The following publication Tsang, Y. P., Wu, C. H., Lin, K. Y., Tse, Y. K., Ho, G. T. S., & Lee, C. K. M. (2022). Unlocking the power of big data analytics in new product development: An intelligent product design framework in the furniture industry. Journal of Manufacturing Systems, 62, 777-791 is available at https://doi.org/10.1016/j.jmsy.2021.02.003.

Unlocking the power of big data analytics in new product development: an intelligent product design framework in the furniture industry

Abstract: New product development to enhance companies' competitiveness and reputation is one of the leading activities in manufacturing. At present, achieving successful product design has become more difficult, even for companies with extensive capabilities in the market, because of disorganisation in the fuzzy front end (FFE) of the innovation process. Tremendous amounts of information, such as data on customers, manufacturing capability, and market trend, are considered in the FFE phase to avoid common flaws in product design. Because of the high degree of uncertainties in the FFE, multidimensional and high-volume data are added from time to time at the beginning of the formal product development process. To address the above concerns, deploying big data analytics to establish industrial intelligence is an active but still under-researched area. In this paper, an intelligent product design framework is proposed to incorporate fuzzy association rule mining (FARM) and a genetic algorithm (GA) into a recursive association-rule-based fuzzy inference system to bridge the gap between customer attributes and design parameters. Considering the current incidence of epidemics, such as the COVID-19 pandemic, communication of information in the FFE stage may be hindered. Through this study, a recursive learning scheme is established, therefore, to strengthen market performance, design performance, and sustainability on product design. It is found that the industrial big data analytics in the FFE process achieve greater flexibility and self-improvement mechanism on the evolution of product design.

Keywords— **Product design; Fuzzy front end; Fuzzy inference system; Big data analytics; Industrial intelligence**

1. Introduction

The current business environment, in which many companies are struggling to strengthen their competitiveness, is filled with numerous uncertainties and changes. Ways to approach the design of cyber-physical systems in Industry 4.0 are under development [1]. Shorter development lead times and the uncertainty of increasingly complex markets require intelligent product design in manufacturing systems [2]. Under these turbulent circumstances, merely offering fundamental functions without considering customer needs in the product design is no longer appropriate to facilitate trade and order fulfilment along the entire supply chain. Apart from production cost and capability, market and customer perspectives are major considerations for companies to gain extraordinary competitive advantages among peer competitors [3]. However, the focal inertia in production process design and modularization makes manufacturing sectors performing disappointingly in customization and customer satisfaction [4]. Also, manufacturing of ready-made products brings boundaries and expectation discrepancies, and thus the ability to capture customer needs is lacking. Therefore, research for new product development (NPD) has garnered considerable attention to introduce new products to the demanding market successfully.

Schneider and Hall [5] stated that the most critical problem in NPD is that companies are too focused on product design and manufacturing, but lack consideration and preparation in the market. For example, the Facebook Phone launched in 2013 was a failure of NPD because of the failure to investigate market and customer perspectives, and thus it was classified as merely an Android-skinned device manufactured by HTC. Moreover, Google Glass was launched in 2013 as a revolutionary and innovative product in the market but encountered many challenging issues, such as privacy, lack of consensus, and questionable value to customers [6]. In view of such examples of the failure of NPD, business organisations are striving to design products that best suit customer needs at the stage of the fuzzy front end (FFE) of product innovation, as shown in Figure 1. The FFE is regarded as the initial phase in the entire NPD process, which requires substantial investigation to increase the likelihood of success of the product innovation [7]. However, numerous sources of data and information, such as business cases, market research, and customer preference, are involved in the FFE, while the study of big data in the FFE is lacking to customise the product design through examining data patterns and relationship.

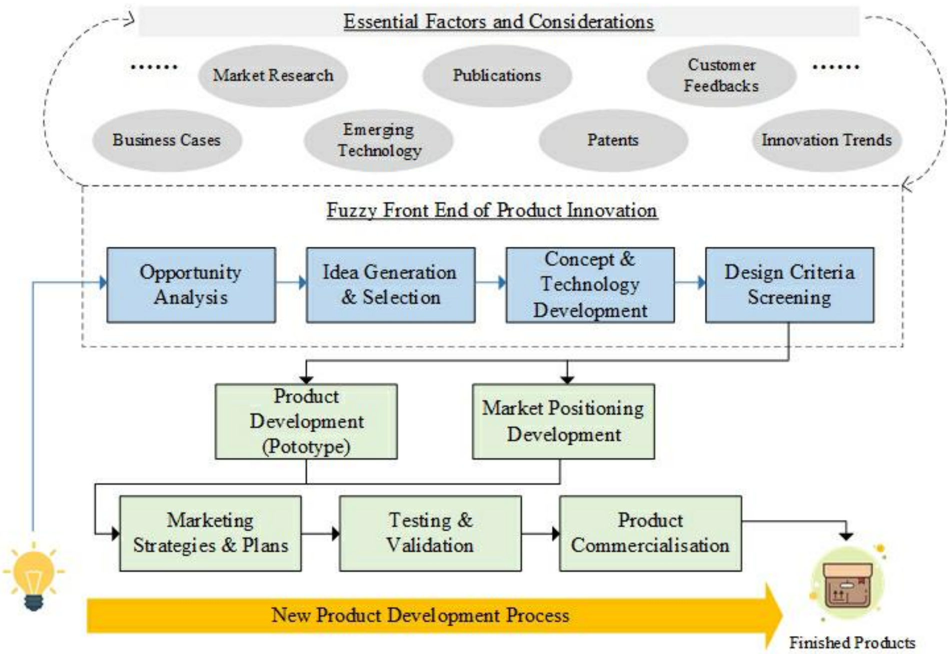

Fig. 1 Illustration of fuzzy front end in the new product development process

In addition, traditional product design is concerned mainly with the world of innovation and up-to-date manufacturing technologies, without much integration of customer needs [8]. As a result, designed products are not always needed by end customers, and thus imbalances exist between created and perceived value in product innovation. In recent years, more and more research has focused on the integration of customer perspectives and requirements in the NPD process [9-10]. Moreover, considering the current COVID-19 pandemic, data collection at the FFE stage has become extraordinarily difficult as the exchange of ideas, work arrangements, and business communication are changing continuously. Self-adaptive methodologies utilising big data analytics on factors of the FFE are more appropriate for digesting the available data to determine core feature parameters on product design. Once additional data are available, the methodologies can be effectively fine-tuned to improve the product design in a recursive manner. With consideration of FFE data, companies can better match the product design process with customer requirements and launch the "right" products for maximising the value delivered.

To address these issues, an intelligent fuzzy product design framework (I-FPDF) is proposed in this study, which incorporates fuzzy association rule mining (FARM) and a genetic algorithm (GA) into the optimal Mamdani's fuzzy inference system. Implicit fuzzy rules between FFE data and product design can be revealed through the use of FARM, such that a comprehensive knowledge repository is established. Considering the extracted fuzzy knowledge from the big data in FFE, system parameters, including types, partitions, and ranges of membership functions, can be fine-tuned. Until the convergence of minimising the system error is reached, the optimal Mamdani's inference system for the fuzzy logic can be formulated to customise the product design in the NPD process. To verify the proposed framework, a case study was conducted of a furniture manufacturer in China that produces various kinds of furniture products. The case study looked at curtain rods with suspended curtains, which are relatively mature products in the market, with limited breakthroughs. The main challenges in this industry are related to increasing customer satisfaction, expanding market share, and differentiating from other competitors by introducing new products. Therefore, successful NPD and effective product design are essential for enriching manufacturing systems in the curtain rod manufacturing industry.

This paper is organised as follows. Section 2 reviews the recent literature on NPD, big data analytics, and machine learning for fuzzy inference systems. Section 3 introduces the proposed framework of intelligent fuzzy product design. A case study of a furniture manufacturer in China is presented in Section 4. Results and discussion are provided in Section 5. Finally, conclusions are drawn in Section 6.

2. Literature Review

In today's competitive business world, it is more difficult and challenging for companies to survive as there is no single solution to win customer loyalty and market share. Recently, researchers have concluded that customer satisfaction and loyalty, driven by product performance and customer perceived value, are positively correlated and reflected in profitability performance [11-12]. To achieve better customer satisfaction, a recent study [13] proposed analysing market demand and customers' needs by using quality function deployment to establish a win-win situation between business organisations and customers. It has also been suggested that customer engagement can be fostered through enhancing the customisation and personalisation of products and services along the customer lifecycle [14]. Customer-oriented products resulting from comprehensive studies during NPD fit customers' needs and earn their loyalty [15]. In addition, developing new products is crucial to improve competitive advantages, long-term firm performance, and viability [16].

The importance of NPD is widely recognised, and recent research has actively investigated key success factors to mitigate the risks and failures of NPD. New products account for about 50% of the total sales of a company, yet only half of them can succeed in the market. Apart from the novel features of new products, various non-product-related factors, such as support from top management, external collaboration, and market analysis, can influence the success of new product designs in the market [17]. Florén et al. [18] summarised the phases of evaluating, defining, and formalising as success factors at the front end of new product development to enrich the stage-gate product development model. It was found that the front end of the NPD process is critical to successfully introducing new products, but the complexity and uncertainties lead to the occurrence of FFE. The stage-gate model is regarded as the pioneer to manage NPD-related activities and smooth the process, which can be classified into a set of discrete tasks and events [19].

Recently, the agile-stage-gate hybrid model was proposed to further improve flexibility, speed, and communication in the NPD process [20]. The exploratory product development model (ExPD) is another promising approach to reduce uncertainties and risks for adapting to the ever-changing business environment due to changes in the market, technology, regulations, and so on [21]. However, exploring the use of big data analytics and intelligent methods in managing the FFE for the NPD process is relatively rare in recent studies. Using these can contribute to formulating a self-adaptive learning mechanism for decision support in the NPD process.

Among all the phases in the NPD, the FFE, that is, phase zero, is the most complicated, which involves the processes of (i) scoping and generation of ideas, (ii) evaluation of ideas, and (iii) early prototyping for iteration [22]. Numerous opportunities and barriers, for example, information technology, market situation, and customer perspectives, must be evaluated, in which large quantities of data from multiple dimensions are considered. Referring to industrial big data, the features of 3V-3M (namely, volume, velocity, variety, multisource, multinoise, and multidimension) should be further applied in the FFE stage to generate a new synergy in the NPD process [23-24]. Table 1 summarises the features of industrial big data for the FFE stage of NPD. It shows that the data considered in the FFE stage are matched to industrial big data, and therefore corresponding big data analytics should be explored to support the decision-making process in the FEE and entire NPD process.

