 2017.

[96] Abbie Griffin. PDMA Research on New Product Development Practices: Updating Trends and Benchmarking Best Practices. *Journal of Product Innovation Management*, 14(6):429–458, 1997.

[97] Linda Hamdi-Kidar, Peter Keinz, Emmanuelle Le Nagard, and Eric Vernette. Comparing Lead Users to Emergent-Nature Consumers as Sources of Innovation at Early Stages of New Product Development. *Journal of Product Innovation Management*, 36(5):616–631, 2019.

[98] Colleen M. Harmeling, Jordan W. Moffett, Mark J. Arnold, and Brad D. Carlson. Toward a theory of customer engagement marketing. *Journal of the Academy of Marketing Science*, 45(3):312–335, 2017.

[99] Johann D. Harnoss and Ramon Baeza. Overcoming the Four Big Barriers to Innovation Success. Quarterly, Boston Consulting Group, USA, 2019.

[100] R. Helm, M. Steiner, A. Scholl, and L. Manthey. A comparative empirical study on common methods for measuring preferences. *International Journal of Management and Decision Making*, 9(3):242, 2008.

[101] David H. Henard and David M. Szymanski. Why Some New Products are More Successful than Others. *Journal of Marketing Research*, 38(3):362–375, 2001.

[102] Joachim Henkel and Eric von Hippel. Welfare Implications of User Innovation. *The Journal of Technology Transfer*, 30(1):73–87, 2004.

[103] Thorsten Hennig-Thurau, Kevin P. Gwinner, Gianfranco Walsh, and Dwayne D. Gremler. Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the Internet? *Journal of Interactive Marketing*, 18(1):38–52, 2004.

[104] Christoph Hienerth, Christopher Lettl, and Peter Keinz. Synergies among Producer Firms, Lead Users, and User Communities: The Case of the LEGO Producer–User Ecosystem. *Journal of Product Innovation Management*, 31(4):848–866, July 2014.

[105] Dorit S. Hochbaum and Asaf Levin. Methodologies and Algorithms for Group-Rankings Decision. *Management Science*, 52(9):1394–1408, 2006.

[106] Martin Hoegl and Hans Georg Gemuenden. Teamwork Quality and the Success of Innovative Projects: A Theoretical Concept and Empirical Evidence. *Organization Science*, 12(4):435–449, 2001.

[107] Donna L. Hoffman, Praveen K. Kopalle, and Thomas P. Novak. The "right" consumers for better concepts: Identifying consumers high in emergent nature to develop new product concepts. *Journal of Marketing Research*, 47(5):854–865, 2010.

[108] Robin M. Hogarth. *Educating Intuition*. University of Chicago Press, June 2001. ISBN 978-0-226-34860-5.

[109] Jeff Howe. The Rise of Crowdsourcing. *Wired Magazine*, (14):5, 2006.

[110] Wayne D. Hoyer, Rajesh Chandy, Matilda Dorotic, Manfred Krafft, and Siddharth S. Singh. Consumer Cocreation in New Product Development. *Journal of Service Research*, 13(3):283–296, 2010.

[111] Ting-Ya Hsieh, Shih-Tong Lu, and Gwo-Hshiung Tzeng. Fuzzy MCDM approach for planning and design tenders selection in public office buildings. *International Journal of Project Management*, 22(7):573–584, 2004.

[112] Dongling Huang and Lan Luo. Consumer Preference Elicitation of Complex Products Using Fuzzy Support Vector Machine Active Learning. *Marketing Science*, 35 (3):445–464, 2016.

[113] Yan Huang, Param Vir Singh, and Kannan Srinivasan. Crowdsourcing New Product Ideas Under Consumer Learning. *Management Science*, 60(9):2138–2159, 2014.

[114] Erik Jan Hultink, Abbie Griffin, Susan Hart, and Henry S. J. Robben. Industrial new product launch strategies and product development performance. *Journal of Product Innovation Management*, 14(4):243–257, 1997.

[115] Ching-Lai Hwang and Kwangsun Yoon. Methods for Multiple Attribute Decision Making. *Multiple Attribute Decision Making*, pages 58–191, 1981.

[116] Oriol Iglesias, Nicholas Ind, and Manuel Alfaro. The Organic View of the Brand: A Brand Value Co-creation Model. In John M. T. Balmer, Shaun M. Powell, Joachim Kernstock, and Tim Oliver Brexendorf, editors, *Advances in Corporate Branding*, Journal of Brand Management: Advanced Collections, pages 148–174. Palgrave Macmillan UK, London, 2017. ISBN 978-1-352-00008-5.

[117] Subin Im, Barry L. Bayus, and Charlotte H. Mason. An Empirical Study of Innate Consumer Innovativeness, Personal Characteristics, and New-Product Adoption Behavior. *Journal of the Academy of Management Science*, 31(No. 1):61–73, 2003.

