Partner selection in sustainable supply chains: a fuzzy ensemble learning model

Wu, C., Lin, C., Barnes, D. and Zhang,Y.

Partner selection in sustainable supply chains: a fuzzy ensemble learning model

Chong Wu [a], Chuanlin Lin [a], David Barnes [b], Yi Zhang [a]

[a] *School of Management, Xiamen University, Xiamen, 361005 P.R. China*

[b] *Westminster Business School, University of Westminster, London, NW1 5LS U.K.*

Corresponding author:

David Barnes

Address:

Westminster Business School, University of Westminster, London, NW1 5LS U.K.

Telephone: +44 (0) 20 350 66531

Fax: +44 (0) 20 7911 5839

Email: d.barnes@westminster.ac.uk

Partner selection in sustainable supply chains: a fuzzy ensemble learning model

Abstract: With the increasing demands on businesses to operate more sustainably, firms must ensure that the performance of their whole supply chain in sustainability is optimized. As partner selection is critical to supply chain management, focal firms now need to select supply chain partners that can offer a high level of competence in sustainability. This paper proposes a novel multi-partner classification model for the partner qualification and classification process, combining ensemble learning technology and fuzzy set theory. The proposed model enables potential partners to be classified into one of four categories (strategic partner, preference partner, leverage partner and routine partner), thereby allowing distinctive partner management strategies to be applied for each category. The model provides for the simultaneous optimization of both efficiency in its use of multi-partner and multi-dimension evaluation data, and effectiveness in dealing with the vagueness and uncertainty of linguistic commentary data. Compared to more conventional methods, the proposed model has the advantage of offering a simple classification and a stable prediction performance. The practical efficacy of the model is illustrated by an application in a listed electronic equipment and instrument manufacturing company based in southeastern China.

Keywords: Partner selection; Sustainable supply chains; Ensemble learning; Fuzzy set theory; Machine learning

1. Introduction

Growing public pressures from regulations, policies, NGOs, customers, as well as competitors have all caused companies to pay more attention to sustainable performance and apply the concept of sustainable supply chains (SSCs) (Soleimani et al. 2017; Shafiq et al. 2017). Environmental and social dimension requirements have extended from a single firm to the whole supply chain (Zimmer et al. 2016). In other words, these requirements not only require companies themselves to balance social, economic and environmental performance, but also require their supply chain partners to scrutinize their performance in energy efficiency, environmental protection, and corporate social responsibility (CSR) while pursuing economic objectives (Kannan 2018; Banaeian et al. 2018; Awasthi and Omrani 2019). The performance of potential partners influences the performance of the whole supply chain in terms of purchasing (Yin et al. 2016; Kazemi et al. 2018), production (Gharaei et al. 2019a; 2019b), inventory (Gharaei et al. 2019c; 2019d), distribution (Hoseini Shekarabi et al. 2019), and logistics (Rabbani et al. 2018). Also, cooperation and trust between partners is crucial for supply chain design (Sarkar and Giri 2018; Hao et al. 2018) and planning (Duan et al. 2018). Therefore, partner selection has become a critical issue for complying with sustainable supply chain management (Govindan et al. 2015; Oelze 2017). Considering how to evaluate and select appropriate partners is one of the most crucial challenges confronted by decision-makers in SSCs (Lima Junior et al. 2014; Rabbani et al. 2019).

The full process of partner selection in SSCs includes: a) criteria formulation, b) partner qualification and classification, c) final selection, and d) selection feedback (Luo et al. 2009; Wu and Barnes 2012; Zimmer et al. 2016). Most existing research in this field focuses on the third step (final selection). However, the quality of decision-making in final selection (on supply chain network design, lot-sizing coordination and other supply chain practices) has a significant relationship with the quality of previous decision-making steps. Partner qualification and classification is an indispensable prerequisite in the partner selection process (Wu and Barnes 2014).

Ineffective or inappropriate partner qualification and classification decision-making can result in many CSR and environmental incidents among supply chains. A tragic example of this was the Bangladesh garment factory collapse in 2013, which killed thousands of workers (Fox 2013; Friedman 2015). UK fashion brand Primark also suffered manufacturing disruption and brand reputation damage because of its neglect of CSR and safety issues in the process of partner qualification and classification (Hendriksz 2017). Furthermore, several other fashion brands (for instance, Zara, H&M, and Forever 21) have also been questioned on environmental concerns and labor issues involving their supply chain partners according to the 2015 documentary “The True Cost” (Morgan 2015).

Therefore, this paper addresses this problem by proposing a novel multi-partner classification model for the partner qualification and classification process. Use of the model will enable potential partners to be classified into one of four categories (strategic partner, preference partner, leverage partner and routine partner), thereby allowing distinctive partner management strategies to be applied to each category.

Historically, companies have tended to only concentrate on economic criteria such as cost, quality and lead time when evaluating partners. This approach is now considered to be outdated and there is a need to also consider sustainability criteria, especially in social and environmental factors (Kannan et al. 2013). This research follows the triple bottom line (TBL) approach proposed by Elkington (1998) to consider criteria from the economic, environmental and social dimensions of sustainability during the partner qualification and classification process. Selecting appropriate partners under triple bottom line criteria requires an effective and efficient approach capable of analyzing both qualitative and quantitative data. Accordingly, this paper proposes a model to overcome the limitations of existing research for sustainable partner selection, especially in the qualification and classification phase, by combining fuzzy set theory (FST) and ensemble learning technology (ELT).

ELT is well suited to this research problem and has shown excellent classification and prediction capability in financial, medical, social and other applications (Polikar 2006; Liang et al. 2018). Yet, ELT can only process deterministic and numerical data (Polikar 2006), rather than vague and linguistic evaluation data. In contrast, FST can convert qualitative and vague linguistic criteria and data into numeric values very efficiently (Buyukozkan and Cifci, 2012), which ELT cannot. Triple bottom line systems naturally contain criteria, in which partner performances are expressed in terms of linguistic preference, with all their vagueness and uncertainty (Wu and Barnes 2012). Therefore, using these two methods in combination can make their respective advantages complementary to each other. As far as the authors are aware, such a combination has not previously been applied in partner selection problems. Incorporating FST into ELT enhances the ability to handle qualitative performance indicators while improving the efficiency and effectiveness of the decision-making process of partner selection in SSCs.

The research aims to make both theoretical and practical contributions. From the theoretical aspect, firstly, the proposed model will be able to cope with the vagueness and uncertainty of the decision-making environment that is characteristic of partner selection in SSCs. Its use of will enable the vague and imprecise preferences of decision-makers to be captured effectively. Secondly, this is the first time that ensemble learning technology will be applied in partner selection for SSCs. In addition, by introducing FST, the proposed model will overcome the big weakness of the original ensemble learning model, namely that it can typically only handle numerical criteria and data. Thirdly, the combination of ELT and FST will enrich the categories of both qualitative and quantitative data used for inputting and so will widen the applicability of the model. Thereby, the proposed model will be capable of considering, systematically, not only quantitative but also qualitative criteria.

From the practical aspect, firstly, the proposed model will have a considerable ability to handle large quantities of data, which is a fundamental requirement of partner selection decision-making, especially during the early stages of the process (Wu and Barnes, 2011). Secondly, the proposed model will offer the flexibility of expanding or deleting the number of evaluation criteria without having to revise all evaluation knowledge, as would be the case for AHP/ANP (Analytic Hierarchy Process/Analytic Network Process. Thirdly, it will be almost a 'free-parameter' algorithm. As will be shown in Section 4 and 5, the proposed model is able to achieve favorable prediction accuracy in default parameters, which will considerably decrease the complexity of decision-making. Last, but not least, ELT can operate under conditions of partially missing data. This can be vital for partner selection when some historical data is unavailable or when companies refuse to disclose information due to concerns about commercial confidentiality.

Following this Introduction, the paper is organized as follows. Section 2 provides a comprehensive literature review on evaluation criteria construction and partner selection models. Section 3 briefly introduces the concept of FST and ensemble learning strategies and then proposes the fuzzy ensemble learning model. In Section 4, an illustrative application is provided to demonstrate the model in use. This concludes with a sensitivity analysis under various scenarios. Section 5 discusses the outcomes of the application of the model, and particularly considers managerial implications. The paper closes with section 6, which considers both the advantages and limitations of the proposed approach, and discusses future research opportunities.

2. Literature review

2.1 Formulation of criteria for sustainable partner selection and classification

Partner selection in SSCs is a multi-criteria and multi-objective problem under conditions of both vagueness and uncertainty. The analysis of how to screen criteria has been widely discussed by both researchers and practitioners since the 1960s. Since that time, the original consideration of only economic criteria has developed to incorporate a concern for sustainability, particularly as exemplified by the triple bottom line perspective. In seminal work, Dickson (1966) concluded that quality, delivery and performance history were the three most important indicators used by practitioners from a list of 23 possible economic criteria. This list has subsequently been used extensively by scholars in constructing representative and comprehensive supplier evaluation criteria systems (Weber et al. 1991; Ho et al. 2010).

In the last decade, many academics have extended the set of evaluation criteria to include green characteristics. Kuo et al. (2010) identified six dimensions, including certification requirement and the restriction of hazardous substances, within the environment category when extending partner selection criteria beyond the classic cost, delivery and quality categories. Additionally, Hsu et al. (2013) assigned 13 carbon management indicators to 3 dimensions on the basis of a literature review and presented the DEMATEL (Decision-Making Trial and Evaluation Laboratory) approach to identify the most influential indicators in carbon management within green supply chains. Based on the framework of strategic alignment, application, process and context, Jenssen and de Boer (2019) review 39 representative publications on green supplier selection from 1997 to 2017. One of the interesting findings in their research is that life cycle assessment-based criteria are the mostly commonly applied awarding criteria in supplier selection process.