Machine learning in a big data environment is promising for achieving industrial intelligence in the manufacturing industry to facilitate the FFE in product design [25]. When the fuzziness is considered, particularly FFE, fuzzy inference systems are deemed feasible to provide fuzzy decision support through analysing both crisp and linguistic variables. Regarding the self-adaptive and automatic learning in the fuzzy inference systems, adaptive neuro-fuzzy inference system (ANFIS) is the most common model to solve nonlinear problems by integrating acritical neural network and Takagi-Sugeno-type fuzzy logic [26]. Although ANFIS has high generalisability to handle nonlinear and complex control problems, the computational cost is relatively high due to its complex structure and gradient learning process. Also, it is not compatible with Mamdani's fuzzy inference system for supporting the product design process. To optimise the Mamdani's fuzzy inference system for consideration in the FFE, some recent research studies have proposed the self-adaptive mechanism to obtain Mamdani fuzzy rules [27]. However, there is a lack of an integrated systematic approach with self-adaptive and automatic learning capability in obtaining Mamdani fuzzy rules and membership functions. Consequently, an integrated approach with combining big data analytics and optimisation methods to achieve optimised Mamdani fuzzy inference system should be further studied.

Table 1. Deployment of industrial big data in the FFE stage

Aspect of 3V-3M	**Description(s)**
Volume	A large quantity of analytical datasets can be obtained in the FFE stage, where some of the data are recorded in a real-time and time-series manner, for example, historical sales volume of similar products and marketing [29]. Advanced tools, including internet of things, are used to facilitate the online and offline data collection and management.
Velocity	Short data processing time is required to provide an accurate and real-time decision to support go and no-go decisions in the FFE, for example, the crowd and screening new product ideas in real time [30].
Variety	Data types and structures can be varied, in which the data can be generated by computer systems or supply chain stakeholders. The data can be presented in numerical, text, and even video manner [31].
Multisource	Data considered in the FFE can be generated from various sources, such as customer, supply chain parties, and enterprise information systems [32]. A number of internal and external sources are considered to explore the opportunities in the FFE stage.
Multinoise	Data collected in the FFE contain a certain level of random noises. For example, the sales volume and customer satisfaction levels can fluctuate due to seasonal effects [33]. This causes random errors in the collected data so that the decision-making process in FFE becomes more complicated.
Multidimension	Dimensions of data are varied, in which the data are measured by different scales, such as monetary value for sales, Likert scales for customer satisfaction. Data with different scales should be aggregated together to support the decision on product design in the FEE.

A review of recent studies in NPD found that the FFE stage is the most challenging and uncertain part of the entire NPD process. Vast quantities of industrial data have to be considered the selection of ideas and confirming designs for new products. Applying big data analytics to NPD is deemed to be promising. Big data–driven new product development is a systematic process of formulating new products for the marketplace by analysing fuzziness and big data and applying 3V-3M features. A new synergy is therefore generated to establish new products in an intelligent and optimal approach through synthesising these studies [22-24]. The contributions of this study are the integration of industrial big data and FFE in the NPD for the establishment of better product design, where fuzziness related to market, engineering, product, and customers is addressed. Because of the uncertainties and difficulties in data collection, a recursive learning approach to suggest the most appropriate product design is needed, which is investigated through combining the fuzzy-based data mining technique, namely, FARM, and a heuristics method, namely, GA. The aim of this study is to achieve continuous improvements in new product design before entering the formal product development process.

To sum up, the success of NPD has always been an active and challenging research area in the manufacturing industry, in which the complexity and dynamicity in the FFE should be addressed to increase the success rate. Fuzzy inference systems are deemed to be promising tools for analysing fuzziness in the FFE. Building an effective product design is regarded as a Mamdani's fuzzy problem [28]. To our best knowledge, an integrated and closed-loop approach for optimising Mamdani's fuzzy rules and membership functions is lacking, which is the core focus in this study to facilitate and smooth the product design process. In addition, the formulation of an effective product design framework is particularly important to maintain manufacturing companies' competitiveness in the current circumstances.

3. Framework for Intelligent Fuzzy Product Design

In this section, an intelligent fuzzy product design framework is proposed to structure the essential factors and considerations in the FFE for customising product designs in the NPD process. This framework consists of three tiers, namely (i) big data in FFE, (ii) a recursive association-rule-based fuzzy inference system (RAFIS), and (iii) Mamdani's fuzzy logic for product design customisation, as shown in Figure 2.

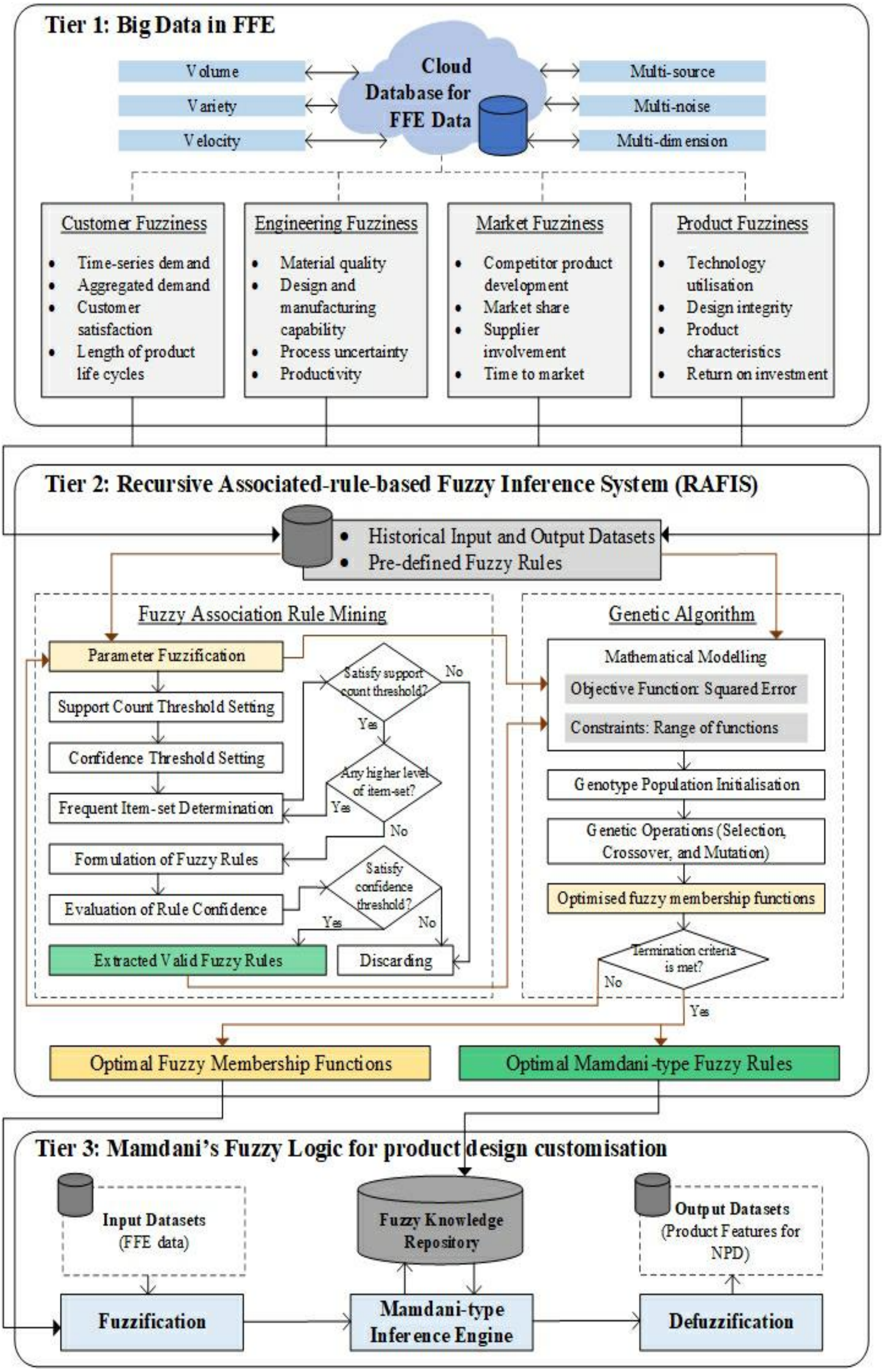

Figure 2. Three-tier structure of the I-FPDF

The data collection and management in the FFE stage are organised. The collected datasets are then inputted to FARM and GA for obtaining the optimal fuzzy rules and membership functions. Subsequently, by making use of the above results, the Mamdani's fuzzy inference system for suggesting customised product design features can be established. With the design of the proposed system, the closed-loop business process for NPD can be established to integrate data in both customer and product perspectives to enhance product competitiveness and company capability in NPD activities.

3.1 Tier 1: Big Data in FFE

The framework starts with the structural formulation of the big data in the FFE, in which four aspects, customer fuzziness, engineering fuzziness, market fuzziness, and product fuzziness, are summarised [34-36]. It has been reported that useful data and information at the FFE stage have to be extracted from a large number of databases, and a high level of uncertainties is involved in making decisions for NPD, such as idea selection and product design features. To make use of the power of big data analytics, it is necessary to quantify the above four aspects of fuzziness into a set of measurable values.

For the customer fuzziness, the uncertainties from customer responses and behaviour are summarised, including time-series demand of similar products, aggregated market demand, customer satisfaction, and length of product life cycles. For the engineering fuzziness, the uncertainties of the production process may affect the success of NPD, which may consider material quality, design and manufacturing capability, process uncertainty, and productivity. For the market fuzziness, successful new product design in an ever-changing business environment becomes challenging, such that the factors, including competitor product development, market share, supplier involvement, and time to market, can be considered. For the product fuzziness, the uncertainties with respect to the success of NPD include technology utilisation, design integrity, product characteristics for customisation, and return on investment. By quantifying the fuzziness in the FFE stage, the uncertainties in the NPD process can be further analysed by intelligent and data-driven approaches to support the product design. In view of the need to handle the large amount of data in the big data environment, the data pipeline consists of two major steps, data processing and data access, where the collected data can be evaluated through the 3V-3M principle for the establishment of industrial intelligence [37].

Data processing includes extraction, transformation, and loading (ETL), such that the data are effectively organised for entry into the structured evaluation process. In real-life situations, big data processing engines, such as Spark and Hadoop, have been developed to provide high scalability in handling big data. Hard disks may be adequate for storage of large quantities of data but there may not be available computational memory to analyse the data, and therefore the big data processing engines are adopted to ensure the data pre-processing for supporting and providing decisions. With data processing engines, the structured data are then stored in data access platforms, for example, data warehousing and NoSQL storage, for future queries and retrieval to designated application platforms. In addition, utilising a cloud-based infrastructure for big data processing and access is preferred to leverage the power of big data process engines and enable dynamic allocation of computational resources.