[118] Sitwala Imenda. Is There a Conceptual Difference between Theoretical and Conceptual Frameworks? *Journal of Social Sciences*, 38(2):185–195, 2014.

[119] Ola Isaksson, Tobias C. Larsson, and Anna Öhrwall Rönnbäck. Development of product-service systems: Challenges and opportunities for the manufacturing firm. *Journal of Engineering Design*, 20(4):329–348, 2009.

[120] Alessio Ishizaka and Sajid Siraj. Are multi-criteria decision-making tools useful? An experimental comparative study of three methods. *European Journal of Operational Research*, 264(2):462–471, 2018.

[121] Stanislav Hristov Ivanov and Craig Webster. Adoption of Robots, Artificial Intelligence and Service Automation by Travel, Tourism and Hospitality Companies – A Cost-Benefit Analysis. SSRN Scholarly Paper ID 3007577, Social Science Research Network, Rochester, NY, 2017.

[122] Raghuram Iyengar, Christophe Van den Bulte, and Thomas W. Valente. Opinion Leadership and Social Contagion in New Product Diffusion. *Marketing Science*, 30 (2):195–212, 2011.

[123] Bernard J. Jaworski and Ajay K. Kohli. Market Orientation: Antecedents and Consequences. *Journal of Marketing*, 57(3):53–70, 1993.

[124] Lars Bo Jeppesen and Måns J. Molin. Consumers as Co-developers: Learning and Innovation Outside the Firm. *Technology Analysis & Strategic Management*, 15(3): 363–383, 2003.

[125] Jon-Arild Johannessen and Bjørn Olsen. The future of value creation and innovations: Aspects of a theory of value creation and innovation in a global knowledge economy. *International Journal of Information Management*, 30(6):502–511, 2010.

[126] M. I. Jordan and T. M. Mitchell. Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245):255–260, 2015.

[127] Daniel Kahneman. A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, 58(9):697–720, 2003.

[128] Vinayak Kalluri and Rambabu Kodali. Analysis of new product development research: 1998-2009 ׀. *Benchmarking: An International Journal*, 21(4):527–618, 2014.

[129] Riitta Katila and Gautam Ahuja. Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction. *Academy of Management Journal*, 45(6):1183–1194, 2002.

[130] Michael L Katz and Carl Shapiro. Network Externalities,Competition, and Compatibility. *The American Economic Review*, 75(3):424–440, 1985.

[131] Harsurinder Kaur, Husanbir Singh Pannu, and Avleen Kaur Malhi. A Systematic Review on Imbalanced Data Challenges in Machine Learning: Applications and Solutions. *ACM Computing Surveys*, 52(4):79:1–79:36, 2019.

[132] Gitae Kim, Bongsug Kevin Chae, and David L. Olson. A support vector machine (SVM) approach to imbalanced datasets of customer responses: Comparison with other customer response models. *Service Business*, 7(1):167–182, 2013.

[133] Namwoon Kim and Kwaku Atuahene-Gima. Using Exploratory and Exploitative Market Learning for New Product Development*. *Journal of Product Innovation Management*, 27(4):519–536, 2010.

[134] YongSeog Kim and W. Nick Street. An intelligent system for customer targeting: A data mining approach. *Decision Support Systems*, 37(2):215–228, 2004.

[135] Ajay K. Kohli and Bernard J. Jaworski. Market Orientation: The Construct, Research Propositions, and Managerial Implications. *Journal of Marketing*, 54(2):1–18, 1990.

[136] Vegard Kolbjørnsrud, Richard Amico, and Robert J Thomas. How Artificial Intelligence Will Redefine Management. *Harvard Business Review*, (November):2–6, 2016.

[137] Maryam Kordi and S. Anders Brandt. Effects of increasing fuzziness on analytic hierarchy process for spatial multicriteria decision analysis. *Computers, Environment and Urban Systems*, 36(1):43–53, 2012.

[138] Philip Kotler. *Marketing Management - Millenium Edition*. Prentice-Hall Inc., New Jersey, USA, tenth edition, 2000. ISBN 0-536-63099-2.

[139] V. Krishnan and Karl T. Ulrich. Product Development Decisions: A Review of the Literature. *Management Science*, 47(1):1–21, 2001.

[140] Per Kristensson, Anders Gustafsson, and Trevor Archer. Harnessing the Creative Potential among Users*. *Journal of Product Innovation Management*, 21(1):4–14, 2004.

[141] Per Kristensson, Jonas Matthing, and Niklas Johansson. Key strategies for the successful involvement of customers in the co-creation of new technology-based services. *International Journal of Service Industry Management*, 19(4):474–491, 2008.