More recently, greater emphasis has been placed on social and sustainability criteria for the whole supply chain. Accordingly, evaluation criteria for potential partners have been extended beyond the original economic indicators to encompass both environmental and social criteria (Awasthi et al. 2018), to include such factors as environmental management systems (Banaeian et al. 2018; Govindan et al. 2015), corporate social responsibility (Ho et al. 2010; Feng et al. 2017) and stakeholder engagement (Kannan 2018; Tseng et al. 2018). Pierre et al. (2019) and Bai et al. (2019) both recognized the importance of green design within supply chain network practices and the application of innovative technologies and devices. In short, there is a benefit to partner selection decision-makers in SSCs in being able to construct their own customized set of evaluation criteria in accordance with their specific decision-making environment.

2.2 Decision models for partner selection and classification

Following the structure of Kannan et al. (2013), it is possible to summarize and divide the existing partner selection multi-criteria decision making (MCDM) models into two broad groupings (shown as Figure 1).

[Take in Figure 1 about here.]

2.2.1 Single MCDM models for partner selection and classification

(1) Conceptual models The conceptual model is a representation of a system, made of the composition of concepts. Ehrgott et al. (2011) tested how pressure from customers, officials and employees determines supplier selection. Pedraza-Acosta et al. (2016) concluded that different competences are needed in selected partners as the product life cycle advances from product innovation, to product adoption to large-scale production. Kannan (2018) provided a decision support system based on critical success factors that incorporated the consideration of stakeholders into sustainable supplier selection. Keivanpour and Kadi (2017) presented an end-of-life complex product model formed of four essential aspects, namely operational, tactical, strategic and sustainability. Sayyadi and Awasthi (2018a) presented a simulation-based optimisation model to determine the key factors in designing sustainable transportation services. Shafiq et al. (2017), Delbufalo (2017) and Chen et al. (2017) have all advanced the understanding of sustainable supply chain performance by highlighting the possibility of all parties in a supply chain obtaining significant returns, without any single party sacrificing their own single profitability, by developing relationships based on long-term cooperation, rather than hostility.

(2) Mathematical models Mathematical models are descriptions of a system using mathematical concepts and language (Giri and Bardhan 2014; Shah et al. 2018). For example, Karaer et al. (2017) modelled the wholesale price premium and cost sharing as effect factors and developed insights into the optimal application strategies under single and competitive situations. Nematollahi et al. (2017) used mathematical experimentation to compare decentralized, centralized and collaborative models in the trade-off between order quantity and CSR investment. In short, these mathematical models have the advantage of enabling problems to be understood in terms of numerical outputs. They also enable the effects of different concepts to be directly studied. However, the disadvantage of these mathematical models is that the process of modelling necessarily simplifies real-world, partial variables in order to overcome the complexity of modelling and to lower computation cost (Nematollahi et al. 2017). Furthermore, this type of model tends to focus on pure methodology analysis, based on hypothetical examples, and so may miss the managerial implications. Greater use of real case study examples should be able to make any proposed method more reliable and convincing (Tsao 2015; Giri and Masanta 2018).

(3) Multi-objective programming models Multi-objective programming offers a capable and effective approach to simultaneously balancing the conflicting requirements associated with resource constraints (Gharaei et al. 2019a). It has been widely adopted in the field of supply chain management research, particularly to address the challenges associated with increased globalization, where it is important to trade-off the demands of customers, the profitability of companies and the environmental pressures from government legislators (Sgarbossa and Russo 2017). Nurjanni et al. (2017) developed a multi-objective model that enables the optimization of both supply chain total cost and total CO_2 emission, whilst maximizing service level and customer satisfaction. Govindan et al. (2017) provide a multi-objective model to minimize carbon emissions and the overall cost of supply chain activities, whilst maximizing the performance of all members of the supply chain. Yu et al. (2018) consider both economic and environmental criteria in order to facilitate a trade-off between carbon emissions, total supply chain profit and green factors. How to find and explain the optimal solutions is the key for the above multi-objective programming models. Under operational and disruption risks, Vahidi et al. (2018) construct a bi-objective mixed programming model for sustainable supplier selection and order allocation. Yet, if the proposed model could incorporate the quantity discounts and inventory control issues, it would be closer to the complex business practice for sustainable supplier selection and order allocation.

(4) **Fuzzy set theory models** Galo et al. (2018) propose a hesitant fuzzy group decision model for supplier categorization based on the application of ELECTRE TRI. The application of FST in the proposed model can capture the uncertainty of judgments due to the possible lack of complete information and the qualitative nature of some criteria. Khan et al. (2018) apply fuzzy Shannon entropy to formulate the sustainability criteria while using fuzzy-inference system to evaluate and select potential sustainable suppliers. Yet, one of limitations of the above research is that the finding is sensitive to the assumptions of a single evaluation framework. Liu et al. (2019) develop a three stage multi-criteria decision-making approach to select sustainable suppliers under fuzzy environment. Their proposed model can effectively identify both advantages and disadvantages of the performance of potential suppliers. However, the subjectivity of expert selection and data interpretation is one of the main limitations of any FST related models.

(5) **Grey system and rough set models** Bai and Sarkis (2010) proposed a grey system and rough number theory based multi-stage approach for partner selection. The application of rough set theory allows distillation of a larger set of suppliers into a smaller set of preferred suppliers under an uncertainty decision-making environment. More recently, Badi and Ballem (2018) integrated rough number theory with BWM, in which rough-BWM could determine intervals of expert evaluations without additional

information. To treat uncertainties and imprecisions in MCDM process, Chatterjee et al. (2018) incorporated DEMATEL and AHP method in the rough context to determine evaluation criteria weights, and finally to evaluate the potential suppliers' green performance.

2.2.2 Integrated MCDM models for partner selection and classification

(1) DEA integrated models DEA (Data Envelopment Analysis) is a non-parametric method used in operations research and economics for the estimation of performance and resource utilization efficiency (Emrouznejad and Yang, 2018). Kuo et al. (2010) proposed a green supplier selection model that integrates ANN (Artificial Neural Network) and DEA to evaluate the green performance of supply partners. The proposed model overcomes traditional DEA limitations of data accuracy and the number of decision-making units constraint. To rank sustainable suppliers and select benchmarks, Shabanpour et al. (2017) proposed a decision-making model by combining goal programming and DEA. They applied a robust CCR (Charnes-Cooper-Rhodes) inefficiency model for ranking the potential sustainable suppliers. Thus, the uncertainty of goals can be considered by running robust optimization technique.

(2) AHP/ANP integrated models Analytic network process (ANP) and analytic hierarchy process (AHP) have been applied in multi-criteria decision analysis in partner selection in the past few years in order to identify decision-makers' preferences through split pairwise comparisons (Lima Junior et al. 2014; Sayyadi and Awasthi 2018b). Shaw et al. (2012) also combined fuzzy AHP and multi-objective programming in a model that simultaneously considers classic partner evaluation criteria such as cost, lead time and carbon emission. Like Shaw et al. (2012), Ahmadi et al. (2017) combined so-called improved grey relational analysis (IGRA) into AHP approach. They firstly used AHP to calculate the selected criteria weights in sustainable partner evaluation. Then, IGRA was introduced to handle the interval number and uncertainty in the evaluation stage. The final integrated model was verified in an application to the Iranian telecom industry. AHP/ANP approaches enable complex situations to be simplified using repeated pairwise comparisons. This is effective, but not efficient as it cannot handle more than approximately fifteen criteria due to the exponential increase in comparison iterations.

(3) TOPSIS integrated models TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) is a multi-criteria decision analysis method, originally developed by Hwang and Yoon (1981). Chen (2011) applied the TOPSIS model in constructing a two-stage method for partner selection. Stage one uses SWOT analysis to identify the company's competitive strategy, which is then used to derive a framework for evaluation criteria. Then, stage two applies DEA to reduce potential suppliers to a smaller number, which are ranked using a multi-attribute decision-making approach. Govindan et al. (2013) applied FST to the process of weighting criteria in interpreting the preferences and vagueness of decision-makers based on triple-bottom-line principles. The final supplier order was calculated by using a fuzzy TOPSIS model. Through their applications to realistic problems, both Chen (2011) and Govindan et al. (2013) were able to demonstrate the capability of TOPSIS approaches in the ranking stage of their models as part of a series of partner selection procedures. More recently, Li et al. (2019) develop an extended TOPSIS method for sustainable supplier selection. Their model has noteworthy advantages in manipulating uncertainty of randomness and handling interpersonal uncertainty.

(4) DEMATEL integrated model The decision-making trial and evaluation laboratory (DEMATEL) method has been seen as one of the most appropriate methods to analyze the importance and causal relationships among different criteria. Hsu et al. (2013) applied DEMATEL to obtain the most significant indicator, as an information system in carbon management. Regarding the importance and relationships among criteria, Zhou et al. (2018) combined DEMATEL and VIKOR (VlseKriterijumska Optimizacija I Kompromisno Resenje) techniques to choose the best candidates for small-and-medium enterprises. Liu et al. (2018) applied a modification of DEMATEL and the single valued neutrosophic number (SVNN) to ranking alternative transportation providers. Both Zhou et al. (2018) and Liu et al. (2018) integrated DEMATEL with fuzzy set theory to deal with the incompleted and uncertain information and simplified the decision making process.