3.2 Tier 2: Recursive Association-Rule-Based Fuzzy Inference System (RAFIS)

To optimise the settings and system parameters for Mamdani's fuzzy inference system, a new intelligent method through the integration of FARM and GA is designed to obtain the proper settings for Mamdani's fuzzy logic. The new method is called a recursive association-rule-based fuzzy inference system (RAFIS). FARM integrates fuzzy set theory and a data mining technique, that is, association rule mining, to extract useful relationships between antecedents and consequences, where the fuzzification is improved iteratively by solving the optimisation problem of fuzzy membership functions. Before beginning RAFIS, product engineers extract historical input and output datasets, which are related to FFE characteristics and product design features, respectively. Also, predefined fuzzy rules based on the business logic are organised as the prerequisite for the entire system mechanism. The procedure for the proposed system is illustrated step by step in this section. The notations used to explain the mechanism of RAFIS are presented in Table 2. An example application of RAFIS in the furniture industry for product design customisation is presented in the next section.

Step 1: Initialise triangular membership functions F_{zi} by evenly assigning base lengths b_{zi} and ranges $r_z = [(z_k)\ , max\ (z_k)]$ based on historical input and output datasets.

Step 2: Define the threshold values of support count SC_{tz} and confidence CF_t for the rule mining process.

Step 3: Convert the crisp values of parameters Z showing the relationship between FFE data and product design features into fuzzy sets by using the predefined membership functions F_{zi} so that each parameter is represented as $\frac{\mu_{Fzi}(FC_{z1})}{FC_{z1}} + \frac{\mu_{Fzi}(FC_{z2})}{FC_{z2}} + \cdots + \frac{\mu_{Fzi}(FC_{zu})}{FC_{zu}}$ based on each row of historical input and output datasets.

Table 2. List of notations used in the RAFIS

Variables/Sets	Description(s)
Z	A set of input and output parameters, where $Z^n = \{z_k\}$ representing *n* data rows and *m* parameters for *p* input and *m* - *p* output
X	A set of input parameter which is a subset of Z, where X = $\{x_1, x_2, \ldots, x_p\}$
Y	A set of output parameter which is a subset of Z, where Y = $\{y_{m-p}, y_{m-p+1}, \ldots, y_m\}$
F_{zi}	Triangular membership function for parameter *z* at fuzzy class *i*
b_{zi}	Base length of the triangular membership function for parameter *z* at fuzzy class *i*
r_z	Range of the parameter *z*
SC_{tz}	Threshold value of support count in the rule mining process
CF_{tz}	Threshold value of confidence in the rule mining process
$\mu_{Fzi}(FC_{zu})$	Values of fuzzy membership functions F_{zi} at specific fuzzy class FC_{zu}
FC_{zu}	Fuzzy class for parameter *z* at fuzzy class *u*
SC_{zij}	Support count value for parameter *z* at fuzzy class *i* for the itemset *j*
R	A set of fuzzy rules considered in the RAFIS
$b_q^{x_i}, b_q^{y_j}$	Base length for input and output parameters for the number of fuzzy classes *q* in each parameter

Step 4: Sum the support count values SC_{zij} with the same fuzzy class FC_{ziu} across different data row z_k, that is, $SC_{zij} = \sum_{z \in N} \mu_{Fzi}(FC_{zu})$, and select the fuzzy class with the largest support count value to represent the parameters as the 1-itemset in which $SC_{zij} \geq SC_{tz}$.

Step 5: Generate higher levels of feasible itemset until (*m*-1)-itemset in which possible combinations between input and output parameters are considered, and repeat the above step 4 to calculate support count values which have to be greater than $max\ [SC_{tz1}, SC_{tz2}, \ldots, SC_{tz(m-1)}]$.

Step 6: Compute the confidence values from 2-itemset to (m-1)-itemset, where support count values of the itemset are divided by support count values of the antecedent as ($\frac{SC_{zij}^{itemset}}{SC_{zij}^{antecedent}}$), and sort out the rules that cannot meet the value of CF_{tz}.

Step 7: Obtain a list of fuzzy association rules, which are combined with the predefined fuzzy rules as a set of fuzzy rules **R** to be the prerequisite in the optimisation problem of fuzzy membership functions.

Step 8: Optimise the base length of the triangular membership functions where the optimisation problem is formulated as follows:

$$Min. \sum_{j=m-p}^{m} (y_j - \hat{y}_j)^2 \tag{1}$$

Subject to:

$$\sum\nolimits_{q \in FC_{zu}+1} b_q^{x_i} = (x_i) - min\,(x_i)\,, \forall i \in (1, \dots, p)\ and\ x_i \in Z \tag{2}$$

$$\sum\nolimits_{q \in FC_{zu}+1} b_q^{y_j} = \left(y_j\right)\ - \left(y_j\right), \forall j \in (m-p, \dots, m)\ and\ y_j \in Z \tag{3}$$

$$x_i, y_j, b_q^{x_i}, b_q^{y_j} \in R^+ \cup \{0\} \tag{4}$$

Equation (1) is the objective function of the optimisation problem to minimise the error between actual and estimated output. It examines the appropriateness of fuzzy membership functions. The estimated $\hat{y}_j$ is calculated by applying Mamdani's fuzzy inference engine based on input datasets, fuzzy rules, and membership functions [38]. Equations (2) and (3) show the constraints to the base lengths of input and output that are limited to the ranges of parameters, as shown in Figure 3. Equation (4) presents the non-negativity constraint to the input, output, and base lengths.

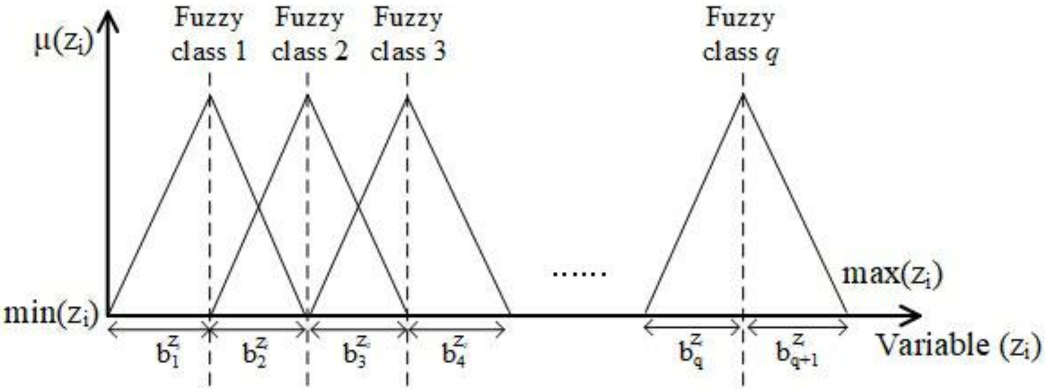

Figure 3. Graphical illustration of base length partitioning in triangular membership functions

Step 9: Establish the phenotype chromosome for the population initialisation, where the base lengths are the dependent variables in the optimisation problem. The chromosome, or the individual representation of the population, is illustrated in Figure 4. The base lengths stored in the double vector are optimised using genetic algorithms.

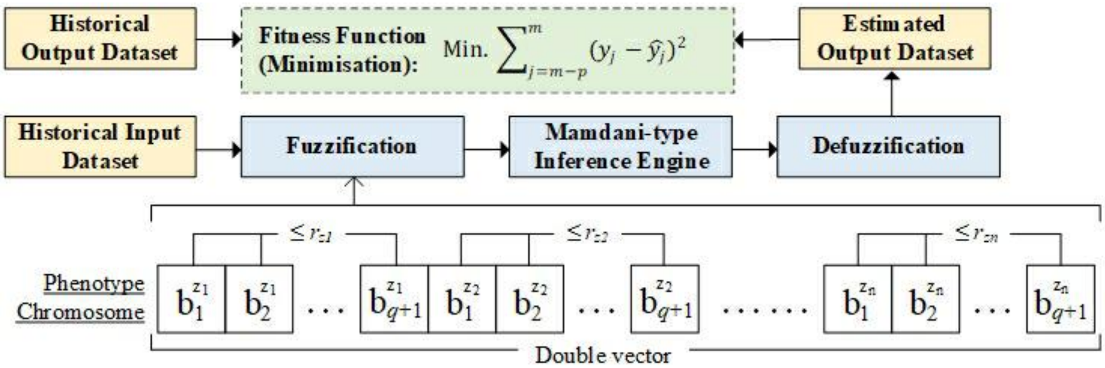

Figure 4. Graphical illustration of the individual representation

Step 10: Conduct genetic operations to search near-to-optimal solutions, including selection, crossover, and mutation, among chromosomes [39]. Subsequently, optimal settings of triangular membership functions are obtained.

Step 11: If the termination criteria, for example, errors of the fuzzy inference engine and number of iterations, are met, obtain the optimal Mamdani's fuzzy inference system. Otherwise, repeat Steps 3 to 9.

Step 12: Store the optimal settings of fuzzy membership functions and fuzzy rules in the knowledge repository for building the Mamdani's fuzzy logic further.

3.3 Tier 3: Mamdani's Fuzzy Logic for Product Design Customisation

After the operations in the RAFIS, the optimal settings in aspects of fuzzy membership functions and fuzzy IF-THEN rules (i.e., *antecedent* → *consequence*) are obtained to formulate the appropriate Mamdani's fuzzy logic as an intelligent decision support system for facilitating product design customisation. In the Mamdani's fuzzy logic, there are fuzzification, Mamdani-type fuzzy inference engine, and defuzzification to provide fuzzy decision support functionalities [38].

3.3.1 Fuzzification

Although conducting surveys and research does not make success certain, it is the only strategy to understand customer trends and behaviour, as no one can accurately predict whether new products will succeed or fail. However, the data analysed from FFE can be used as input to generate some readable outcomes through the adoption of RAFIS. As stated before, customer needs are now linked with product features so that rich customer information can be directly inputted into a fuzzy knowledge repository to support fuzzification. The FFE data, which are crisp variables, are then transformed from numbers into the degrees of membership by using optimal fuzzy membership functions within the range [0,1] for inferencing fuzzy sets and rules.

3.3.2 Mamdani-Type Fuzzy Inference Engine

The inference engine is regarded as the core element in performing a series of inference processes, including rule block formation, rule composition, rule firing, implication, and aggregation. With the adoption of the RAFIS, an optimal number of fuzzy IF-THEN rules are generated, where the rules are composited and fired in a systematic manner. When new input datasets are entered, the IF part of the rules are considered to evaluate the composited degree of belongingness μ_r^c in the membership functions to obtain composition weights for all the rules in the knowledge repository, as in equation (5). The implication operation is to select the minimum membership function values among all the weights as the composited weights for the rule **R**. Afterwards, with a set of composited weights and THEN rules, areas in the output membership functions can be aggregated.

$$\mu_r^c = \min_r\left[\mu_r(x_1), \mu_r(x_2), \ldots, \mu_r\left(x_p\right)\right] \tag{5}$$

3.3.3 Defuzzification

Defuzzification is the process of determining crisp values from the aggregated areas in the inference engine. The fuzzy sets derived from FFE data are defuzzified into measurable values related to features in product design. Regarding recent fuzzy decision support systems, various defuzzification methods are available for evaluating the aggregated areas in the output membership functions, including centre of sums (CoS), centre of gravity (CoG), centre of area (CoA), and the weighted average method. The appropriate defuzzification method can be selected according to industrial specifications to convert the fuzzy sets back to crisp values to provide fuzzy reasoning.