[142] Vincent S Lai, Bo K Wong, and Waiman Cheung. Group decision making in a multiple criteria environment: A case using the AHP in software selection. *European Journal of Operational Research*, 137(1):134–144, 2002.

[143] Anja Lambrecht, Catherine Tucker, and Caroline Wiertz. Advertising to Early Trend Propagators: Evidence from Twitter. *Marketing Science*, 2018.

[144] Fred Langerak, Erik Jan Hultink, and Henry S. J. Robben. The Impact of Market Orientation, Product Advantage, and Launch Proficiency on New Product Performance and Organizational Performance. *Journal of Product Innovation Management*, 21(2):79–94, 2004.

[145] Kun Chang Lee, Habin Lee, Namho Lee, and Jaehoon Lim. An agent-based fuzzy cognitive map approach to the strategic marketing planning for industrial firms. *Industrial Marketing Management*, 42(4):552–563, 2013.

[146] Matthew K.O. Lee, Christy M.K. Cheung, Kai H. Lim, and Choon Ling Sia. Understanding customer knowledge sharing in web-based discussion boards: An exploratory study. *Internet Research*, 16(3):289–303, May 2006.

[147] Yikuan Lee and Gina Colarelli O'Connor. The Impact of Communication Strategy on Launching New Products: The Moderating Role of Product Innovativeness. *Journal of Product Innovation Management*, 20(1):4–21, 2003.

[148] Guillaume Lema{{\î}}tre, Fernando Nogueira, and Christos K. Aridas. Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. *Journal of Machine Learning Research*, 18(17):1–5, 2017.

[149] Katherine N. Lemon and Peter C. Verhoef. Understanding Customer Experience Throughout the Customer Journey. *Journal of Marketing*, 80(November):69–96, 2016.

[150] Dorothy Leonard-Barton. Core capabilities and core rigidities: A paradox in managing new product development. *Strategic Management Journal*, 13(S1):111–125, 1992.

[151] Christopher Lettl, Cornelius Herstatt, and Hans Georg Gemuenden. Users' contributions to radical innovation: Evidence from four cases in the field of medical equipment technology. *R&D Management*, 36(3):251–272, 2006.

[152] Shuliang Li, Barry Davies, John Edwards, Russell Kinman, and Yanqing Duan. Integrating group Delphi, fuzzy logic and expert systems for marketing strategy development: The hybridisation and its effectiveness. *Marketing Intelligence & Planning*, 20(5):273–284, 2002.

[153] Chechen Liao, Jain-Liang Chen, and David C. Yen. Theory of planning behavior (TPB) and customer satisfaction in the continued use of e-service: An integrated model. *Computers in Human Behavior*, 23(6):2804–2822, 2007.

[154] Nan Lin. Social Network and Status attainment. *Annual Review of Sociology*, 25: 467–487, 1999.

[155] Shixi Liu, Cuiqing Jiang, Zhangxi Lin, Yong Ding, Rui Duan, and Zhicai Xu. Identifying effective influencers based on trust for electronic word-of-mouth marketing: A domain-aware approach. *Information Sciences*, 306:34–52, 2015.

[156] June Lu. Are personal innovativeness and social influence critical to continue with mobile commerce? *Internet Research*, 24(2):134–159, 2014.

[157] June Lu, James E. Yao, and Chun-Sheng Yu. Personal innovativeness, social influences and adoption of wireless Internet services via mobile technology. *The Journal of Strategic Information Systems*, 14(3):245–268, 2005.

[158] Yifan Lu and Margaret Rucker. Apparel acquisition via single vs. multiple channels: College students' perspectives in the US and China. *Journal of Retailing and Consumer Services*, 13(1):35–50, 2006.

[159] C. Lüthje. Characteristics of innovating users in a consumer goods field: An empirical study of sport-related product consumers. *Technovation*, 24(9):683–695, 2004.

[160] Christian Lüthje and Cornelius Herstatt. The Lead User method: An outline of empirical findings and issues for future research. *R&D Management*, 34(5):553–568, 2004.

[161] Vijay Mahajan and Eitan Muller. When is it worthwhile targeting the majority instead of the Innovators in a new Product launch? *Journal of Marketing Research*, 35:488–495, 1998.

[162] Vijay Mahajan and Jerry Wind. New product models: Practice, shortcomings and desired improvements. *Journal of Product Innovation Management*, 9(2):128–139, 1992.

[163] Dominik Mahr, Annouk Lievens, and Vera Blazevic. The Value of Customer Cocreated Knowledge during the Innovation Process. *Journal of Product Innovation Management*, 31(3):599–615, 2014.