(5) Machine learning Machine learning is a field of computer science that uses statistical techniques to give computer systems the ability to 'learn', which is driven by the data instead of strictly static program instructions (Samuel 1959). Machine learning, which includes a list of distinct algorithms, is regarded as a powerful solution in many scientific research fields (Polikar 2006). Based on radial basis function artificial neural network, Luo et al. (2009) proposed an information-processing model which helps overcome the difficulties inherent in evaluating a large number of potential suppliers in agile supplier selection. However, their proposed model can only process quantitative criteria. To address this disadvantage, Wu and Barnes (2014) developed a fuzzy intelligent approach for partner selection in agile supply chains by combining FST with radial basis function artificial neural network. Both these two pioneering pieces of research above set an interesting direction for the application of machine learning methodology in the field of partner selection, but the effect of ensemble learning is better than the worst single classifier result (Dietterich 2000; Fernandez-Arias et al. 2018).

ELT is something of a current research hotspot in machine learning and has already been applied successfully in many fields. The essential idea of ELT is to maximize performance through combining a number of different models into the integrated application, with each component having different learning characteristics (Kuncheva 2004). Tsai et al. (2011) combined both homogeneous and heterogeneous classifiers to predict stock returns concerning prediction accuracy and Type I & II error. Holimchayachotikul et al. (2014) initiated an integrated intelligent algorithm and machine learning model to enhance the value of the whole supply chain. They also introduced new value-creation concepts from a collaborative perspective. Geng et al. (2015), Zhao et al. (2017), Fernandez-Arias et al. (2018), and Halteh et al. (2018) all applied machine learning technology to financial distress prediction, achieving comparable or better performances than other methods, thereby demonstrating the capability of ensemble learning in dealing with multi-criteria decision problems.

2.3 Summary of literature review

Partner selection is one of the most importance aspects of SSC Management because of its contribution to the sustainability profile of an organization (Bai and Sarkis 2010;

Ahmadi et al. 2017). Previously, different models have been proposed to deal with this vital problem. Table 1 offers a representative summary of models and algorithms that have been applied in the field of partner selection in recent literature. The distinctive features of each is outlined and compared to the model proposed in this paper. From the above detailed literature review and the comparisons of representative models in Table 1, we can conclude that there have been some good achievements in partner selection in SSCs in recent years, but some important research gaps remain to be addressed by further research.

[Take in Table 1 about here.]

In summary, the following four main research gaps can be identified in the existing literature:

1) Most of current research focuses only on the final selection stage of partner selection in SSCs (Zimmer et al. 2016). However, to a large extent, the decision-making quality of final selection depends on the prior decision-making phases (Wu and Barnes 2011). In addition, although Kraljic's matrix (1983) has been widely applied in purchasing and supplier management but it is a qualitative framework in nature, which restricts its veracity and objectivity. Therefore, more attention needs to be paid to the early stages of partner selection process. Specifically, in the qualification and classification phase, it would be beneficial to extend the practicability of, and enhance the objectivity of the Kraljic matrix. This presents an interesting research gap.
2) Most existing research has been predicated on conditions of certain and precise information (Shafiq et al. 2017; Sgarbossa and Russo 2017; Soleimani et al. 2017). However, decision-making in partner selection in SSCs, especially in the qualification and classification stage, is subject to much vagueness and uncertainty (Wu and Barnes 2011). As such, decision-making models and methods need to be able to cope under such conditions. At the same time, in actual business practice, a fundamental requirement of decision-making during this specific stage of partner selection process is that decision-making models and methods should have the ability to handle large quantities of data. However, most of existing approaches

cannot handle large quantities of data within a decision-making environment of vagueness and uncertainty.

3) Although ensemble learning technology has been widely applied in many research areas and has proved its worth (Polikar 2006; Barsacchi et al. 2017), it has not previously been applied in partner selection. It seems to offer great opportunities to improve the efficiency of such decision-making. Yet, ELT also has its own shortcomings when handling the vagueness of qualitative criteria and evaluation. Therefore, how to enhance its strong points and to overcome its weakness in the field of partner selection is an urgent research gap for further research.
4) Most existing models and methods have little flexibility in changing the number of quantitative and qualitative evaluation criteria without having to revise all evaluation knowledge. How to build in the flexibility of expanding or deleting the number of evaluation criteria without having to revise both quantitative and qualitative evaluation results is a very important research gap for further research.

To bridge these gaps, this research proposes a fuzzy ensemble learning model which combines FST and ELT for partner selection in SSCs. On the one hand, FST has been shown to be effective in coping with the vagueness and uncertainty inherent in expressions of decision-makers' preferences (Zadeh 1983; Buyukozkan and Cifci 2011; Govindan et al. 2017), which makes it ideal to deal with the type of qualitative indicators that are usually captured in linguistic terms (Soleimani et al. 2017; Ahmadi et al. 2017). On the other hand, machine learning technology, which has been successfully applied to the field of financial distress prediction (Halteh et al. 2018; Geng et al. 2015; Zhao et al. 2017), typically focuses on deterministic and numerical criteria but not qualitative criteria. In the proposed model, the fuzzy inference system (FIS) has been applied to transform linguistic preferences into numerical evaluations. Thus, with the help of FIS in evaluating the vagueness, ELT can process both qualitative and quantitative indicators. This means that the proposed fuzzy ensemble learning model can consider all the triple-bottom line principles in sustainable partner selection, simultaneously and effectively.

3. Fuzzy ensemble learning model for partner selection in SSCs

In this section, a fuzzy ensemble learning model for partner selection in SSCs is proposed to classify potential partners. It corresponds to the second phase of Wu and Barnes' (2012) four-phase partner selection framework.

The flowchart for the proposed model is shown in Figure 2. It comprises four stages:

Stage 1: Preparation for Partner Evaluation and Selection, in which an evaluation panel is established and a set of customized evaluation criteria are developed.
Stage 2: Establishment of the Fuzzy Inference System, in which the judgments of decision-makers are collected and quantified.
Stage 3: Constructing the Base Learners, in which base learners are trained with combined qualitative and quantitative data.
Stage 4: Application of the Fuzzy Ensemble Learning Model, in which the most appropriate ensemble strategy and settings are identified in order to exploit the efficiency and effectiveness of the proposed model.

[Take in Figure 2 about here.]

3.1 Preparation for partner evaluation and selection

The purpose of this stage is to develop customized evaluation criteria, based on the characteristics of the specific industry under consideration. The model adopts the systematic criteria construction methodology proposed by Wu and Barnes (2010; 2016). This divides the complex processes into three sub-stages: (1) General Hierarchy Criteria construction, (2) Specific Hierarchy Criteria construction, and (3) Optimization Hierarchy Criteria construction. The three-stage model combines both Dempster-Shafer belief acceptability theory and particle swarm optimization technique, which enables optimization of both efficiency in its use of limited resources during the criteria formulation process, and effectiveness in its consideration of the inter-dependence of quantitative and qualitative criteria, to be achieved simultaneously. In this research, the three-stage model incorporates the triple bottom line principle. At this point, partners who cannot satisfy minimum requirements, such as environmental protection laws and production technology standards, will be excluded from further consideration.

It is important to apply distinct partner management strategies according to the characteristics of different types of partners (Wu and Barnes 2016; 2018). Kraljic’s (1983) matrix is one of the most used models to help decision-makers in prioritizing purchase activities and managing relationships with partners. It identifies two variables, supply risk and impact on financial results, as the determining factors for categorizing all procurement relationships, namely, strategic partner, leverage partner, preference partner and routine partner. By using this classic matrix, decision-makers of SSCs can manage their potential partners more easily and effectively.

[Take in Figure 3 about here.]

3.2 Establishment of Fuzzy Inference System

In the process of evaluating performance against qualitative indicators, decision-makers use linguistic terms instead of precise numbers, in accordance with the vagueness and uncertainty which is characteristic of human reasoning (Wu and Barnes, 2014). If such subjectivity is totally ignored, even where precise numbers can be easily obtained, it would damage the effectiveness of decision-making. FST is a way of manipulating data by providing mathematical strengths to resolve the uncertainty associated with the human reasoning and judging processes (Kannan et al. 2013). It has been widely used by researchers to describe and collect the subjective and vague terms used when evaluating qualitative indicators (Buyukozkan and Cifci 2012; Govindan et al. 2017).

Fuzzy Inference System (FIS) is based on FST. It takes as its inputs the evaluation data on partners’ performance given in linguistic term and, by simulating the human reasoning, processes them with the aim of eliminating the obscurity of qualitative information. There are two classic FIS rules, namely Mamdani and TSK (Takagi-Sugeno-Kang). Compared with the outputting crisp number in TSK, the number of ‘if – then’ rules of Mamdani’s is less than TSK’s. In addition, Mamdani has advantage on global semantic definition which can clearly separate the set of possible fuzzy number

(Kroi et al. 2007). All above features promise the Mamdani's FIS achieving a delicate balance between interpretability and accuracy. Therefore, this research follows the four-step Mamdani's FIS as follows:

Sub-step 1: Fuzzify decision-makers' preference. In this research, the linear triangular function is chosen to describe and collect the judgments of decision-makers. There are two main reasons for this choice. Firstly, the triangular function can handle most fuzzy inference situations and achieve satisficed results. Secondly, it is convenient for decision-makers to use and manipulate in subsequent data processing. A triangular fuzzy number can be shown as (a, b, c); the equation and graphic being shown respectively as follows:

$$f_A(x) = \begin{cases} 0, \ x < a, \ x > c \\ \frac{x - a}{b - a}, a \leq x \leq b \\ \frac{c - x}{c - b}, b \leq x \leq c \end{cases} \quad (1)$$

In this study, five linguistic terms are regarded as control molds for describing proper linguistic variable. Table 2 defines the linguistic variables and fuzzy numbers, and Figure 4 demonstrates the setting of fuzzy set with five-level linguistic variables.

[Take in Table 2 about here.]
[Take in Figure 4 about here.]