4. Case Study in the Furniture Industry to Examine I-FPDF

In this section, the feasibility of the proposed framework is validated in a manufacturing company located in China. The company is eager to produce new curtain-related products. The industrial motivations for adopting the proposed system is discussed first, and the I-FPDF is demonstrated in detail with the application case scenario in designing curtain rods.

4.1 Industrial Motivations

In light of capricious customer demands and significant entry barriers, those who fail to capture customer loyalty and offer high customer satisfaction find no place in the market. To introduce successful new products, the right decisions must be made at the FFE stage, and thus data-driven design innovation in the form of decision support systems is preferred in the NPD. Data-driven design innovation with decision support systems that address the above challenges would give the case company superior performance in NPD.

Facing such a highly competitive market environment, the management of the company seeks to design new products that suit potential customers as a core strategy. Currently, customers are forced to select from three ready-made models, which cannot satisfy the requirements of the majority of customers. In addition, the information in the FFE stage is relatively unstructured, which makes it difficult to formulate an effective decision support system to facilitate customisation of the product design. Therefore, manufacturing companies that are actively engaged in NPD need an intelligent approach for analysing big data in the FFE stage to enhance the features in product design.

In the case scenario, product design of curtain rods is considered. The design is subject to six functional parameters, including extendable length (EL in cm), rod perimeter (RP in cm), affordable loads (AL in kg), number of rings on rod (RR in units), ring diameter (RD in cm), and number of screws used (SU in units), as shown in Figure

5. To design the most appropriate curtain rods in the market, some significant and underlying relationships from the FFE data should be investigated for an understanding of market needs and trends.

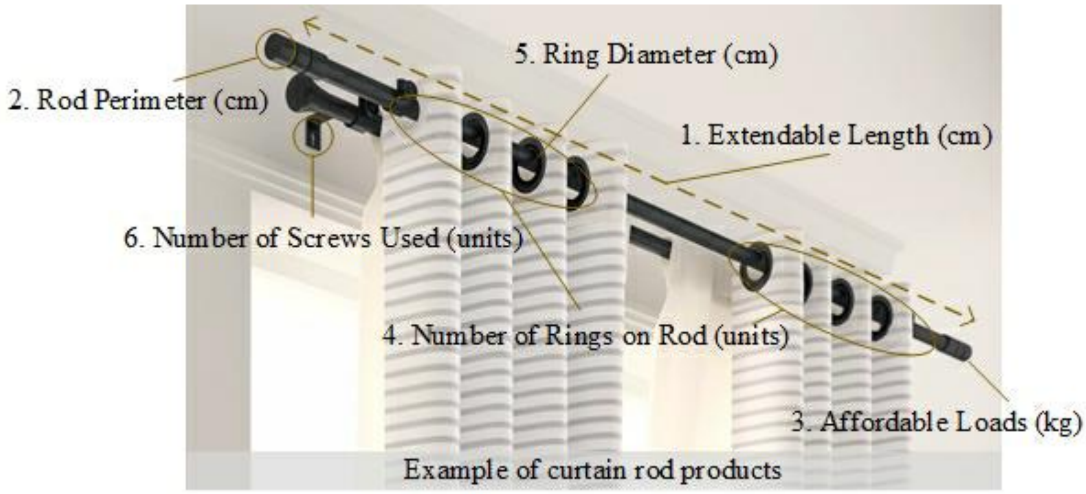

Figure 5. Six functional parameters of the design of curtain rods

4.2 Demonstration of the I-FPDF

According to the three-tier structure of the I-FPDF developed in Section 3, the implementation of the proposed framework is conducted to determine the appropriate design for the curtain rods, which are regarded as essential household appliances. To demonstrate the I-FPDF effectively, an implementation roadmap is established for the case company to walk through the steps in the proposed framework, as shown in Figure 6. A recursive learning scheme can be formulated for the product design at the FFE stage. Regarding the implementation, there are three major stages to demonstrate the application of the I-FPDF, namely (i) formulation of input and output parameters, (ii) recursive learning in the RAFIS, and (iii) adoption of a fuzzy inference system for the product design.

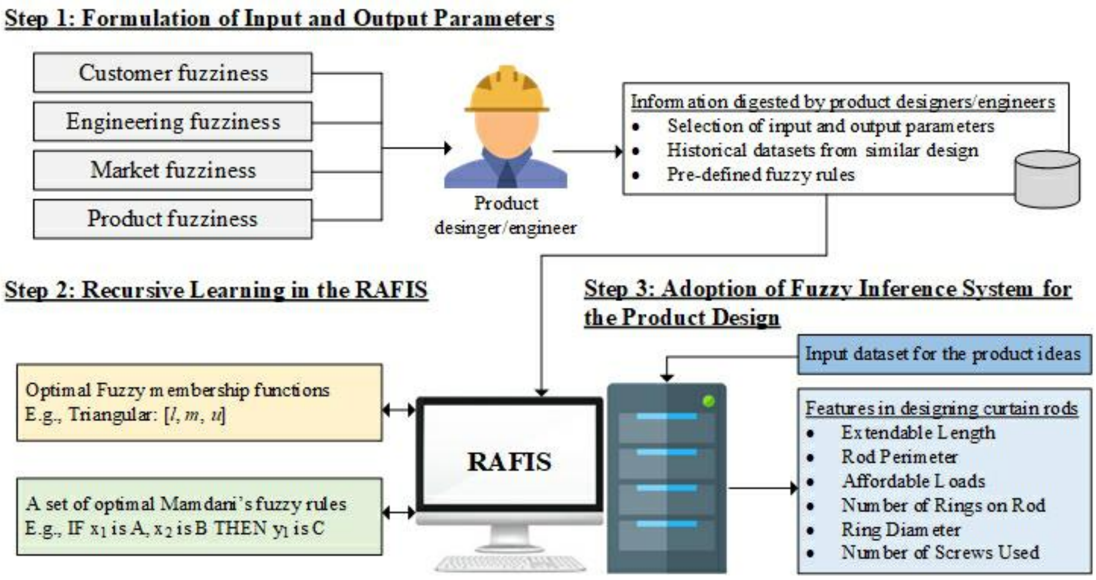

Figure 6. Implementation steps of the I-FPDF for designing curtain rods

4.2.1 Stage 1: Formulation of Input and Output Parameters

First, product designers and engineers in the case company investigate the big data at the FFE stage, including customer fuzziness, engineering fuzziness, market fuzziness, and product fuzziness, as referenced in Section 3.1. To demonstrate the proposed framework, four input parameters are considered: (i) average age of customers (AC in years), (ii) product versatility (PV in units), (iii) price level (PL in HK$), and (iv) average size of property sold (SP in ft^2), to fine-tune the six output parameters, as shown in Figure 4. AC is examined by the average age of active customers, PV is measured by the number of novel features and functions for the product design, PL is the target price of the curtain rods in HK$; SP is evaluated by the average size of property sold within one year. For training the RAFIS in the case company, 15 historical design records are retrieved, which contain the numerical values of input and output parameters, as shown in Table 3.

Table 3. Historical dataset of input and output parameters for curtain rod design

	Input Parameters				Output Parameters					
DN.	**AC**	**PV**	**PL**	**SP**	**EL**	**RP**	**AL**	**RR**	**RD**	**SU**
D1	49.4	6	370	622	45	5	4	8	6	11
D2	36.8	7	290	615	55	6.5	5	11	3.5	13
D3	54.6	6	290	683	50	5.5	4.5	9	5.5	8
D4	67.3	7	370	727	55	5	7.5	16	5	11
D5	31.4	5	230	702	20	7.5	6	14	6	7
D6	40.8	8	140	572	30	5	6	9	6	15
D7	35	7	200	554	50	5	6	10	5.5	13
D8	27.5	4	110	799	70	4.5	8	11	5	14
D9	26.2	5	380	794	40	7	8	9	5.5	13
D10	26.7	6	220	656	45	7	5	15	5	14
D11	34.5	4	310	552	40	6.5	8	10	4.5	15
D12	53.4	5	310	584	20	8	5	12	7	5
D13	55.3	5	160	702	30	6	6.5	15	4.5	11
D14	40.9	7	360	632	45	7.5	8	9	5	10
D15	29.1	3	380	581	25	5	6	15	5	11

According to the data table, the fuzzy classes and ranges for input and output parameters are initialised before entering the recursive learning process, as shown in Table 4. In addition, based on the expertise of product designers and engineers, 10 explicit fuzzy IF-THEN rules are defined in this case study based on the business logic and experience in designing curtain rods.

Table 4. Ranges and initialised membership functions for input and output parameters

Parameter	Range	Fuzzy class	Membership function
AC	[26.2, 67.3]	Young adult	[26.2, 36.475, 46.75]
		Adult	[36.475, 46.75, 57.025]
		Elder	[46.75, 57.025, 67.3]
PV	[3, 8]	Low	[3, 4.25, 5.5]
		Medium	[4.25, 5.5, 6.75]
		High	[5.5, 6.75, 8]
PL	[110, 380]	Low	[110, 177.5, 245]
		Medium	[177.5, 245, 312.5]
		High	[245, 312.5, 380]
SP	[552, 799]	Small	[552, 613.75, 675.5]
		Average	[613.75, 675.5, 737.25]
		Huge	[675.5, 737.25, 799]
EL	[20, 70]	Short	[20, 32.5, 45]
		Medium	[32.5, 45, 57.5]
		Long	[45, 57.5, 70]
RP	[4.5, 8]	Short	[4.5, 5.375, 6.25]
		Medium	[5.375, 6.25, 7.125]
		Long	[6.25, 7.125, 8]
AL	[4, 8]	Low	[4, 5, 6]
		Medium	[5, 6, 7]
		High	[6, 7, 8]
RR	[8, 16]	Few	[8, 10, 12]
		Medium	[10, 12, 14]
		Many	[12, 14, 16]
RD	[3.5, 7]	Short	[3.5, 4.375, 5.25]
		Medium	[4.375, 5.25, 6.125]
		Long	[5.25, 6.125, 7]
SU	[5, 15]	Few	[5, 7.5, 10]
		Medium	[7.5, 10, 12.5]
		Many	[10, 12.5, 15]

4.2.2 Stage 2: Recursive Learning in the RAFIS

To mine the additional fuzzy IF-THEN rules from the historical dataset, the threshold value of support count for all input and output parameters is set at 1.5, and the threshold value of confidence is set at 0.5. To start the FARM process, the crisp values in the historical dataset are converted into fuzzy sets based on the specified membership functions. An example for the first data row of design record D1 is shown in Table 5. All the data on the 15 design records can be converted into fuzzy sets, where FC1, FC2, and FC3 represent the fuzzy classes of the parameters.