[164] Albert E. Mannes, Jack B. Soll, and Richard P. Larrick. The wisdom of select crowds. *Journal of Personality and Social Psychology*, 107(2):276–299, 2014.

[165] Jonas Matthing, Per Kristensson, Anders Gustafsson, and A. Parasuraman. Developing successful technology-based services: The issue of identifying and involving innovative users. *Journal of Services Marketing*, 20(5):288–297, 2006.

[166] Philipp Mayring. Qualitative Inhaltsanalyse. In *Mey G., Mruck K. (eds) Handbuch Qualitative Forschung in der Psychologi*, pages 601–613. VS Verlag für Sozialwissenschaften, Wiesbaden, 2010.

[167] D. Harrison McKnight, Vivek Choudhury, and Charles Kacmar. Developing and Validating Trust Measures for e-Commerce: An Integrative Typology. *Information Systems Research*, 13(3):334–359, 2002.

[168] Regina C. McNally, M. Billur Akdeniz, and Roger J. Calantone. New Product Development Processes and New Product Profitability: Exploring the Mediating Role of Speed to Market and Product Quality. *Journal of Product Innovation Management*, pages 63–77, 2019.

[169] Quinn McNemar. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2):153–157, June 1947. ISSN 1860-0980. doi: 10.1007/BF02295996.

[170] Sheila Mello. *Customer-Centric Product Definition: The Key to Great Product*. PDC Professional Publishing, Boston, M.A., 2002. ISBN 978-0-9745604-0-3.

[171] Jeremy J. Michalek, Fred M. Feinberg, and Panos Y. Papalambros. Linking Marketing and Engineering Product Design Decisions via Analytical Target Cascading*. *Journal of Product Innovation Management*, 22(1):42–62, 2005.

[172] David F. Midgley and Grahame R. Dowling. Innovativeness: The Concept and Its Measurement. *Journal of Consumer Research*, 4:229–241, 1978.

[173] Mintel Group Ltd. Mintel GNPD - Global New Products Database: Features. https://www.mintel.com/global-new-products-database/features, 2019.

[174] Don-Lin Mon, Ching-Hsue Cheng, and Jiann-Chern Lin. Evaluating weapon system using fuzzy analytic hierarchy process based on entropy weight. *Fuzzy Sets and Systems*, 62(2):127–134, 1994.

[175] Mitzl M. Montoya-Weiss and Roger Calantone. Determinants of new product performance: A review and meta-analysis. *Journal of Product Innovation Management*, 11(5):397–417, 1994.

[176] Hakil Moon, Douglas R. Miller, and Sung Hyun Kim. Product Design Innovation and Customer Value: Cross-Cultural Research in the United States and Korea. *Journal of Product Innovation Management*, 30(1):31–43, 2013.

[177] Hakil Moon, Jeongdoo Park, and Sangkyun Kim. The Importance of an Innovative Product Design on Customer Behavior: Development and Validation of a Scale. *Journal of Product Innovation Management*, 32(2):224–232, 2015.

[178] Pamela D Morrison, John H Roberts, and David F Midgley. The nature of lead users and measurement of leading edge status. *Research Policy*, 33(2):351–362, March 2004.

[179] Razieh Mosadeghi, Jan Warnken, Rodger Tomlinson, and Hamid Mirfenderesk. Comparison of Fuzzy-AHP and AHP in a spatial multi-criteria decision making model for urban land-use planning. *Computers, Environment and Urban Systems*, 49:54–65, 2015.

[180] C. D. Mulrow. Systematic Reviews: Rationale for systematic reviews. *BMJ*, 309 (6954):597–599, 1994.

[181] Cheryl Nakata and Yili Huang. Progress and promise: The last decade of international marketing research. *Journal of Business Research*, 58(5):611–618, 2005.

[182] Satish Nambisan. Designing Virtual Customer Environments for New Product Development: Toward a Theory. *Academy of Management Review*, 27(3):392–413, 2002.

[183] Matthew O'Hern and Aric Rindfleisch. Customer Co-creation: A Typology and Research Agenda. In *Review of Marketing Research*, volume 6, pages 84–106. Armonk NY, New York, U.S.A., 2009.

[184] Daniel O'Leary. Artificial Intelligence and Big Data. *IEEE Intelligent Systems*, 28 (March-April):96–99, 2013.

[185] Muammer Ozer. The roles of product lead-users and product experts in new product evaluation. *Research Policy*, 38(8):1340–1349, 2009.

[186] Margherita Pagani. Determinants of adoption of third generation mobile multimedia services. *Journal of Interactive Marketing*, 18(3):46–59, 2004.

[187] Albert L. Page and Gary R. Schirr. Growth and Development of a Body of Knowledge: 16 Years of New Product Development Research, 1989–2004. *Journal of Product Innovation Management*, 25(3):233–248, 2008.