Sub-step 2: Develop knowledge base. It is worth mentioning that Mamdani's method is capable and easily applied when the number of variables is small. Otherwise, as the number of variables in the antecedent or premise increases linearly, the total rules would increase exponentially. Table 3 shows the 'if-then' rules applied in the illustrative application section as an example. Specifically, the 'Cases' column in this table stands for the number of repeated rules due to the ignorance of the decision-makers' evaluation orders.

[Take in Table 3 about here.]

Sub-step 3: Configure inference engine. At this sub-step, the members of the evaluation panel discuss and construct their own customized fuzzy rules. And then, according to the fuzzy rules set (shown in Table 3 as an example), Figure 5 depicts the surface of the fuzzy reasoning processing rules based on forward chaining theory, which starts with known facts and then asserts new facts.

[Take in Figure 5 about here.]

Sub-step 4: Defuzzify input variables.

According to Mamdani and Assilian (1999) and Wu and Barnes (2014), the defuzzification is considered as a 'tuning' process. There are five basic methods for the defuzzying process. In general, COA (centroid of area), MOM (mean of max) and BOA (bisector of area) methods have advantages in modifying the judgment bias compared with SOM (smallest of max) and LOM (largest of max) methods. Furthermore, comparing with MOM and BOA, COA has higher capacity in overcoming the effect from extreme values, and then can aggregate a moderate output crisp number (Famuyiwa et al. 2008). Therefore, COA is the most widely used one (Amindoust et al. 2012) and is adopted as the defuzzification method in corresponding steps. The equation of COA is shown as follows.

$$x_{COA} = \frac{\sum_{i=1}^{n} x_i \times \mu_i(x_i)}{\sum_{i=1}^{n} \mu_i(x_i)} \tag{2}$$

Through the processing of fuzzy inference system, linguistic data is transformed into quantitative data.

3.3 Construction of base learners

By normalizing original quantitative data, a meta evaluation data set which consists of qualitative knowledge and quantitative information is ready for further application. Based on the meta evaluation data set, the next step is to construct base learners through cross-validation. Cross-validation is regarded as a model validation technique for assessing the statistical analysis generalized by an independent dataset and flagging the model complexity, so-called overfitting. And it is commonly used to train Tier 1 classifiers, in which the training set is randomly divided into T blocks. Each one is trained according to the remaining (T-1) blocks and is tested in the T block (the block data is not used for training). The whole cross-validation process is repeated a total of

T times because each single subset is exactly retained as the validation data for testing the model. Then the final estimation can be produced through averaging those models (Wolpert 1992).

At present, Wu et al. (2008) summarizes the top ten data mining algorithms. The supervised learning includes CART (Classification and regression tree) decision tree, SVM (Supporting vector machine), KNN (k-nearest neighbor), MLP (Multilayer perceptron), and Naïve Bayes All five of these algorithms will be tested through cross-validation respectively and demonstrate the distinct performances between models, which are the fundamentals of ELT. This is because ELT aims to mutually complement individual base learners which are characterized with diversity and different accuracy. Theoretically, the base learners' performances results in the performance of the ELT model (Wozniak et al. 2014).

3.4 Application of the Fuzzy Ensemble Learning Model

The diversity among base-learners' construct is the cornerstone of ELT, in which each classifier is expected to be as unique as possible, especially in respect of incorrect prediction or classification (Lior 2010). Based on the number of classifier types, the ensemble strategies could be divided into two clusters: one being homogeneous (Bagging and Boosting) ensemble models and the other heterogeneous (Stacking) ensemble models. `

(1) Bagging, which gains its name from bootstrap aggregating, is also known as the self-help method. It is a re-sampling method in order to get the distribution of statistics and the confidence interval (Quinlan 1996). When the ensemble learning performs in Bagging mode, several sampled weak classifiers are constructed through re-sampling the raw data of potential partner performances. This means taking a bootstrapped replica subset by randomly drawing a certain number of partners from the original dataset after determining the specified iterations, which ensures the difference between those classifiers and regarded as the basis of ensemble learning. Figure 6(a) shows the procedures of Bagging algorithms.

[Take in Figure 6 about here.]

Bagging helps to reduce the error caused by random disturbance of the training data, but if the base classifier is stable, which means little sensitivity to the data, then the Bagging method has no room for improvement in the prediction effect, because the new training sample set does not include all the original samples (Kuncheva 2004). If disturbing the learning set can cause significant changes in predictor construction, then Bagging can improve accuracy (Breiman 1996).

(2) Boosting, includes a variety of algorithms, such as AdaBoost (Adaptive Boosting), Gradient Boosting, etc. Typically, AdaBoost is the representative one, whose procedure is demonstrated in Figure 6(b). Each of its three steps are briefly explained as follows:

Firstly, at the beginning of individual classifier training, each input training sample is given the same initial weight, and the first prediction function is trained with all the data set. The error of the prediction function is calculated, which determines the weight of the prediction function in the final prediction process. Secondly, in the next iteration, the weights of sample are then updated based on the error. If the potential partners are classified into wrong categories, the weight will increase. If the sample is correctly predicted, the weight will decrease. Through the change of weights, the training model of the next round can judge the incorrectly classified potential partners better. Lastly, a new independent prediction function would be trained in each round. The weight of the newly trained prediction classifier in the final prediction is calculated according to the error of classifier prediction in each iteration. In other words, the corresponding weight of sequential classifier is modified according to its accuracy. Iteration continues until the error is less than a certain target or reaches the preset maximum number of iterations. Therefore, Adaboost algorithms use a more democratic voting scheme than Bagging algorithms, in which those base learners receiving better performance in the process of training were given greater weight rather than an equal initial value.

(3) Stacking. This is based on the differences among the model algorithms, and refers to the process of training a multi-level model consisting of distinct base learners designed for obtaining generalization. Stacking is different to Bagging or Boosting

ensemble methods, which are commonly based on majority voting, and are termed 'static' as they lack a philosophy for combining classifiers with an emphasis on training protocol. Stacked generalization is a scheme for deducting the prediction variance and biases of the generalizers. The deduction proceeds via the training set which consists of the guesses of the previous level classifier and searching for the correct prediction (Kuncheva 2004; Polikar 2006).

In Stacking ensemble learning, each training set is obtained through bootstrapped sampling on the entire training data set, which consists of whole partners' information. A series of classification models are obtained, termed Tier 1 classifiers. Then, those classifiers and their outputs are used to train the Tier 2 classifier (Meta-classifier). One of the underlying ideas is that the training data has to classify partners into specific category correctly. For example, a classifier mistakenly learns a specific area in the feature space, so the erroneous classification will come from this area, but the Tier 2 classifier may learn the correct classification based on other classifiers and modify misclassification. Some trainable ensemble strategies, including linear discrimination and machine learning algorithms, can be applied for combining classifiers. The Stacking steps are demonstrated in Figure 6(c).

In short, different training inputs, like Bagging and Boosting, are commonly used to ensure that base learners are distinguishable from each other (in order to ensure diversity). Unlike Bagging, which substantially uses a training subset, Boosting provides the same training, as changeable weights account for learner performance. However, Stacking generates sufficient differences from heterogeneous algorithms. Partner selection in SSCs is commonly confronted with vagueness and uncertainty in qualitative dimensions caused from human decision-makers' preference, unpredicted variance and drastic bias in quantitative evaluation dimensions. Ensemble learning, as a meta-classifier, combines several weak learners; it has the ability to decrease variance (via Bagging), bias (via Boosting) and improve accuracy (via Stacking). Therefore, the three ensemble strategies of Bagging, Boosting and Stacking, plus majority voting as the benchmark, will be applied in practice in partner selection in SSCs to enable realistic managerial implications to be identified.

4. Illustrative application

In this section, the capability and operability of the proposed model is demonstrated by application in the case of a real company. Company X (a pseudonym) is a giant electronic equipment manufacturer, located in southeastern China. Company X not only provides equipment and instrumentation, but also technology solution planning, to enable customers and partners to accelerate their transition to a sustainable future. Over the past two decades, the company has adopted the mantra that: “Climate change can be overcome through innovation and collaboration” and has spared no effort in the pursuit of this aim. It strongly believes that sustainability in business is a core pillar for success. The traditional relationships between partners in a supply chain environment have been adversarial, with the dominant player, the focal firm, having most power and authority and grabbing most of the available profits. Recently, Company X has been prepared to re-evaluate its partners’ performance, using a vendor management strategy, based on Kraljic’s matrix (Figure 3).

This section applies the fuzzy ensemble learning model to classify potential partners for Company X. This research uses MatLab® (from MATHWORK CO) as the platform for programming as it is both powerful and user-friendly. For instance, the defuzzifying mathematic calculation models are programmed and run by the Fuzzy Logic Toolbox within MatLab®. The reasons for choosing it as the programming environment are twofold. Firstly, MatLab® has been widely utilized by millions of users worldwide, from various backgrounds in engineering, science and economics, as their numerical computing environment. Secondly, it allows matrix manipulations, implementation of algorithms and ample machine learning applications. It also provides an interface with programs written in other languages. In short, it provides one of the most credible and compatible calculation environments for this research.

4.1 Preparation for partner evaluation and selection

To illustrate the proposed model effectiveness and efficiency, the formulation of evaluation criteria follows Wu and Barnes (2010; 2016)’s systematic methodology. The economic, social and environmental evaluation criteria set are shown in Tables 4, 5 and 6, respectively.

[Take in Tables 4 to 6 about here.]

An evaluation panel of experts with rich experience in electronic equipment manufacturing industry of China was convened. It comprised experts with different backgrounds, including purchasing managers, production managers and academicswith expertise in partner selection in SSCs. Approximately 200 potential representative companies from the electronic equipment industry in China were identified as a pool of potential partners.