Table 5. Converted fuzzy membership function values for the design record D1

	Input and Output Parameters									
	AC	**PV**	**PL**	**SP**	**EL**	**RP**	**AL**	**RR**	**RD**	**SU**
FC1	0.00	0.00	0.00	0.87	0.00	0.57	0.00	0.00	0.00	0.00
FC2	0.74	0.60	0.00	0.13	1.00	0.00	0.00	0.00	0.14	0.60
FC3	0.26	0.40	0.15	0.00	0.00	0.00	0.00	0.00	0.86	0.40

By combining all 15 fuzzy sets converted from the design records, the 1-itemset can be established through measuring the support counts of each parameter for every fuzzy class, as shown in Table 6, and the maximum support count for every parameter can be identified. After checking the support count threshold (i.e., 1.5), all the items in the 1-itemset have passed and can be used for the formulation of the 2-itemset.

Table 6. Extraction of 1st itemset from the historical data

	FC1	**FC2**	**FC3**
AC	4.744526*	2.381995	2.501217
PV	4.16837	4.23163*	3.6
PL	3.412814*	2.179778	3.185185
SP	4.199139*	3.331225	1.894737
EL	3.76837	5.63163*	1.6
RP	4.96837*	1.888773	3.142857
AL	3.46837	4.53163*	1
RR	5.46837*	1.53163	2.5
RD	4.111227	6.460202*	3.428571
SU	2.56837	3.63163	4*
Remark: Values with the sign (*) represent the highest support count value for the parameters.			

Based on the results, the co-occurrence of the members in the 1-itemset are examined to formulate the 2-itemset. The support count thresholds for the 2-itemset are compiled. Repeatedly, the higher level of itemset can be investigated until no outstanding co-occurrence itemset can be generated. According to the historical dataset, the mining process of fuzzy rules stops at the 3-itemset, and no 4-itemset can fulfil the requirement of the support count thresholds. The results of the fuzzy IF-THEN rules are summarised in Table 7. The confidence values for the rules are measured to determine the useful rules to be entered into the knowledge repository in the Mamdani's fuzzy inference system. For the fuzzy rules generated, the antecedent part contains the input parameters, while the consequence part contains only the output parameters. Therefore, five additional fuzzy rules are obtained by using the FARM, and the rules are handled in the knowledge repository to optimise the membership functions.

Table 7. Mined fuzzy rules using FARM for the Mamdani's fuzzy inference system

Generated fuzzy rules	**Rule confidence**
IF {AC is young adult} THEN {EL is medium}	0.4253
IF {AC is young adult} THEN {AL is medium}	0.4687
IF {AC is young adult} THEN {RR is few}	0.6936$^{\#}$
IF {AC is young adult} THEN {RD is medium}	0.4573
IF {AC is young adult} THEN {SU is many}	0.4336
IF {PV is medium} THEN {EL is medium}	0.5672$^{\#}$
IF {PV is medium} THEN {RD is medium}	0.5266$^{\#}$
IF {PV is medium} THEN {SU is many}	0.4254
IF {PL is low} THEN {AL is medium}	0.5372$^{\#}$
IF {SP is small} THEN {EL is medium}	0.5046$^{\#}$
IF {SP is small} THEN {RD is medium}	0.4306
IF {SP is small} THEN {SU is many}	0.4639
IF {AC is young adult} THEN {EL is medium and RR is few}	0.4005
IF {AC is young adult} THEN {RR is few and RD is medium}	0.3428
IF {PV is medium} THEN {EL is medium and RD is medium}	0.4591

Remark: the rule confidence values with the sign ($^{\#}$) denote that the threshold in rule confidence value is met.

By storing the mined fuzzy rules in the knowledge repository, the optimisation of fuzzy membership can be conducted using the historical dataset, shown in Table 3. The optimisation model defined in Steps 8 and 9 of Section 3 is built in the MATLAB environment to determine the 40 optimal base lengths for 10 input and output parameters. To optimise the double vector of the base lengths in a GA, the population size is 200; the selection function is tournament; the mutation rate is 0.01; the crossover function is single point; the maximum number of generations is 4,000. Table 8 shows the optimised base length through the GA, and therefore the fuzzy membership functions for input and output parameters can be updated accordingly, as shown in Figure 7. The fitness value is 2541.76, which is acceptable to the case company. Otherwise, the recursive learning process can be continued with additional historical datasets to locate the optimal settings for building the fuzzy inference system. With the finalised set of fuzzy rules and membership functions, the optimal design for the Mamdani's fuzzy inference system for design curtain rods can be formulated.

Table 8. Optimised base lengths for the membership functions by the GA

	Base Length			
	b_1	**b_2**	**b_3**	**b_4**
AC	30.5968	1.5626	1.5460	7.3936
PV	3.8225	0.3844	0.3690	0.4241
PL	81.9697	54.7144	54.8093	78.5066
SP	173.5293	21.1613	21.1444	31.1650
EL	0.6381	16.4340	16.4366	16.4913
RP	1.2165	0.5912	1.1060	0.5863
AL	1.6509	0.7989	0.8911	0.6573
RR	3.8666	1.3695	1.2226	1.5413
RD	2.4021	0.3546	0.3975	0.3458
SU	4.1767	1.9746	2.0189	1.9368

4.2.3 Stage 3: Adoption of Fuzzy Inference System for the Product Design

Based on the results from the RAFIS, the corresponding Mamdani's fuzzy inference system for suggesting product design features can be established. In this case study, the company is attempting to adopt the generated fuzzy inference system to determine a set of features of new curtain rods to be sold in the next year. For the new product, the configurations of the input parameters are AC is 45, PV is 4, PL is 199, and SP is 600.

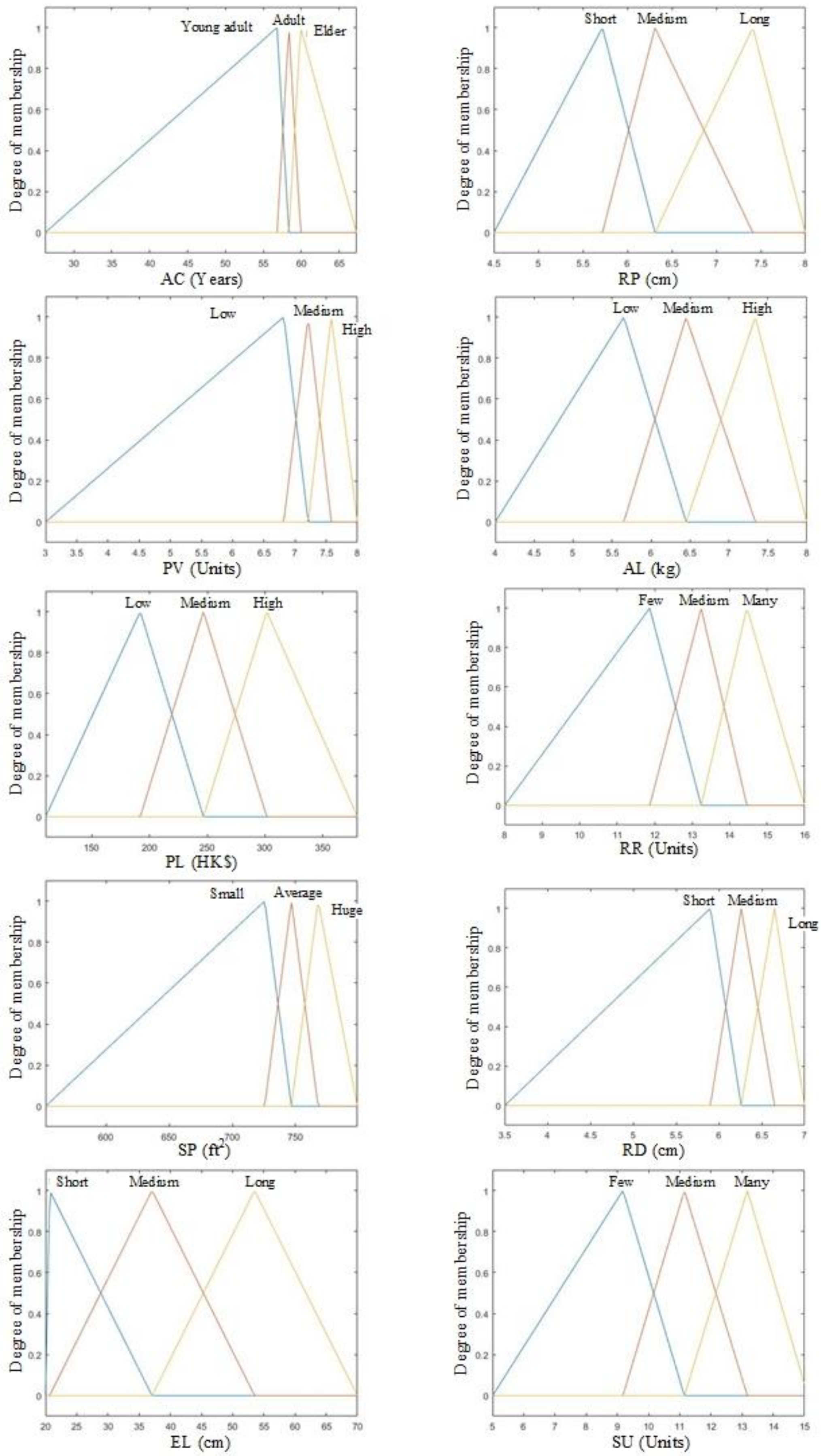

Figure 7. Optimised fuzzy membership functions from the recursive learning

By applying the optimised fuzzy inference engine, the product design features are EL is 37.1 (~38 cm), RP is 7.8 (~8 cm), AL is 6 kg, RR is 11.3 (~12 units), RD is 5.3 (~6 cm), and SU is 8.2 (~9 units). Consequently, the product designers and engineers can modify the product design based on the estimated model to further ensure its practicality and productivity in the production lines. Among numerous new ideas and market opportunities in the FFE stage of the NPD process, the proposed framework provides an intelligent and autonomous approach to formulate the appropriate product design to enter the formal product development stage.

5. Results and Discussion

This section presents the results and a discussion of the research. The design performance before and after applying the I-FPDF is measured to evaluate the performance of the proposed model. A comparison of the RAFIS with existing approaches is presented, and industrial implications are discussed.