[188] Hyun-Hee Park and Mi-Jin Noh. The Influence of Innovativeness and Price Sensitivity on Purchase Intention of Smart Wear. *Journal of the Korean Society of Clothing and Textiles*, 36(2):218–230, 2012.

[189] Fabian Pedregosa, Gael Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss,

Vincent Dubourg, Jake Vanderplas, Alexandre Passos, and David Cournapeau. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12: 2825–2830, 2011.

[190] Eyal Peer, Laura Brandimarte, Sonam Samat, and Alessandro Acquisti. Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70:153–163, 2017.

[191] Kenneth J. Petersen, Robert B. Handfield, and Gary L. Ragatz. Supplier integration into new product development: Coordinating product, process and supply chain design. *Journal of Operations Management*, 23(3):371–388, 2005.

[192] Leif E. Peterson. K-nearest neighbor. *Scholarpedia*, 4(2):1883, 2009.

[193] Frank T. Piller, Christoph Ihl, and Alexander Vossen. A Typology of Customer Co-Creation in the Innovation Process. SSRN Scholarly Paper ID 1732127, Social Science Research Network, Rochester, NY, December 2010.

[194] Marion K. Poetz and Martin Schreier. The Value of Crowdsourcing: Can Users Really Compete with Professionals in Generating New Product Ideas? *Journal of Product Innovation Management*, 29(2):245–256, 2012.

[195] C. K. Prahalad and Venkat Ramaswamy. Co-creation experiences: The next practice in value creation. *Journal of Interactive Marketing*, 18(3):5–14, 2004.

[196] C K Prahalad and Venkatram Ramaswamy. Co-opting Customer Competence. *Harvard Business Review*, 78(1):79–88, 2000.

[197] Sebastian Raisch and Sebastian Krakowski. Artificial Intelligence and Management: The Automation–Augmentation Paradox. *Academy of Management Review*, 46(1):192–210, 2021.

[198] Kumar Ranjan and Stuart Read. Value co-creation: Concept and measurement. *Journal of the Academy of Marketing Science*, 44(3):290–315, 2016.

[199] Sam Ransbotham, Philipp Gerbert, Martin Reeves, David Kiron, and Michael Spira. Artificial Intelligence in Business Gets Real. MIT Sloan Management review, MIT, Boston, M.A., 2018.

[200] Virginia P. Richmond. The Relationship Between Opinion Leadership and Information Acquisition. *Human Communication Research*, 4(1):38–43, 1977.

[201] George Ritzer and Nathan Jurgenson. Production, Consumption, Prosumption: The nature of capitalism in the age of the digital 'prosumer'. *Journal of Consumer Culture*, 10(1):13–36, 2010.

[202] Thomas S. Robertson and Yoram Wind. Organizational Psychographics and Innovativeness. *Journal of Consumer Research*, 7(1):24–31, 1980.

[203] Everett M. Rogers. *Diffusion of Innovations*. The Free Press - A division of Simon & Schuster Inc., U.S.A., first edition, 1962.

[204] Everett M. Rogers. Diffusion of preventive innovations. *Addictive Behaviors*, 27: 989–993, 2002.

[205] Everett M. Rogers. *Diffusion of Innovation*. The Free Press, New York, 5th edition edition, 2003.

[206] E. Roghanian, J. Rahimi, and A. Ansari. Comparison of first aggregation and last aggregation in fuzzy group TOPSIS. *Applied Mathematical Modelling*, 34(12):3754–3766, 2010.

[207] Ribana Roscher, Bastian Bohn, Marco F. Duarte, and Jochen Garcke. Explainable Machine Learning for Scientific Insights and Discoveries. *IEEE Access*, 8:42200–42216, 2020.

[208] Frank T. Rothaermel and Andrew M. Hess. Building Dynamic Capabilities: Innovation Driven by Individual-, Firm-, and Network-Level Effects. *Organization Science*, 18(6):898–921, December 2007.

[209] Thomas L. Saaty. Analytic Hierarchy Process. In *Encyclopedia of Biostatistics*. American Cancer Society, 2005. ISBN 978-0-470-01181-2.

[210] Thomas L. Saaty. Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, 1(1):83–98, 2008.

[211] Soren Salomo, Fee Steinhoff, and Volker Trommsdorff. Customer orientation in innovation projects and new product development success - the moderating effect of product innovativeness. *International Journal of Technology Management*, 26(5-6): 442–463, 2003.

[212] Vella Verónica Somoza Sánchez, Davide Giacalone, and René Chester Goduscheit. Digital anthropology as method for lead user identification from unstructured big data. *Creativity and Innovation Management*, 27(1):32–41, 2018.