For the quantitative criteria, relevant data was selected from the database provided by Wind Information Co. Ltd, which is one of the best information providers on listed companies in China. In consideration of space limitations, only the first 5 sample companies' data are shown in Table 7. For the qualitative criteria, it was decided that the evaluation panel should use one of five-level linguistic terms, 'Very Low', 'Low', 'Average', 'High' and 'Very High', to assess each company, based on their rich experience and knowledge. After determining the evaluation criteria set, all panel members' preferences were captured in linguistic terms for maximum retention of the vagueness and uncertainty in human reasoning, as discussed in section 3.2. The original evaluations are partially shown in Table 8 (social criteria) and Table 9 (environmental criteria), due to space limitations.

[Take in Tables 7 to 9 about here.]

4.2 Establishment of Fuzzy Inference system

The next stage in the process is to establish the fuzzy inference system. With the help of the Fuzzy Logic Toolbox in MATLAB, 'if-then' rules are then constructed by the Graphical Users Interface edition. Ensemble learning consisting of base learners is also constructed in the same calculation platform, to ensure model compatibility and focus on model performance.

All the 'if-then' rules are listed in Table 3. As five membership functions and three input variables are considered for this model, the total number of rules sum to 125, which represents an acceptable and moderate opinion on the selected criteria. The three decision-makers' perception of partners' performances were given the same weight. Therefore, the number of 'if-then' rules combination dropped to 60 rules only. It is worth mentioning that the rules must be verified through fuzzy engine surface, which is shown in Figure 4. From this, we can see that the combined output result increases with the increase of the inputs.

4.3 Construction of base learners

It is necessary to normalize quantitative data first and then combine them with the transformed qualitative data. A Z-Score linear normalization procedure is applied for the whole economic dataset. The normalized results are partially presented in Table 10. The results of the defuzzification of the social and environmental evaluation are partially shown in Table 11.

[Take in Tables 10 and 11 about here.]

After integrating the qualitative and quantitative data, it is advisable to review data quality for both accuracy and completeness. As for accuracy, the numerical data derives from the Wind® database, which is collected from company annual reports. As for completeness, approximately 10 values are missing in different criteria. Yet, ELT has the capability to tolerate missing values.

In this research, 10-fold cross-validation is done by partitioning the full dataset into 10 random and independent sub-datasets, using one-fold to validate the algorithm performance and using the remaining subsets to train. One round of cross validation involves all the training and testing procedures and those processes are repeated 10 times to make sure that each subset was certainly used for validation. The average cross-validation error is used as a key performance indicator. This cross-validation process is also done in MatLab to ensure compatibility.

Firstly, all the five base learners are used, MLP, SVM, CART, KNN and Naïve Bayes, which are the most popular supervised learning algorithms in pattern recognition and classification identified by Wu et al. (2008). Table 12 lists the parameters for setting individual base learners.

[Take in Table 12 about here.]

In order to assess the performance of each single algorithm, this research constructs the evaluation metrics, the so-called confusion metrics, presented in Table 13, to define the accuracy in multi-class classification problems.

[Take in Table 13 about here.]

The accuracy can be calculated by the following equation:

$$\text{Algorithm accuracy} = (A_{11} + A_{22} + A_{33} + A_{44}) / \sum A_{ij} \quad (i = 1\ldots4, j = 1\ldots4)$$

Theoretically, the more satisfying outputs the single base learner performed, the higher the probability that the final combination could remedy single learner limitations. The test subsets are a randomly selected as a 20% sample of the whole dataset. Then, the cross-validation method was used for preventing model over-fitting and obtaining the average accuracy. The initial base learner results for all five algorithms are shown in Table 14.

[Take in Table 14 about here.]

Based on Table 14, the boxplots of each base learner performances are plotted in Figure 7. The black lines on the top and bottom of the boxplot represent the inner boundary range from lower quartile minus 1.5 times interquartile to upper quartile plus 1.5 times interquartile, and red crosses represent the outliers. The tighter the boundary ranges, the more stable the base learner performances. For instance, the median and highest accuracy of MLP algorithms dominated other base learners, while its stability was less than SVM or Naïve Bayes. This information provides an intuitive perspective for comparing prediction performance of each base learner.

[Take in Figure 7 about here.]

4.4 Application of the Fuzzy Ensemble learning Model

According to the statistical data of single classifiers, shown in Table 14 and Figure 7, we can see that all algorithm performances, except SVM, fluctuate within a certain range. This is caused by imbalanced class distribution. More specifically, it is too difficult for any single classifier to solve the decision-making problem perfectly, even when plenty of data is available. Therefore, a divide-and-conquer strategy is applied in ELT modeling. This study applies the re-sample method in the form of Bagging and Adaboost to achieve diversity in homogeneous classifier ensembles. For stacking, stacked heterogeneous classifiers provide inherent diversity. The settings for the Bagging, Boosting and Stacking parameters are listed in Table 15.

[Take in Table 15 about here.]

By using Bagging to achieve the base learner diversities for constructing homogeneous combiners, this research conducted 15 rounds of bootstrapped iterations of the training data to construct distinct base learners. In addition, each base learner has the same initial parameters. Individual classifiers use majority voting to obtain the final prediction results. In other words, each base learner has the same weight at the voting stage.

Similar to Bagging, Boosting achieves its diversity through re-sampling the dataset. This research ran 20 iterations in order to generate the most informative or the most confused training data for consecutive calculation procedures. Then the Adaboost and weighted majority voting were used to construct this boosting model.

For Stacking, the research combined all five heterogeneous weak learners. Firstly, the training dataset was used to construct a single model and save the prediction results. Secondly, the first-level classifier results are used as input to train the meta classifier to bridge the gap between inputs and actual class labels. From previous literature, the ensemble strategies in the meta-classifier must be instance specific. The reason for selecting Naïve Bayes as the ensemble strategy in meta-classifier is twofold. On the one hand, Naïve Bayes algorithms could deduce the final result accounting for prior probability that is easily collected in the level-one stage (Rokach, 2010). On the other hand, Naïve Bayes could trade certain acceptable accuracy off prediction stability when applying training data (shown in the Table 14).

It should be noted that combination strategies, which provide the essential characters among three different ensemble learnings, must be compared with the benchmark from single classifier predictions. Since the purpose of this paper is to exploit and apply the ensemble learning method in partner classifications, the ensemble learning has a higher computational cost and more intricate procedures than single base learner. More detailed information about the algorithms can be seen in Polikar (2006). The prediction accuracy for various ensemble strategies is shown in Table 16.

[Take in Table 16 about here.]

According to the prediction accuracies in Table 16, all the ensemble learning models, except SVM, outperform the single classifier majority voting. Constructing classifier ensembles through the bagging strategy does not obtain better prediction accuracy than using the boosting strategy. Comparing homogeneous and heterogeneous classifiers ensembles, it is clear that heterogeneous classifier ensembles outperform homogeneous ones.

Above all, the proposed ensemble learning model achieves better performance than any single base learner in dealing with quantitative and qualitative data combined. The final best prediction accuracy reached 92.31% under the Boosting ensemble strategy (see Table 16.).

4.5 Sensitivity analysis

The purpose of sensitivity analysis is to evaluate the influence of base learners' prediction accuracy and combination strategies. The ensemble learning consists of many distinct base learners and is directly related to its component performances. The individual learner performance can be affected by initial settings, so-called key parameters. For instance, when applying MLP to conduct partner classification, the structures of MLP, learning epochs (the number of training each item of whole dataset), and the fold numbers of cross-validation should be well defined in advance (Tsai et al. 2010).

This research investigates the key parameters for each base learner, and then chooses the most influential one as the variable for analysis. For MLP algorithms, the number of 'hidden layers' (that transform inputs into those needed in the next layer) greatly affects the final prediction accuracy. Similar situations appear in CART, KNN and Naïve Bayes algorithms. Obviously, the SVM algorithms that are conducted on the modular toolbox called LibSVM, which is widely applied because of its stability, illustrate high tolerance on kernel function types (that transform original data in another dimension for maximizing the margin between classes of data). The key parameters of different learners are listed in Table 17.

[Take in Table 17 about here.]

Figure 8 illustrates five graphics that show the volatility of results when applying the different base learners. From Figure 8, we can see that all five base learners have volatility under different parameter settings. Therefore, the process of parameter setting needs more attention for initializing the base learners.

[Take in Figure 8 about here.]

After exploring the effect that the key parameters play on model performances, the best and worst scenarios are compared in Table 18. Figure 9 illustrates a comparison of different strategies within a certain classifier. Figures 10 and 11 show the results under default, worst and best parameter setting scenarios.

[Take in Table 18 about here.]
[Take in Figures 9 to 11 about here.]

As for effectiveness, all ensemble strategies whether Bagging, Boosting (see Figure 9) or Stacking (see Table 16) dominated the majority, which could be regarded as the benchmark for comparing the effect of different models. The best scenarios undoubtedly exceeded the worst ones in accuracy (both Figure 10 and 11). Specifically, the accuracy of MLP algorithms under the default and worst settings are in the same position but both achieved acceptable prediction accuracy. As for Boosting, the accuracy under default scenarios showed little difference with the best one, thanks to the high parameter tolerance within ensemble learnings. In short, by combining different base learners with reasonable ensemble strategies, the proposed ensemble learning model is more effective and efficient than existing single learner models for the classification of potential partners of SSCs.