5.1 Performance measurements of the I-FPDF

The design and market performance of the proposed framework are measured quantitatively before and after the use of I-FPDF. The specific key performance indicators (KPIs) were defined. The number of prototypes accepted for mass production (in units) and time to market (in months) were taken as the KPIs for design. Market performance was measured as the monthly recurring revenue (in HK$). It included customer retention rate (in %), average net promoter score (on a scale of 0–10), and average customer satisfaction (on a scale of 0–10), where the customer retention rate was calculated by (number of customers at the end of the period – number of new customers)/(number of customers at the start of the period). To measure the performance of the I-FPDF, the data were collected in a 3-month timeframe during the product design process in the case company. A simple survey was conducted to evaluate the net promoter score (i.e., to what extent customers were willing to recommend the products to others) and customer satisfaction (i.e., the level of customer satisfaction with the products). The measurement results are summarised in Table 9. It was found that the positive effects in the aspects of design and market performance can be obtained with 16.7% and 17.15% on average, respectively. It shows that the adoption of the proposed framework can effectively facilitate new product design, with a higher number of prototypes accepted and shorter time to market. Also, since the new product design is generated with consideration of customer preferences and requirements, the market performance of new products is improved, with higher monthly recurring revenue and customer retention rate. More customers were satisfied with new products in the case company and willing to recommend them to others, such as friends and relatives. Consequently, the proposed framework has a positive influence on the FFE stage to foster the data-driven product design process with the use of a recursive learning approach.

Table 9. Performance measurement before and after the use of I-FPDF

Area	**Before**	**After**	**% change**
Design performance:			
-Number of prototypes accepted	8	10	25.0%
-Time to market	~12	~11	8.33%
Market performance:			
-Monthly recurring revenue (HK$'000)	~280.8	~345.1	22.9%
-Customer retention rate	55.8%	63.5%	13.8%
-Average net promoter score	6.1	6.9	13.1%
-Average customer satisfaction	6.4	7.6	18.8%

5.2 Comparison between the RAFIS and existing approaches

Regarding recursive learning for handling fuzziness at the FFE stage, the RAFIS was established to create adequate reasoning on the product design features. A qualitative comparison between the proposed fuzzy inference system, ANFIS, and typical fuzzy inference systems was conducted to elucidate the value of the RAFIS. Table 10 summarises the results of the comparison for the three fuzzy inference systems.

The typical fuzzy logic can be designed for the Mamdani-type and Sugeno-type fuzzy inference engines in a multiple-input multiple-output (MIMO) manner, but the training and learning mechanism is lacking. Its effectiveness relies heavily on the domain knowledge of experts to define sufficient membership functions and fuzzy rules. The ANFIS is designated to cater to the needs of a Sugeno-type fuzzy inference engine by making use of an artificial neural network (ANN) to optimise the membership functions and Sugeno-type fuzzy rules. However, it merely supports a multiple-input single-output (MISO) system design such that multiple ANFISs must be operated concurrently if more than one output parameter is considered. Regarding the proposed RAFIS, it is targeted to build an optimal MIMO-based Mamdani-type fuzzy inference system abandoning high complexity in ANN to mine the fuzzy rules and to adjust membership. Instead, a recursive loop integrating FARM and GA is established, and thus the most appropriate fuzzy inference engine can be located.

Implementing such a recursive loop for the proposed system can reduce the time to locate the appropriate settings in fuzzy inference systems. Also, the recursion is an efficient method to develop and debug the systems by decluttering the codes. The above advantages from the proposed fuzzy inference system align with the directions of eco-innovation in the system development. Recursive learning-based approaches, as environmentally friendly technological advancements, have positive effects on sustainable NPD, particularly in the FFE stage. In future, additional modules or advanced methods to strengthen the functionalities in mining fuzzy rules and optimise membership functions can be easily integrated into the current decluttered system environment to fulfil the need for eco-design of the systems for sustainability in the manufacturing industry.

Table 10. Qualitative comparison between typical fuzzy logic, ANFIS, and RAFIS

Aspect	Typical fuzzy logic	ANFIS	RAFIS
Type of fuzzy inference	Mamdani-type & Sugeno-type	Sugeno-type	Mamdani-type
Training data required	No	Yes	Yes
Self-adaptive training	No	Yes	Yes
Learning method	N/A	ANN	Recursive loop in FARM & GA
System I/O	MIMO	MISO	MIMO
Eco-innovation	Low	Medium	High

Remarks: MIMO stands for multiple-input multiple-output; MISO stands for multiple-input single-output; ANN stands for artificial neural network; FARM stands for fuzzy association rule mining; GA stands for genetic algorithm.

5.3 Managerial implications

It has been challenging to make adequate decisions at the FFE stage for product and service innovation before beginning the formal product development process. Contributing factors are high level of uncertainties and unstructured data that is high volume, high velocity, high variety, multisource, multinoise, and multidimensional. Under the proposed framework, namely I-FPDF, adopting industrial big data aids in the evaluation of four sources of fuzziness at the FFE stage, namely customer, engineering, market, and product fuzziness. The Mamdani’s fuzzy inference engine optimised by the recursive learning-based RAFIS can provide adequate reasoning on features of product and service design. Such a recursive-learning approach offers the advantages of improving product design continuously such that ultimate product design can be effectively generated when the companies are ready to produce the prototypes and to develop market positioning. Compared with the traditional approach to determine the finalised product design in one step, the proposed framework provides greater flexibility to investigate the evolution of product design, which is subject to four defined sources of fuzziness for the FFE stage.

Recursive learning for artificial intelligence provides a self-improvement mechanism to capture ever-changing market needs for fine-tuning product design features. Because epidemics such as COVID-19 pandemic create extraordinary uncertainties in the market, the proposed framework can effectively adapt to the changes by itself to maintain a competitive edge and to understand changing market positions. Subsequently, the impact from such occurrences, which may increase the complexity and uncertainties at the FFE stage, can be reduced.

6. Concluding Remarks

This study develops a generic methodology, the intelligent fuzzy product design framework (I-FPDF), for supporting the decision-making process at the FFE stage of product design. The recursive association-rule-based fuzzy inference system (RAFIS) is established by integrating fuzzy association rule mining and a genetic algorithm for identifying fuzzy rules and optimising membership functions, respectively. Subsequently, an optimal design of Mamdani's fuzzy inference engine is formulated to provide adequate reasoning in the product design process, based on the designated input parameters.

The proposed framework was demonstrated in a case study that showed the functions serving to identify hidden and useful patterns among historical product design records and consideration of user preferences and requirements. Also, using the RAFIS, in which customer knowledge and industrial big data were effectively transformed into information for decision making concerning product features, is applicable to the development of new functional products or modification of existing ones. The methodology provides a way for firms to better understand the evolution of the business landscape and facilitate strategic decision making. The main objective is to assist companies in achieving a greater success rate, higher customer satisfaction, and, most importantly, improved overall firm performance. After the feasibility of the proposed system was verified through the case study, the main results are summarised as follows:

(i) Industrial big data, in terms of volume, velocity, variety, multi-source, multi-noise, and multi-dimensional (3V-3M), contributes to the innovation process in the FFE stage, which provides solid foundation to customise product design.

(ii) A recursive learning scheme is proposed and demonstrated to bridge the gap between customer attributes and design parameters such that the design of curtain rods is successfully determined by average age of customers, product versatility, price level, and average size of property sold.

(iii) The implementation of the proposed system for the product design can bring the positive influence to design performance and market performance so as to strengthen the capability on the evolution of product design, which further enhances customer value in the market.

(iv) The proposed RAFIS introduces the self-adaptive and automatic capabilities in the fuzzy inference process by means of fuzzy association rule mining and genetic algorithm, which drives the eco-innovation of the system development.

Although the proposed system contributing to the context of product innovation of the FFE stage from the perspective of industrial big data has been verified and evaluated, the investigation is limited in the furniture industry with limited dataset for training and verification. For future study, the proposed framework can be applied to other industries that require an intelligent NPD process to deliver new products and services, using large amounts of historical data. With the consideration of industrial big data, the customisation of product and service innovation can be further explored in other manufacturing and service businesses. Moreover, due to the eco-innovation of the RAFIS design, other optimisation algorithms, such as particle swarm optimisation, pattern search, and simulated annealing algorithm, can be plugged and replaced to achieve automatic rule learning and parameter tuning in Mamdani's fuzzy inference systems.

Acknowledgements

The authors would like to thank the Big Data Intelligence Centre of The Hang Seng University of Hong Kong, and Department of ISE, Hong Kong Polytechnic University for supporting the research. Gratitude is extended to Cardiff Business School for supporting the project. Lastly, we would like to acknowledge the financial support from the Laboratory of Artificial Intelligence in Design Limited (AiDLab) under the research programme of "Innovation in Product and Service Design" (Project Code: RP2-2).

Appendices

In this appendix, the numerical illustration of the case study is provided to better understand the whole computation process of the RAFIS. Given that the historical dataset and initial fuzzy membership functions are shown in Tables 3 and 4, the processes of deriving the results in Tables 5, 6, and 7 are presented in Appendix A.

Moreover, the computation to optimise the base length of the membership functions is presented in Appendix B.

Appendix A. Deployment of FARM for Rule Extraction

According to the historical dataset of D1 to D15 in Table 3, the corresponding membership function values are evaluated through substituting to the membership functions stated in Table 4. For example, the value of AC at D1 is 49.4, which are between the fuzzy classes "Adult" (FC2) and "Elder" (FC3). Thus, its FC2's membership function value is $(49.4 - 57.025)/(46.75 - 57.025) = 0.74$, while its FC3's membership function value is $(49.4 - 46.75)/(57.025 - 46.75) = 0.26$. Also, its FC1's membership function value (i.e., "Young adult") is zero. By repeating the above calculation, Table 5 for the design record D1 is established, while Table A.1 summarises the converted membership function values for the design records D2 to D15. Subsequently, by combining all 15 design records, the results shown in Table 6 presents the 1st itemset between the input and output parameters. Based on the results in Table 6, the evaluation of the 2nd itemset to consider all the pair combination between input and output parameters with the selected fuzzy classes which have the largest support count in the 1st itemset. For example, considering the pair combination between AC.YoungAdult and EL.Medium, its combined support count is calculated by $\min(0, 1) + \min(0.97, 0.2) + \min(0, 0.6) + \min(0, 0.2) + \min(0.51, 0) +$ $\min(0.58, 0) + \min(0.86, 0.6) + \min(0.13, 0) + \min(0, 0.6) + \min(0.05, 1) +$ $\min(0.81. 0.6) + \min(0, 0) + \min(0, 0) + \min(0.57, 1) + \min(0.28, 0) = 2.02$.