[213] Elizabeth B.-N. Sanders and Pieter Jan Stappers. Co-creation and the new landscapes of design. *CoDesign*, 4(1):5–18, 2008.

[214] Alexander Sänn and Daniel Baier. Lead User Identification in Conjoint Analysis Based Product Design. In Wolfgang A. Gaul, Andreas Geyer-Schulz, Lars Schmidt-Thieme, and Jonas Kunze, editors, *Challenges at the Interface of Data Analysis, Computer Science, and Optimization*, Studies in Classification, Data Analysis, and Knowledge Organization, pages 521–528, Berlin, Heidelberg, 2012. Springer. ISBN 978-3-642-24466-7.

[215] Mohanbir Sawhney, Gianmario Verona, and Emanuela Prandelli. Collaborating to create: The Internet as a platform for customer engagement in product innovation. *Journal of Interactive Marketing*, 19(4):4–17, 2005.

[216] Chelsea Schelly. Residential solar electricity adoption: What motivates, and what matters? A case study of early adopters. *Energy Research & Social Science*, 2:183–191, 2014.

[217] Brita Schemmann, Andrea M. Herrmann, Maryse M. H. Chappin, and Gaston J. Heimeriks. Crowdsourcing ideas: Involving ordinary users in the ideation phase of new product development. *Research Policy*, 45(6):1145–1154, July 2016.

[218] Kenneth C. Schneider and William C. Rodgers. Generalized marketplace influencers' (market mavens') attitudes toward direct mail as a source of... *Journal of Direct Marketing*, 7(4):20–28, September 1993. ISSN 08920591.

[219] Martin Schreier and Reinhard Prügl. Extending Lead-User Theory: Antecedents and Consequences of Consumers' Lead Userness*. *Journal of Product Innovation Management*, 25(4):331–346, 2008.

[220] Joseph A. Schumpeter. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*. Oxford University Press, London, 1934.

[221] Tim G. Schweisfurth. Comparing internal and external lead users as sources of innovation. *Research Policy*, 46(1):238–248, 2017.

[222] Stefan Seuring and Stefan Gold. Conducting content-analysis based literature reviews in supply chain management. *Supply Chain Management: An International Journal*, 17(5):544–555, 2012.

[223] Stefan Seuring and Martin Müller. From a literature review to a conceptual framework for sustainable supply chain management. *Journal of Cleaner Production*, 16 (15):1699–1710, 2008.

[224] Sonali K. Shah. Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development. *Management Science*, 52(7):1000–1014, 2006.

[225] Sonali K. Shah and Mary Tripsas. The accidental entrepreneur: The emergent and collective process of user entrepreneurship. *Strategic Entrepreneurship Journal*, 1 (1-2):123–140, 2007.

[226] Jie Sheng, Joseph Amankwah-Amoah, and Xiaojun Wang. A multidisciplinary perspective of big data in management research. *International Journal of Production Economics*, 191:97–112, 2017.

[227] Jagdish N. Sheth, Rajendra S. Sisodia, and Arun Sharma. The antecedents and consequences of customer-centric marketing. *Journal of the Academy of Marketing Science*, 28(1):55–66, 2000.

[228] Duncan Simester. Why great new products fail. *MIT Sloan Management Review*, 57 (3):32–40, 2016.

[229] Duncan Simester, Artem Timoshenko, and Spyros I. Zoumpoulis. Targeting Prospective Customers: Robustness of Machine-Learning Methods to Typical Data Challenges. *Management Science*, 2019.

[230] Nagaraj Sivasubramaniam, S. Jay Liebowitz, and Conway L. Lackman. Determinants of New Product Development Team Performance: A Meta-analytic Review. *Journal of Product Innovation Management*, 29(5):803–820, 2012.

[231] Alina B. Sorescu and Jelena Spanjol. Innovation's Effect on Firm Value and Risk: Insights from Consumer Packaged Goods. *Journal of Marketing*, 72(2):114–132, 2008.

[232] Keith E. Stanovich and Richard F. West. Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences*, 23(5):645–665, 2000.

[233] Jan-Benedict E. M. Steenkamp and Katrijn Gielens. Consumer and Market Drivers of the Trial Probability of New Consumer Packaged Goods. *Journal of Consumer Research*, 30(3):368–384, 2003.

[234] Stephen L. Vargo and Robert F. Lusch. Evolving to a New Dominant Logic for Marketing. *Journal of Marketing*, 68(1):1–17, 2004.

[235] Chao-Ton Su, Yung-Hsin Chen, and D. Y. Sha. Linking innovative product development with customer knowledge: A data-mining approach. *Technovation*, 26(7): 784–795, 2006.

[236] Chia-Chi Sun. A performance evaluation model by integrating fuzzy AHP and fuzzy TOPSIS methods. *Expert Systems with Applications*, 37(12):7745–7754, 2010.