5. Discussion

The proposed model has successfully incorporated the triple bottom line principle into decision-making for supply chain partner selection (Sarkis and Dhavale 2015). It does so by combining FST and ELT for partner qualification and classification decision-making. Specifically, in the decision-making process of classifying potential partners, it enables both numerical data and natural linguistic variables, which represent decision-maker's preferences, to be taken into consideration simultaneously. In addition, the model's capability and effectiveness in considering a large number of potential partners under dozens of criteria was verified through the illustrative application in the case of the Chinese electrical equipment and instrument manufacturing industry. The decision-making outputs enabled managers to focus on suitable suppliers and prevented considerable time and resource being wasted. In respect of applicability, the results of the sensitivity analysis demonstrate that a single base learner is susceptible to the initial key parameter. However, the final ensemble outputs show no significant change in prediction accuracy even though all the base learners were under the worst situations. This important finding means that managers could set themselves free from complex classifier selecting and time-consuming parameter setting. In comparison with other existing partner selection classification models, for instance Luo et al. (2009) and Wu and Barnes (2014), the proposed model is more reliable and applicable in SSCs decision-making environments and could be automatically operated even if supply chain managers have little knowledge of ELT or FST.

The application of the procedures of the model and the analysis of the results shown above enable some useful managerial implications to be drawn. On the one hand, from the suppliers' perspective, one of the essential characteristics of the KNN model is to classify the closest neighboring companies into the same category. This feature clearly illustrates the clustering center of distinct labels and provide geometric insights into measuring the quantitative difference. At the same time, the CART model defines the significance of evaluation attribute in partner selection through the levels of node: the closer to root node, the more important role the indicator plays. As the upstream component of the whole SSC, supplier organizations should concentrate on those 'root note' attributes and improve their shortcomings with the help of other partners, and upstream corporations should prioritize their resources and attention on reducing the negative impacts during operational practices. On the other hand, from the buyers' perspective, the clustering centers provided by the KNN algorithms can be regarded as a benchmark for buyers in the downstream supply chain to compare potential partners' performance and classify them into a select group. Furthermore, 'root node' attributes in CART algorithm should be the focus for purchasing managers, who should continuously monitor these attributes of upstream suppliers.

This research had two main objectives for decision-makers wishing to select suitable partners when considering sustainability requirements. The first is to construct a set of sustainable partner qualification and classification criteria, encompassing economic, social and environmental dimensions, considering both quantitative and qualitative criteria. The second is to enable distinctive partner management strategies to be applied in accordance with the different partner categories. For instance, once a supplier is deemed as a strategic partner, a long-term close cooperative relationship between buyer and supplier should be carefully built and maintained to enable buyers to concentrate on securing continuous and reliable supply. Such a relationship requires buyers and suppliers to work collaboratively, sharing demand and supply information to minimize supply interruption risk. In contrast, once a supplier has been identified as a non-critical and routine partner, buyers would build an arms-length formal relationship, applying a bid-by-bid policy, to reduce purchasing costs as much as possible. For a supplier classified as a preference partner, the most important issue for buyers to consider is how to control and reduce supply risk. Thus, the information provided by the proposed model offers useful managerial guidance and insights for both buyers and suppliers.

6. Conclusions

With the increased customer consciousness of sustainability and greater pressures from outside stakeholders, supply chains are gradually giving greater importance to improving their performance against sustainability criteria (Govindan et al. 2013; Dubey et al. 2015). This paper has presented a model, based on fuzzy set theory and ensemble learning technology, which enables companies to make more informed decisions about partner selection in SSCs. The practical efficacy of the model is illustrated by an application in a listed electronic equipment and instrument manufacturing company based in southeastern China.

The potential advantages of the proposed model can be summarized as follows: Firstly, the proposed model considers economic, environmental and social criteria simultaneously, thereby achieving a good balance between all triple bottom line criteria. This will benefit decision-making in partner selection, making it more sustainable. Secondly, the proposed model can consider the vagueness and uncertainty in the natural linguistic preferences given by decision-makers. On the one hand, this feature is a user-friendly setting for decision-makers. It enhances both the effectiveness and the efficiency of the process of evaluation of potential partners. On the other hand, it is also beneficial for capturing the vagueness and uncertainty of information during the decision-making process. Thirdly, the proposed model overcomes the big weakness of the existing ensemble learning model, which can typically only handle numerical quantitative data. In other words, the combination of ELT and FST widens the applicability of the proposed model by using the inputs of both qualitative and quantitative criteria and data. Fourthly, the proposed model has the capability of handling large quantities of data. This is a fundamental requirement for the early stages of partner selection (Wu and Barnes, 2011). Fifthly, without having to revise all evaluation calculations, the proposed model has the flexibility to expand or reduce the number of evaluation criteria. More importantly, ELT can operate under conditions of partially missing data, which is vital for partner selection when some historical data is unavailable or when companies refuse to disclose confidential information. Last, but not least, the proposed model can achieve favorable prediction accuracy with default parameters. This feature considerably decreases the complexity of decision-making.

There are also some limitations with the proposed model. Firstly, it requires rich and representative data for training base learners. This research acquired evaluation data from an open database, which only provides information from listed company reports. For potential partners not included in such databases, manual data collection and processing would be required. Secondly, the ensemble learning model has a higher calculation cost and more complex evaluation procedures than some other models. For instance, decision-makers need to train single learners and then combine those results using an appropriate ensemble strategy. Finally, although the proposed method is 'parameters free', decision-makers still need to construct their own customized evaluation criteria in advance.

There are a number of potential avenues for future research. Firstly, as the proposed model applies the classic fuzzy inference system to deal with qualitative criteria, new fuzzy information transforming algorithms or concepts, such as the interval linguistic model or the trapezoidal fuzzy method, could be introduced. Secondly, scholars and managers could explore more classification and intelligent algorithms and add them into the pool of base learners. More trainable ensemble strategies could also be explored to further improve classification accuracy. Thirdly, the adaptability of the proposed model could be tested through its application in different companies, industries and countries beyond the Chinese electrical equipment manufacturing industry in which this research was conducted. Fourthly, as partner evaluation requires consistent monitoring and improvement, it would be beneficial to develop dynamic evaluation techniques suitable for use in rapidly changing environments. Finally, as there are many existing models and frameworks for final partner selection after the classification phase, the issue of how the proposed model might best be used in conjunction with these requires further research.

Acknowledgement

This work was supported by the National Natural Science Foundation of China under grant number 71872155 and 71502153. We are also grateful to the anonymous reviewers for their insightful comments and constructive suggestions, which helped us improve the quality of this work.

References

Ahmadi, H. B., S. H. H. Petrudi, and X. Wang. 2017. "Integrating sustainability into supplier selection with analytical hierarchy process and improved grey relational analysis: a case of telecom industry." International Journal of Advanced Manufacturing Technology 90 (9-12): 2413-2427.

Amindoust, A., S. Ahmed, A. Saghafinia, and A. Bahreininejad. 2012. "Sustainable supplier selection: A ranking model based on fuzzy inference system." Applied Soft Computing 12 (6):1668-77.

Awasthi, A., K. Govindan, and S. Gold. 2018. "Multi-tier sustainable global supplier selection using a fuzzy AHP-VIKOR based approach." International Journal of Production Economics 195: 106-117.

Awasthi, A., and H. Omrani. 2019. "A goal-oriented approach based on fuzzy axiomatic design for sustainable mobility project selection." International Journal of Systems Science: Operations & Logistics 6 (1): 86-98.

Azadnia, A. H., M. Z. M. Saman, and K. Y. Wong. 2015. "Sustainable supplier selection and order lot-sizing: an integrated multi-objective decision-making process." International Journal of Production Research 53 (2): 383-408.

Badi, I., and M. Ballem. 2018. "Supplier selection using the rough BWM-MAIRCA model: A case study in pharmaceutical supplying in Libya." Decision Making: Applications in Management and Engineering 1 (2): 16-33.

Bai, C., and J. Sarkis. 2010a. "Green supplier development: analytical evaluation using rough set theory." Journal of Cleaner Production 18 (12):1200-10.

Bai, C., and J. Sarkis. 2010b. "Integrating sustainability into supplier selection with grey system and rough set methodologies." International Journal of Production Economics 124 (1): 252-264.

Bai, Q., Y. Gong, M. Jin, and X. Xu. 2019. "Effects of carbon emission reduction on supply chain coordination with vendor-managed deteriorating product inventory." International Journal of Production Economics 208: 83-99.

Banaeian, N., H. Mobli, B. Fahimnia, I. E. Nielsen, and M. Omid. 2018. "Green supplier selection using fuzzy group decision making methods: A case study from the agri-food industry." Computers & Operations Research 89: 337-347.

Barsacchi, M., A. Bechini, and F. Marcelloni. 2017. Multi-class Boosting with Fuzzy Decision Trees. Paper presented at the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Naples, ITALY, 2017, July 09-12.

Breiman, L. 1996. "Bagging predictors." Machine Learning 24 (2): 123-140.

Buyukozkan, G. 2012. "An integrated fuzzy multi-criteria group decision-making approach for green supplier evaluation." International Journal of Production Research 50 (11): 2892-2909.

Buyukozkan, G., and G. Cifci. 2011. "A novel fuzzy multi-criteria decision framework for sustainable supplier selection with incomplete information." Computers in Industry 62 (2): 164-174.

Buyukozkan, G., and G. Cifci. 2012. "A novel hybrid MCDM approach based on fuzzy DEMATEL, fuzzy ANP and fuzzy TOPSIS to evaluate green suppliers." Expert Systems with Applications 39 (3): 3000-3011.

Chatterjee, K., D. Pamucar, and E. K. Zavadskas. 2018. "Evaluating the performance of suppliers based on using the R'AMATEL-MAIRCA method for green supply chain implementation in electronics industry." Journal of Cleaner Production 184: 101-29.

Chen, J., Q. Hu, and J. S. Song. 2017. "Supply Chain Models with Mutual Commitments and Implications for Social Responsibility." Production and Operations Management 26 (7): 1268-1283.