Table A.1. Converted fuzzy membership function values for the design records D2 to D15

		AC	PV	PL	SP	EL	RP	AL	RR	RD	SU
D2	**FC1**	0.97	0.00	0.00	0.98	0.00	0.00	1.00	0.50	0.00	0.00
	FC2	0.03	0.00	0.33	0.02	0.20	0.71	0.00	0.50	0.00	0.00
	FC3	0.00	0.80	0.67	0.00	0.80	0.29	0.00	0.00	0.00	0.80
D3	**FC1**	0.00	0.00	0.00	0.00	0.00	0.86	0.50	0.50	0.00	0.80
	FC2	0.24	0.60	0.33	0.88	0.60	0.14	0.00	0.00	0.71	0.20
	FC3	0.76	0.40	0.67	0.12	0.40	0.00	0.00	0.00	0.29	0.00
D4	**FC1**	0.00	0.00	0.00	0.00	0.00	0.57	0.00	0.00	0.29	0.00
	FC2	0.00	0.00	0.00	0.17	0.20	0.00	0.00	0.00	0.71	0.60
	FC3	0.00	0.80	0.15	0.83	0.80	0.00	0.50	0.00	0.00	0.40
D5	**FC1**	0.51	0.40	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.80
	FC2	0.00	0.60	0.78	0.57	0.00	0.00	1.00	0.00	0.14	0.00
	FC3	0.00	0.00	0.00	0.43	0.00	0.57	0.00	1.00	0.86	0.00
D6	**FC1**	0.58	0.00	0.44	0.32	0.80	0.57	0.00	0.50	0.00	0.00
	FC2	0.42	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.14	0.00
	FC3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.86	0.00
D7	**FC1**	0.86	0.00	0.67	0.03	0.00	0.57	0.00	1.00	0.00	0.00
	FC2	0.00	0.00	0.33	0.00	0.60	0.00	1.00	0.00	0.71	0.00
	FC3	0.00	0.80	0.00	0.00	0.40	0.00	0.00	0.00	0.29	0.80
D8	**FC1**	0.13	0.80	0.00	0.00	0.00	0.00	0.00	0.50	0.29	0.00
	FC2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.71	0.00
	FC3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.40
D9	**FC1**	0.00	0.40	0.00	0.00	0.40	0.00	0.00	0.50	0.00	0.00
	FC2	0.00	0.60	0.00	0.00	0.60	0.14	0.00	0.00	0.71	0.00
	FC3	0.00	0.00	0.00	0.08	0.00	0.86	0.00	0.00	0.29	0.80
D10	**FC1**	0.05	0.00	0.37	0.32	0.00	0.00	1.00	0.00	0.29	0.00
	FC2	0.00	0.60	0.63	0.68	1.00	0.14	0.00	0.00	0.71	0.00
	FC3	0.00	0.40	0.00	0.00	0.00	0.86	0.00	0.50	0.00	0.40
D11	**FC1**	0.81	0.80	0.00	0.00	0.40	0.00	0.00	1.00	0.86	0.00
	FC2	0.00	0.00	0.04	0.00	0.60	0.71	0.00	0.00	0.14	0.00
	FC3	0.00	0.00	0.96	0.00	0.00	0.29	0.00	0.00	0.00	0.00
D12	**FC1**	0.00	0.40	0.00	0.52	0.00	0.00	1.00	0.00	0.00	0.00
	FC2	0.35	0.60	0.04	0.00	0.00	0.00	0.00	1.00	0.00	0.00
	FC3	0.65	0.00	0.96	0.00	0.00	0.00	0.00	0.00	0.00	0.00
D13	**FC1**	0.00	0.40	0.74	0.00	0.80	0.29	0.00	0.00	0.86	0.00
	FC2	0.17	0.60	0.00	0.57	0.00	0.71	0.50	0.00	0.14	0.60
	FC3	0.83	0.00	0.00	0.43	0.00	0.00	0.50	0.50	0.00	0.40
D14	**FC1**	0.57	0.00	0.00	0.70	0.00	0.00	0.00	0.50	0.29	0.00
	FC2	0.43	0.00	0.00	0.30	1.00	0.00	0.00	0.00	0.71	1.00
	FC3	0.00	0.80	0.30	0.00	0.00	0.57	0.00	0.00	0.00	0.00
D15	**FC1**	0.28	0.00	0.00	0.47	0.40	0.57	0.00	0.00	0.29	0.00
	FC2	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.71	0.60
	FC3	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.50	0.00	0.40

By doing so, the item-sets at higher levels, such as 3rd itemset and 4th itemset, can be generated until no valid pair combinations can be obtained, as shown in Table A.3. For the valid pair combinations which pass the threshold of the support count, the rule confidence between antecedent and consequence is calculated, and the results are summarised in Table 7. For example. The confidence of "IF {AC is young adult} THEN {EL is medium}" is calculated by $2.018/4.744526 = 0.4253$.

Table A.2. Extraction of 2nd itemset from the historical data

Pair Combination	SC	Valid?	Pair Combination	SC	Valid?
{AC.YoungAdult, PV.Medium}	0.5547	No	{SP.Small, EL.Medium}	2.1190	Yes
{AC.YoungAdult, PL.Low}	1.3820	No	{SP.Small, RP.Short}	1.3973	No
{AC.YoungAdult, SP.Small}	2.2249	Yes	{SP.Small, AL.Medium}	0.8259	No
{AC.YoungAdult, EL.Medium}	2.0180	Yes	{SP.Small, RR.Few}	1.3563	No
{AC.YoungAdult, RP.Short}	1.4251	No	{SP.Small, RD.Medium}	1.8080	Yes
{AC.YoungAdult, AL.Medium}	2.2238	Yes	{SP.Small, SU.Many}	1.9482	Yes
{AC.YoungAdult, RR.Few}	3.2908	Yes	{EL.Medium, RP.Short}	1.9429	Yes
{AC.YoungAdult, RD.Medium}	2.1696	Yes	{EL.Medium, AL.Medium}	0.6000	No
{AC.YoungAdult, SU.Many}	2.0574	Yes	{EL.Medium, RR.Few}	2.9000	Yes
{PV.Medium, PL.Low}	1.1926	No	{EL.Medium, RD.Medium}	3.7143	Yes
{PV.Medium, SP.Small}	1.4340	No	{EL.Medium, SU.Many}	2.4000	Yes
{PV.Medium, EL.Medium}	2.4000	Yes	{RP.Short, AL.Medium}	2.0000	Yes
{PV.Medium, RP.Short}	1.4571	No	{RP.Short, RR.Few}	1.5714	Yes
{PV.Medium, AL.Medium}	1.1000	No	{RP.Short, RD.Medium}	2.8571	Yes
{PV.Medium, RR.Few}	1.0000	No	{RP.Short, SU.Many}	2.0571	Yes
{PV.Medium, RD.Medium}	2.2286	Yes	{AL.Medium, RR.Few}	1.5000	Yes
{PV.Medium, SU.Many}	1.8000	Yes	{AL.Medium, RD.Medium}	1.8571	Yes
{PL.Low, SP.Small}	0.6721	No	{AL.Medium, SU.Many}	1.6000	Yes
{PL.Low, EL.Medium}	0.9704	No	{RR.Few, RD.Medium}	3.0000	Yes
{PL.Low, RP.Short}	1.3016	No	{RR.Few, SU.Many}	2.2000	Yes
{PL.Low, AL.Medium}	1.8333	Yes	{RD.Medium, SU.Many}	3.3143	Yes
{PL.Low, RR.Few}	1.1111	No			
{PL.Low, RD.Medium}	1.4656	No			
{PL.Low, SU.Many}	1.4370	No			

Table A.3. Extraction of 3rd and higher-level item-sets from the historical data

Pair Combination	SC	Valid?	Pair Combination	SC	Valid?
{AC.YoungAdult, SP.Small, EL.Medium}	0.8504	No	{EL.Medium, RP.Short, RR.Few}	1.0714	No
{AC.YoungAdult, SP.Small, RD.Medium}	1.0755	No	{EL.Medium, RP.Short, RD.Medium}	1.5143	Yes
{AC.YoungAdult, SP.Small, SU.Many}	1.1633	No	{EL.Medium, RP.Short, SU.Many}	1.1714	No
{AC.YoungAdult, EL.Medium, RR.Few}	1.9000	Yes	{EL.Medium, RR.Few, RD.Medium}	2.2429	Yes
{AC.YoungAdult, EL.Medium, RD.Medium}	1.3609	No	{EL.Medium, RR.Few, SU.Many}	1.3000	No
{AC.YoungAdult, EL.Medium, SU.Many}	0.8487	No	{EL.Medium, RD.Medium, SU.Many}	1.9429	Yes
{AC.YoungAdult, AL.Medium, RR.Few}	1.3564	No	{RP.Short, AL.Medium, RR.Few}	1.0714	No
{AC.YoungAdult, AL.Medium, RD.Medium}	1.2822	No	{RP.Short, AL.Medium, RD.Medium}	1.4286	No
{AC.YoungAdult, AL.Medium, SU.Many}	1.0822	No	{RP.Short, AL.Medium, SU.Many}	1.2571	No
{AC.YoungAdult, RR.Few, RD.Medium}	1.6265	Yes	{RP.Short, RR.Few, RD.Medium}	1.2143	No
{AC.YoungAdult, RR.Few, SU.Many}	1.4265	No	{RP.Short, RR.Few, SU.Many}	0.5714	No
{AC.YoungAdult, RD.Medium, SU.Many}	1.1717	No	{RP.Short, RD.Medium, SU.Many}	1.6571	Yes
{PV.Medium, EL.Medium, RD.Medium}	1.9429	Yes	{AL.Medium, RR.Few, RD.Medium}	0.8571	No
{PV.Medium, EL.Medium, SU.Many}	1.4000	No	{AL.Medium, RR.Few, SU.Many}	0.8000	No
{PV.Medium, RD.Medium, SU.Many}	1.2857	No	{AL.Medium, RD.Medium, SU.Many}	1.2571	No
{SP.Small, EL.Medium, RD.Medium}	1.1955	No	{RR.Few, RD.Medium, SU.Many}	1.6143	Yes
{SP.Small, EL.Medium, SU.Many}	0.9482	No	{AC.YoungAdult, EL.Medium, RR.Few, RD.Medium}	1.2429	No
{SP.Small, RD.Medium, SU.Many}	0.8910	No	{PV.Medium, EL.Medium, RP.Short, RD.Medium}	0.7429	No
			{PV.Medium, EL.Medium, RD.Medium, SU.Many}	1.1429	No

Appendix B. Membership Function's Base Length Optimisation by GA

The optimisation of the base lengths of the membership functions regarding the proposed Mamdani's fuzzy inference system is conducted in MATLAB® environment. Based on the extracted fuzzy rules from the FARM and the existing fuzzy rules built by domain expert intuitively, a Mamdani's fuzzy inference system for the I-FPDF can be initiated, and thus the objective function to minimise the errors between actual and estimated output. In this study, a MATLAB function y = gaInMF(x) is formulated where the formulation of Mamdani's fuzzy inference system is used to build the objective function, and its pseudo code is illustrated as follows:

Pseudo Code 1: function y = gaInMF(x)
Load the historical dataset *hData*, and initial membership function *mf* Initiate a Mamdani's fuzzy inference system (mamfis) mamfis.Name ← "I-FPDF" range ← [min(*hData*), min(*hData*)] mamfis.addInput.Name ← AC/PV/PL/SP mamfis.addInput.dataRange ← range of AC/PV/PL/SP mamfis.addOutput.Name ← EL/RP/AL/RR/RD/SU mamfis.addOutput.dataRange ← range of EL/RP/AL/RR/RD/SU Add the initial membership functions for AC/PV/PL/SP/EL/RP/AL/RR/RD/SU Declare the type of membership function ← "trimf" Define the membership function values [l, $l+x_i$, $l+x_i+x_j$] with specific fuzzy class from *mf*, where l is the lower bound of the membership function of specific parameters, $l+x_i$ is the mid-point of the membership function with the distance x_i to the lower bound, and $l+x_i+x_j$ is the upper bound of the membership function with the distance x_j to the mid-point. Build a matrix of fuzzy rules *ruleList*, with designated weight and operator of the rules Add the *ruleList* to the mamfis Obtain the estimated output by entering the set of input parameters from *hData* Evaluate the sum of square error between actual and estimated output as y

For each parameter illustrated in Figure 3, it has three corresponding fuzzy classes, and therefore four base lengths can be adjusted to achieve the optimal settings of fuzzy membership functions. Since there are ten input and output parameters, the number of variables in the optimisation problem of @gaInMF is forty in total. The optimisation problem is subject to constraints presented in Figure 4. First, the sum of base lengths in each parameter is less than or equal to the designated parameter's range. Second, the non-negativity restriction is required for all the variables considered in the optimisation problem. When using the default settings in GA, the results presented in Table 8 are obtained to determine the optimal base lengths of the triangular membership functions.