[237] Robert I. Sutton and Barry M. Staw. What Theory is Not. *Administrative Science Quarterly*, 40(3):371–384, 1995.

[238] Morgan Swink, Srinivas Talluri, and Temyos Pandejpong. Faster, better, cheaper: A study of NPD project efficiency and performance tradeoffs. *Journal of Operations Management*, 24(5):542–562, 2006.

[239] Kim Hua Tan and Yuanzhu Zhan. Improving new product development using big data: A case study of an electronics company. *R&D Management*, 47(4):570–582, 2017.

[240] Christian Terwiesch and Yi Xu. Innovation Contests, Open Innovation, and Multiagent Problem Solving. *Management Science*, 54(9):1529–1543, 2008.

[241] Rakhi Thakur and Mala Srivastava. Adoption readiness, personal innovativeness, perceived risk and usage intention across customer groups for mobile payment services in India. *Internet Research*, 24(3):369–392, 2014.

[242] Stefan Thomke and Eric von Hippel. Customers as Innovators: A new way to creat value. *Harvard Business Review*, (April):5–11, 2002.

[243] Artem Timoshenko and John R. Hauser. Identifying Customer Needs from User-Generated Content. *Marketing Science*, 38(1):1–20, 2019.

[244] Liem Tran and Lucien Duckstein. Comparison of fuzzy numbers using a fuzzy distance measure. *Fuzzy Sets and Systems*, 130(3):331–341, 2002.

[245] David Tranfield, David Denyer, and Palminder Smart. Towards a Methodology for Developing Evidence-Informed Management Knowledge by Means of Systematic Review. *British Journal of Management*, 14(3):207–222, 2003.

[246] Jakob Trischler, Simon J. Pervan, Stephen J. Kelly, and Don R. Scott. The Value of Codesign: The Effect of Customer Involvement in Service Design Teams. *Journal of Service Research*, 21(1):75–100, 2018.

[247] Paul Trott. *Innovation Management and New Product Development*. Pearson Education, England, third edition, 2005. ISBN 978-0-273-68643-9.

[248] Lisa C. Troy, Tanawat Hirunyawipada, and Audhesh K. Paswan. Cross-Functional Integration and New Product Success: An Empirical Investigation of the Findings. *Journal of Marketing*, 72(6):132–146, 2008.

[249] V. V. Tsyganok, S. V. Kadenko, and O. V. Andriichuk. Significance of expert competence consideration in group decision making using AHP. *International Journal of Production Research*, 50(17):4785–4792, 2012.

[250] Amos Tversky and Daniel Kahneman. Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4):293–315, 1983.

[251] Glenn L. Urban, Fareena Sultan, and William J. Qualis. Placing Trust at the Center of Your Internet Strategy. *Sloan Management review*, 42(1):39–48, 2000.

[252] Christophe Van den Bulte and Yogesh V. Joshi. New Product Diffusion with Influentials and Imitators. *Marketing Science*, 26(3):400–421, 2007.

[253] Christophe Van den Bulte and Gary L. Lilien. Medical Innovation Revisited: Social Contagion versus Marketing Effort. *American Journal of Sociology*, 106(5):1409–1435, 2001.

[254] Stephen L. Vargo and Robert F. Lusch. Evolving to a New Dominant Logic for Marketing. *Journal of Marketing*, 68(1):1–17, 2004.

[255] Mark Velasquez and Patrick T. Hester. An Analysis of Multi-Criteria Decision Making Methods. *International Journal of Operations research*, 10(2):56–66, 2013.

[256] Eric Vernette and Linda Hamdi Kidar. Co-creation with Consumers: Who has the Competence and Wants to Cooperate? *International Journal of Market Research*, 55 (4):539–561, 2013.

[257] Eric von Hippel. Lead users: A source of novel product concepts. *Management Science*, 32(7):791–805, 1986.

[258] Eric von Hippel. "Sticky Information" and the Locus of Problem Solving: Implications for Innovation. *Management Science*, 40(4):429–439, 1994.

[259] Eric von Hippel. *Democratizing Innovation*. MIT Press, Cambridge, Mass, 2005. ISBN 978-0-262-00274-5.

[260] Eric von Hippel and Ralph Katz. Shifting Innovation to Users via Toolkits. *Management Science*, 48(7):821–833, 2002.

[261] Eric von Hippel, Nikolaus Franke, and Reinhard Prügl. "Pyramiding: Efficient search for rare subjects". *Research Policy*, 38(9):1397–1406, 2009.

[262] Maximilian von Zedtwitz and Oliver Gassmann. Managing Customer Oriented Research. *International Journal of Technology Management*, 24:165–193, 2002.