Chen, Y. J. 2011. "Structured methodology for supplier selection and evaluation in a supply chain." Information Sciences 181 (9): 1651-1670.

Delbufalo, E. 2017. "The effects of suppliers' trust on manufacturers' innovation capability: an analysis of direct versus indirect relationships." Production Planning & Control 28 (14): 1165-1176.

Dickson, G. W. 1966. "An analysis of vendor selection systems and decisions." Journal of Purchasing 2 (1): 5-17.

Dietterich, T. G. 2000. Ensemble Methods in Machine Learning. International Workshop on Multiple Classifier Systems. Springer, Berlin, Heidelberg.

Duan, C., C. Deng, A. Gharaei, J. Wu, and B. Wang. 2018. "Selective maintenance scheduling under stochastic maintenance quality with multiple maintenance actions." International Journal of Production Research 56(23): 7160-7178.

Dubey, R., A. Gunasekaran, Sushil, and T. Singh. 2015. "Building theory of sustainable manufacturing using total interpretive structural modelling." International Journal of Systems Science: Operations & Logistics 2 (4): 231-247.

Ehrgott, M., F. Reimann, L. Kaufmann, and C. R. Carter. 2011. "Social Sustainability in Selecting Emerging Economy Suppliers." Journal of Business Ethics 98 (1): 99-119.

Elkington, J. 1998. "Partnerships from cannibals with forks: The triple bottom line of 21st□century business." Environmental Quality Management 8 (1): 37-51.

Emrouznejad A., and G. L. Yang. 2017. "A survey and analysis of the first 40 years of scholarly literature in DEA: 1978–2016." Socio-Economic Planning Sciences 61: 4-8.

Famuyiwa, O., L. Monplaisir, and B. Nepal. 2008. "An integrated fuzzy-goal-programming-based framework for selecting suppliers in strategic alliance formation." International Journal of Production Economics 113 (2): 862-875.

Feng, Y., Q. Zhu, and K. H. Lai. 2017. "Corporate social responsibility for supply chain management: A literature review and bibliometric analysis." Journal of Cleaner Production 158: 296-307.

Fernandez-Arias, D., M. Lopez-Martin, T. Montero-Romero, F. Martinez-Estudillo, and F. Fernandez-Navarro. 2018. "Financial Soundness Prediction Using a Multi-classification Model: Evidence from Current Financial Crisis in OECD Banks." Computational Economics 52 (1): 275-297.

Fox, E. J., 2013. "Bangladesh factory collapse kills at least 160, reviving safety questions." https://money.cnn.com/2013/04/24/news/companies/bangladesh-factory-collapse/index.html

Friedman, V. 2015. "The True Cost, a Different Kind of Fashion Documentary." https://www.nytimes.com/2015/05/29/fashion/the-true-cost-a-different-kind-of-fashion-documentary.html

Galo, N. R., L. D. Del Rosso Calache, and L. C. Ribeiro Carpinetti. 2018. "A group decision approach for supplier categorization based on hesitant fuzzy and ELECTRE TRI." International Journal of Production Economics 202: 182-196.

Geng, R., I. Bose, and X. Chen. 2015. "Prediction of financial distress: An empirical study of listed Chinese companies using data mining." European Journal of Operational Research 241 (1): 236-247.

Ghadimi, P., A. H. Azadnia, C. Heavey, A. Dolgui, and B. Can. 2016. "A review on the buyer-supplier dyad relationships in sustainable procurement context: past, present and future." International Journal of Production Research 54 (5): 1443-1462.

Gharaei A., S. A. Hoseini Shekarabi, and M. Karimi. 2019a. "Modelling and optimal lot-sizing of the replenishments in constrained, multi-product and bi-objective EPQ models with defective products: Generalised Cross Decomposition." International Journal of Systems Science: Operations & Logistics 1-13.

Gharaei, A., S. A. Hoseini Shekarabi, M. Karimi, E. Pourjavad, and A. Amjadian. 2019b. "An integrated stochastic EPQ model under quality and green policies: generalised cross decomposition under the separability approach." International Journal of Systems Science: Operations & Logistics 1-13.

Gharaei, A., M. Karimi, and S. A. H. Shekarabi. 2019c. "An integrated multi-product, multi-buyer supply chain under penalty, green, and quality control polices and a vendor managed inventory with consignment stock agreement: The outer approximation with equality relaxation and augmented penalty algorithm." Applied Mathematical Modelling 69: 223-254.

Gharaei, A., Karimi, M., and S. A. H. Shekarabi. 2019d. "Joint economic lot-sizing in multi-product multi-level integrated supply chains: Generalized benders decomposition." International Journal of Systems Science: Operations & Logistics 1-17.

Giri, B. C., and S. Bardhan. 2014. "Coordinating a supply chain with backup supplier through buyback contract under supply disruption and uncertain demand." International Journal of Systems Science: Operations & Logistics 1(4): 193-204.

Giri, B. C., and M. Masanta. 2018. "Developing a closed-loop supply chain model with price and quality dependent demand and learning in production in a stochastic environment." International Journal of Systems Science: Operations & Logistics 1-17.

Govindan, K., J. D. Darbari, V. Agarwal, and P. C. Jha. 2017. "Fuzzy multi-objective approach for optimal selection of suppliers and transportation decisions in an eco-efficient closed loop supply chain network." Journal of Cleaner Production 165: 1598-619.

Govindan, K., M. Kadzinski, and R. Sivakumar. 2017. "Application of a novel PROMETHEE-based method for construction of a group compromise ranking to prioritization of green suppliers in food supply chain." Omega-International Journal of Management Science 71: 129-145.

Govindan, K., R. Khodaverdi, and A. Jafarian. 2013. "A fuzzy multi criteria approach for measuring sustainability performance of a supplier based on triple bottom line approach." Journal of Cleaner Production 47: 345-354.

Govindan, K., S. Rajendran, J. Sarkis, and P. Murugesan. 2015. "Multi criteria decision making approaches for green supplier evaluation and selection: a literature review." Journal of Cleaner Production 98: 66-83.

Halteh, K., K. Kumar, and A. Gepp. 2018. "Financial distress prediction of Islamic banks using tree-based stochastic techniques." Managerial Finance 2018.

Hao, Y., P. Helo, and A. Shamsuzzoha. 2018. "Virtual factory system design and implementation: Integrated sustainable manufacturing." International Journal of Systems Science: Operations & Logistics 5(2): 116-132.

Hendriksz, V., 2017. "A closer look at Primark's stance on responsible fashion." https://fashionunited.uk/primark-sustainability

Ho, W., X. Xu, and P. K. Dey. 2010. "Multi-criteria decision making approaches for supplier evaluation and selection: A literature review." European Journal of Operational Research 202 (1): 16-24.

Holimchayachotikul, P., R. Derrouiche, D. Damand, and K. Leksakul. 2014. "Value creation through collaborative supply chain: holistic performance enhancement road map." Production Planning & Control 25 (11): 912-922.

Hoseini Shekarabi, S. A., A. Gharaei and M. Karimi. 2019. "Modelling and optimal lot-sizing of integrated multi-level multi-wholesaler supply chains under the shortage and limited warehouse space: Generalised outer approximation." International Journal of Systems Science: Operations & Logistics 6(3): 237-257.

Hsu, C. W., T. C. Kuo, S. H. Chen, and A. H. Hu. 2013. "Using DEMATEL to develop a carbon management model of supplier selection in green supply chain management", Journal of Cleaner Production 56: 164-172.

Hwang, C.L. and Yoon, K. 1981. "Multiple Attribute Decision Making: Methods and Applications". Springer-Verlag, New York.

Jenssen, M.M. and L. de Boer. 2019. "Implementing life cycle assessment in green supplier selection: A systematic review and conceptual model." Journal of Cleaner Production 229: 1198-1210.

Kafa, N., Y. Hani, and A. El Mhamedi. 2015. "An integrated sustainable partner selection approach with closed-loop supply chain network configuration." IFAC Papersonline 48 (3): 1840-5.

Kannan, D. 2018. "Role of multiple stakeholders and the critical success factor theory for the sustainable supplier selection process." International Journal of Production Economics 195: 391-418.

Kannan, D., R. Khodaverdi, L. Olfat, A. Jafarian, and A. Diabat. 2013. "Integrated fuzzy multi criteria decision making method and multi-objective programming approach for supplier selection and order allocation in a green supply chain." Journal of Cleaner Production 47: 355-367.

Karaer, O., T. Kraft, and J. Khawam. 2017. "Buyer and Nonprofit Levers to Improve Supplier Environmental Performance." Production and Operations Management 26 (6): 1163-1190.

Kazemi, N., S. H. Abdul-Rashid, R. A. R. Ghazilla, E. Shekarian, and S. Zanoni. 2018. "Economic order quantity models for items with imperfect quality and emission considerations." International Journal of Systems Science: Operations & Logistics 5 (2): 99-115.

Keivanpour, S., and D. A. Kadi. 2017. "Modelling end of life phase of the complex products: the case of end of life aircraft." International Journal of Production Research 55 (12): 3577-3595.

Khan, S.A., S. Kusi-Sarpong, F.K. Arhin, and H. Kusi-Sarpong 2018. "Supplier sustainability performance evaluation and selection: A framework and methodology." Journal of Cleaner Production 205: 964-979.

Kraljic, P. (1983). "Purchasing must become supply management." Harvard business review 61 (5): 109-117.

Kroi, D., T. Lasota, B. Trawinski, and K. Trawinski. 2007. "Comparison of mamdani and TSK fuzzy models for real estate appraisal." Knowledge-Based Intelligent Information and Engineering Systems: Part III, Proceedings 4694:1008

Kuncheva, L. I. 2004. Combining pattern classifiers: methods and algorithms: John Wiley & Sons.