References

[1] Peralta M, Soltero VM. Analysis of fractal manufacturing systems framework towards industry 4.0. J Manuf Syst 2020; 57: 46-60. https://doi.org/10.1016/j.jmsy.2020.08.004
[2] Zheng C, Qin X, Eynard B, Bai J, Li J, Zhang Y. SME-oriented flexible design approach for robotic manufacturing systems. J Manuf Syst 2019; 53: 62-74. https://doi.org/10.1016/j.jmsy.2019.09.010
[3] Odoom R. Brand marketing programs and consumer loyalty–evidence from mobile phone users in an emerging market. J. Prod. Brand. Manag. 2016; 25(7): 651-662. https://doi.org/10.1108/JPBM-04-2016-1141
[4] Barbosa LM, Lacerda DP, Piran FAS, Dresch A. Exploratory analysis of the variables prevailing on the effects of product modularization on production volume and efficiency. Int J Prod Econ 2017; 193: 677-690. https://doi.org/10.1016/j.ijpe.2017.08.028
[5] Schneider J, Hall J. Why Most Product Launches Fail, https://hbr.org/2011/04/why-most-product-launches-fail; 2014 [accessed 31 July 2020].
[6] Yoon C. Assumptions that led to the failure of Google Glass, https://medium.com/nyc-design/the-assumptions-that-led-to-failures-of-google-glass-8b40a07cfa1e; 2018 [accessed 31 July 2020].
[7] Takey SM, Carvalho MM. Fuzzy front end of systemic innovations: A conceptual framework based on a systematic literature review. Technol Forecast Soc Change 2016; 111: 97-109. https://doi.org/10.1016/j.techfore.2016.06.011
[8] Goffin K, Koners U, Baxter D, Van der Hoven C. Managing lessons learned and tacit knowledge in new product development. Res. Technol. Manag. 2010; 53(4): 39-51. https://doi.org/10.1080/08956308.2010.11657639
[9] Chang W, Taylor SA. The effectiveness of customer participation in new product development: A meta-analysis. J Mark 2016; 80(1): 47-64. https://doi.org/10.1509/jm.14.0057
[10] La Rocca A, Moscatelli P, Perna A, Snehota I. Customer involvement in new product development in B2B: The role of sales. Ind. Mark. Manag. 2016; 58: 45-57. https://doi.org/10.1016/j.indmarman.2016.05.014
[11] Kasiri LA, Cheng KTG, Sambasivan M, Sidin SM. Integration of standardization and customization: Impact on service quality, customer satisfaction, and loyalty. J. Retail. Consum. Serv. 2017; 35: 91-97. https://doi.org/10.1016/j.jretconser.2016.11.007
[12] Hama Kareem JA. The impact of intelligent manufacturing elements on product design towards reducing production waste. Int. J. Eng. Bus. Manag. 2019; 11. https://doi.org/10.1177/1847979019863955
[13] Fargnoli M, Costantino F, Di Gravio G, Tronci M. Product service-systems implementation: A customized framework to enhance sustainability and customer satisfaction. J. Clean. Prod. 2018; 188: 387-401. https://doi.org/10.1016/j.jclepro.2018.03.315
[14] Bleier A, De Keyser A, Verleye K. Customer engagement through personalization and customization. In: Palmatier RW, Kumar V, Harmeling CM, editors. Customer engagement marketing, Palgrave Macmillan, Cham; 2018, p. 75-94.
[15] Calantone RJ, Di Benedetto A, Rubera G. Launch activities and timing in new product development. J Glo Sch Mark Sci 2018; 28(1): 33-41. https://doi.org/10.1080/21639159.2017.1410771

[16] Arnett DB, Sandvik I., Sandvik K. Two paths to organizational effectiveness–Product advantage and life-cycle flexibility. J Bus Res 2018; 84: 285-292. https://doi.org/10.1016/j.jbusres.2017.11.010
[17] Roy S, Modak N, Dan PK. Role of Product Development Process for NPD Success in Indian Manufacturing Industries: Quality, Cost and Technological Aspects. In Advances in Simulation, Product Design and Development, Singapore: Springer; 2020, p. 583-596.
[18] Florén H, Frishammar J, Parida V, Wincent J. Critical success factors in early new product development: a review and a conceptual model. Int. Entrepreneurship Manag. J. 2018; 14(2): 411-427. https://doi.org/10.1007/s11365-017-0458-3
[19] Smolnik T, Bergmann T. Structuring and managing the new product development process-review on the evolution of the Stage-Gate® process. J Bus Chem 2020; (2): 41-57.
[20] Cooper RG. Agile–Stage-Gate Hybrids: The Next Stage for Product Development Blending Agile and Stage-Gate methods can provide flexibility, speed, and improved communication in new-product development. Res. Technol. Manag. 2016; 59(1): 21-29. https://doi.org/10.1080/08956308.2016.1117317
[21] PDMA. Exploratory PD: An Adaptive Product Development Process for a Complex World, https://community.pdma.org/knowledgehub/bok/process18/exploratory-pd; 2016 [accessed 31 July 2020].
[22] Tate M, Bongiovanni I, Kowalkiewicz M, Townson P. Managing the "Fuzzy front end" of open digital service innovation in the public sector: A methodology. Int J Inf Manage 2018; 39: 186-198. https://doi.org/10.1016/j.ijinfomgt.2017.11.008
[23] Zhong RY, Xu X, Klotz E, Newman ST. Intelligent manufacturing in the context of industry 4.0: a review. Engineering 2017; 3(5): 616-630. https://doi.org/10.1016/J.ENG.2017.05.015
[24] Wang J, Zheng P, Zhang J. Big data analytics for cycle time related feature selection in the semiconductor wafer fabrication system. Comput Ind Eng 2020; 106362. https://doi.org/10.1016/j.cie.2020.106362
[25] Liu Z, Pu J. Analysis and research on intelligent manufacturing medical product design and intelligent hospital system dynamics based on machine learning under big data. Enterp Inf Syst 2019. https://doi.org/10.1080/17517575.2019.1701713
[26] Parsaie A, Haghiabi AH, Saneie M, Torabi H. Predication of discharge coefficient of cylindrical weir-gate using adaptive neuro fuzzy inference systems (ANFIS). Front. Struct. Civ. Eng. 2017; 11(1): 111-122. https://doi.org/10.1007/s11709-016-0354-x
[27] Kacimi MA, Guenounou O, Brikh L, Yahiaoui F, Hadid N. New mixed-coding PSO algorithm for a self-adaptive and automatic learning of Mamdani fuzzy rules. Eng Appl Artif Intell 2020; 89: 103417. https://doi.org/10.1016/j.engappai.2019.103417
[28] Gologlu C, Mizrak C. An integrated fuzzy logic approach to customer-oriented product design. J. Eng. Des. 2011; 22(2): 113-127. https://doi.org/10.1080/09544820903032519
[29] Yerpude S, Singhal TK. Enhancing new product development effectiveness with internet of things origin real time data. J. Cases Inf. Technol. 2018; 20(3): 21-35. https://doi.org/10.4018/JCIT.2018070102
[30] Hoornaert S, Ballings M, Malthouse EC, Van den Poel D. Identifying new product ideas: waiting for the wisdom of the crowd or screening ideas in real time. J Prod Innov Manage 2017; 34(5): 580-597. https://doi.org/10.1111/jpim.12396

[31] Seo W, Yoon J, Park H, Coh BY, Lee JM, Kwon OJ. Product opportunity identification based on internal capabilities using text mining and association rule mining. Technol Forecast Soc Change 2016; 105: 94-104. https://doi.org/10.1016/j.techfore.2016.01.011
[32] Wowak KD, Craighead CW, Ketchen Jr DJ, Hult GTM. Toward a “theoretical toolbox” for the supplier-enabled fuzzy front end of the new product development process. J. Supply Chain Manag. 2016; 52(1): 66-81. https://doi.org/10.1111/jscm.12084
[33] Pavlyshenko BM. Machine-learning models for sales time series forecasting. Data 2019; 4(1): 15. https://doi.org/10.3390/data4010015
[34] Zhang Q, Doll WJ. The fuzzy front end and success of new product development: a causal model. Eur. J. Innov. Manag. 2001; 4(2): 95-112. https://doi.org/10.1108/14601060110390602
[35] Thanasopon B, Papadopoulos T, Vidgen R. The role of openness in the fuzzy front-end of service innovation. Technovation 2016; 47: 32-46. https://doi.org/10.1016/j.technovation.2015.11.007
[36] Zhang Q, Cao M, Doll W. Fuzzy front end of innovation: a dual theoretical rationale. J. Bus. Ind. Mark. 2019; 34(1): 176-191. https://doi.org/10.1108/JBIM-06-2017-0144
[37] Bharti AK, NehaVerma DKV. A Review on Big Data Analytics Tools in Context with Scalability. Int J Comp Sci Eng 2019; 7(2): 273-277. https://doi.org/10.26438/ijcse/v7i2.273277
[38] Erozan İ. A fuzzy decision support system for managing maintenance activities of critical components in manufacturing systems. J Manuf Syst 2019; 52: 110-120. https://doi.org/10.1016/j.jmsy.2019.06.002
[39] Chen JC, Chen YY, Chen TL, Kuo YH. Applying two-phase adaptive genetic algorithm to solve multi-model assembly line balancing problems in TFT–LCD module process. J Manuf Syst 2019; 52: 86-99. https://doi.org/10.1016/j.jmsy.2019.05.009