[263] Michael Wade and John Hulland. Review: The Resource-Based View and Information Systems Research: Review, Extension, and Suggestions for Future Research. *MIS Quarterly*, 28(1):107–142, 2004.

[264] Christian Wagner and Ann Majchrzak. Enabling Customer-Centricity Using Wikis and the Wiki Way. *Journal of Management Information Systems*, 23(3):17–43, 2006.

[265] Samuel Fosso Wamba, Angappa Gunasekaran, Shahriar Akter, Steven Ji-fan Ren, Rameshwar Dubey, and Stephen J. Childe. Big data analytics and firm performance: Effects of dynamic capabilities. *Journal of Business Research*, 70:356–365, 2017.

[266] Tien-Chin Wang and Hsien-Da Lee. Developing a fuzzy TOPSIS approach based on subjective weights and objective weights. *Expert Systems with Applications*, 36 (5):8980–8985, 2009.

[267] Duncan J. Watts and Peter Sheridan Dodds. Influentials, Networks, and Public Opinion Formation. *Journal of Consumer Research*, 34(4):441–458, 2007.

[268] Michel Wedel and P.k. Kannan. Marketing Analytics for Data-Rich Environments. *Journal of Marketing*, 80(6):97–121, 2016.

[269] David A Whetten. What Constitutes a Theoretical Contribution? *Academy of Management Review*, 14(4):490–495, 1989.

[270] Richard Whittington. Completing the Practice Turn in Strategy Research. *Organization Studies*, 27(5):613–634, 2006.

[271] Brian C. Williams and Christopher R. Plouffe. Assessing the evolution of sales knowledge: A 20-year content analysis. *Industrial Marketing Management*, 36:408–419, 2007.

[272] H James Wilson, Paul R Daugherty, and Nicola Morini-Bianzino. The Jobs That Artificial Intelligence Will Create. *MIT Sloan Management Review*, 58(4):13–16, 2017.

[273] Yoram Wind and Thomas L. Saaty. Marketing Applications of the Analytic Hierarchy Process. *Management Science*, 26(7):641–658, 1980.

[274] Lars Witell, Per Kristensson, Anders Gustafsson, and Martin Löfgren. Idea generation: Customer co-creation versus traditional market research techniques. *Journal of Service Management*, 22(2):140–159, 2011.

[275] WoS. Web of Science: Journal Title abbreviations. Technical report, Web of Science, 2018.

[276] Cheng-Shiung Wu, Chin-Tsai Lin, and Chuan Lee. Optimal marketing strategy: A decision-making with ANP and TOPSIS. *International Journal of Production Economics*, 127(1):190–196, 2010.

[277] Ye-Jun Xu and Qing-Li Da. Standard and mean deviation methods for linguistic group decision making and their applications. *Expert Systems with Applications*, 37 (8):5905–5912, 2010.

[278] Nese Yalcin, Ali Bayrakdaroglu, and Cengiz Kahraman. Application of fuzzy multi-criteria decision making methods for financial performance evaluation of Turkish manufacturing industries. *Expert Systems with Applications*, 39(1):350–364, 2012.

[279] Xin Yao. Evolving artificial neural networks. *Proceedings of the IEEE*, 87(9):1423–1447, 1999.

[280] Hema Yoganarasimhan. Search Personalization Using Machine Learning. *Management Science*, 66(3):1045–1070, 2020.

[281] Lofti A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.

[282] Debra Zahay, Nick Hajli, and Debika Sihi. Managerial perspectives on crowdsourcing in the new product development process. *Industrial Marketing Management*, 71:41–53, 2018.

[283] Judith L. Zaichkowsky. Measuring the Involvement Construct. *Journal of Consumer Research*, 12(1):341–352, 1985.

[284] Edmundas Kazimieras Zavadskas, Abbas Mardani, Zenonas Turskis, Ahmad Jusoh, and Khalil MD Nor. Development of TOPSIS Method to Solve Complicated Decision-Making Problems — An Overview on Developments from 2000 to 2015. *International Journal of Information Technology & Decision Making*, 15(03):645–682, 2016.

[285] Yuanzhu Zhan, Kim Hua Tan, Yina Li, and Ying Kei Tse. Unlocking the power of big data in new product development. *Annals of Operations Research*, 270(1): 577–595, 2018.

[286] Hengjie Zhang, Yucheng Dong, and Xin Chen. The 2-Rank Consensus Reaching Model in the Multigranular Linguistic Multiple-Attribute Group Decision-Making. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 48(12):2080–2094, 2018.

[287] Lina Zhou, Shimei Pan, Jianwu Wang, and Athanasios V. Vasilakos. Machine learning on big data: Opportunities and challenges. *Neurocomputing*, 237:350–361, 2017.