Kuo, R. J., Y. C. Wang, and F. C. Tien. 2010. "Integration of artificial neural network and MADA methods for green supplier selection." Journal of Cleaner Production 18 (12): 1161-1170.

Li, J., H. Fang, and W. Song. 2019. "Sustainable supplier selection based on SSCM practices: A rough cloud TOPSIS approach." Journal of Cleaner Production 222: 606-621.

Liang, D., C.-F. Tsai, A.J. Dai, and W. Eberle. 2018. "A novel classifier ensemble approach for financial distress prediction." Knowledge and Information Systems 54 (2): 437-462.

Lima Junior, F. R., L. Osiro, and L. C. Ribeiro Carpinetti. 2014. "A comparison between Fuzzy AHP and Fuzzy TOPSIS methods to supplier selection." Applied Soft Computing 21: 194-209.

Liu, A., Y. Xiao, H. Lu, S. B. Tsai, and W. Song. 2019. "A fuzzy three-stage multi-attribute decision-making approach based on customer needs for sustainable supplier selection." Journal of Cleaner Production 239: 118043

Liu, F., G. Aiwu, V. Lukovac, and M. Vukic. 2018. "A multicriteria model for the selection of the transport service provider: A single valued neutrosophic DEMATEL multicriteria model." Decision Making: Applications in Management and Engineering 1 (2): 121-130.

Lior, R. 2010. "Ensemble-based classifiers." Artificial Intelligence Review 33 (1-2): 1-39.

Luo, X., Wu, C., Rosenberg, D., Barnes, D., 2009. "Supplier selection in agile supply chains: an information processing model and an illustration." Journal of Purchasing and Supply Management 15: 249-262.

Luthra, S., K. Govindan, D. Kannan, S. K. Mangla, and C. P. Garg. 2017. "An integrated framework for sustainable supplier selection and evaluation in supply chains." Journal of Cleaner Production 140: 1686-1698.

Luthra, S., S. K. Mangla, R. Shankar, C. P. Garg, and S. Jakhar. 2018. "Modelling critical success factors for sustainability initiatives in supply chains in Indian context using Grey-DEMATEL." Production Planning & Control 29 (9): 705-728.

Mamdani, E. H. 1974. Application of fuzzy algorithms for control of simple dynamic plant. Paper presented at the Proceedings of the institution of electrical engineers.

Mamdani, E. H., and S. Assilian. 1999. "An experiment in linguistic synthesis with a fuzzy logic controller." International Journal of Human-Computer Studies 51 (2): 135-47.

Morgan, A., 2015. "The True Cost." https://www.imdb.com/title/tt3162938/

Nematollahi, M., S. M. Hosseini-Motlagh, and J. Heydari. 2017. "Coordination of social responsibility and order quantity in a two-echelon supply chain: A collaborative decision-making perspective." International Journal of Production Economics 184: 107-121.

Nurjanni, K. P., M. S. Carvalho, and L. Costa. 2017. "Green supply chain design: A mathematical modeling approach based on a multi-objective optimization model." International Journal of Production Economics 183: 421-432.

Oelze, N. 2017. "Sustainable Supply Chain Management Implementation-Enablers and Barriers in the Textile Industry." Sustainability 9 (8): 1435.

Pedraza-Acosta, I., A. Pilkington, and D. Barnes. 2016. "Production systems and supplier selection: a multi-phase process model." Production Planning & Control 27 (9): 717-726.

Pierre, C., P. Francesco, and N. Theo. 2019. "Towards low carbon global supply chains: A multi-trade analysis of CO2 emission reductions in container shipping." International Journal of Production Economics 208: 17-28.

Polikar, R. 2006. "Ensemble based systems in decision making." IEEE Circuits and systems magazine 6 (3): 21-45.

Quinlan, J. R. 1996. "Bagging, boosting, and C4.S." Thirteenth National Conference on Artificial Intelligence & Eighth Innovative Applications of Artificial Intelligence Conference. DBLP.

Rabbani M., N. Foroozesh, S. M. Mousavi, and H. Farrokhi-Asl. 2019. "Sustainable supplier selection by a new decision model based on interval-valued fuzzy sets and possibilistic statistical reference point systems under uncertainty." International Journal of Systems Science: Operations & Logistics 6(2): 162-178.

Rabbani, M., S. A. A. Hosseini-Mokhallesun, A. H. Ordibazar, and H. Farrokhi-Asl, 2018. "A hybrid robust possibilistic approach for a sustainable supply chain location-allocation network design." International Journal of Systems Science: Operations & Logistics 1-16.

Rokach, L. 2010. "Ensemble-based classifiers." Artificial Intelligence Review 33 (1-2): 1-39.

Samuel, A. 1959. "Some Studies in Machine Learning Using the Game of Checkers." IBM Journal, 3 (3): 535-554.

Sarkar, S. and B. C. Giri. 2018. "Stochastic supply chain model with imperfect production and controllable defective rate." International Journal of Systems Science: Operations & Logistics 1-14.

Sarkis, J., and D. G. Dhavale. 2015. "Supplier selection for sustainable operations: A triple bottom line approach using a Bayesian framework." International Journal of Production Economics 166: 177-191.

Sayyadi, R., and A. Awasthi. 2018a. "A simulation-based optimisation approach for identifying key determinants for sustainable transportation planning." International Journal of Systems Science: Operations & Logistics 5 (2): 161-174.

Sayyadi, R., and A. Awasthi. 2018b. "An integrated approach based on system dynamics and ANP for evaluating sustainable transportation policies." International Journal of Systems Science: Operations & Logistics 1-10.

Sgarbossa, F., and I. Russo. 2017. "A proactive model in sustainable food supply chain: Insight from a case study." International Journal of Production Economics 183: 596-606.

Shabanpour, H., S. Yousefi, and R. F. Saen. 2017. "Future planning for benchmarking and ranking sustainable suppliers using goal programming and robust double frontiers DEA." Transportation Research Part D-Transport and Environment 50: 129-143.

Shafiq, A., P. F. Johnson, R. D. Klassen, and A. Awaysheh. 2017. "Exploring the implications of supply risk on sustainability performance." International Journal of Operations & Production Management 37 (10): 1386-407.

Shah, N. H., U. Chaudhari, and L. E. Cárdenas-Barrón. 2018. "Integrating credit and replenishment policies for deteriorating items under quadratic demand in a three echelon supply chain." International Journal of Systems Science: Operations & Logistics 1-12.

Shaw, K., R. Shankar, S. S. Yadav, and L. S. Thakur. 2012. "Supplier selection using fuzzy AHP and fuzzy multi-objective linear programming for developing low carbon supply chain." Expert Systems with Applications 39 (9): 8182-8192.

Soleimani, H., K. Govindan, H. Saghafi, and H. Jafari. 2017. "Fuzzy multi-objective sustainable and green closed-loop supply chain network design." Computers & Industrial Engineering 109: 191-203.

Tsai, C. F., Y. C. Lin, D. C. Yen, and Y.-M. Chen. 2011. "Predicting stock returns by classifier ensembles." Applied Soft Computing 11 (2): 2452-2459.

Tsao, Y. C. 2015. "Design of a carbon-efficient supply-chain network under trade credits." International Journal of Systems Science: Operations & Logistics 2 (3): 177-186.

Tseng, M. L., K. J. Wu, J. Hu, and C. H. Wang. 2018. "Decision-making model for sustainable supply chain finance under uncertainties." International Journal of Production Economics 205: 30-36.

Vahidi, F., S.A. Torabi, and M.J. Ramezankhani. 2018. "Sustainable supplier selection and order allocation under operational and disruption risks." Journal of Cleaner Production 174: 1351-1365.

Weber, C. A., J. R. Current, and W. C. Benton. 1991. “Vendor Selection Criteria and Methods”, European Journal of Operational Research 50: 2-18.

Wolpert, D. H. 1992. "Stacked generalization." Neural Networks 5 (2): 241-259.

World Commission on Environment and Development. 1987 *“Our Common Future”, Oxford University Press, Oxford.*

Wozniak, M., M. Grana, and E. Corchado. 2014. "A survey of multiple classifier systems as hybrid systems." Information Fusion 16: 3-17.

Wu, C., D. Barnes. D. Rosenberg, and X. Luo. 2009. "An analytic network process-mixed integer multi-objective programming model for partner selection in agile supply chains." Production Planning & Control 20 (3), 254-275.

Wu, C., and D. Barnes. 2010. "Formulating partner selection criteria for agile supply chains: A Dempster-Shafer belief acceptability optimisation approach." International Journal of Production Economics 125 (2): 284-293.

Wu, C., and D. Barnes. 2011. "A literature review of decision-making models and approaches for partner selection in agile supply chains." Journal of Purchasing and Supply Management 17 (4): 256-274.

Wu, C., and D. Barnes. 2012 "A dynamic feedback model for partner selection in agile supply chains." International Journal of Operations & Production Management 32 (1): 79-103.

Wu, C., and D. Barnes. 2014. "Partner selection in agile supply chains: a fuzzy intelligent approach." Production Planning & Control 25 (10): 821-839.

Wu, C., and D. Barnes. 2016. "Partner selection in green supply chains using PSO - a practical approach." Production Planning & Control 27 (13): 1041-1061.

Wu, C., and D. Barnes. 2018. "Design of ASCs including the trade-off between number of partners and reliability." International Journal of Advanced Manufacturing Technology, 97(9-12): 3683-3700.

Wu, X., V. Kumar, J. R. Quinlan, J. Ghosh, Q. Yang, H. Motoda, and G. J. McLachlan. 2008. "Top 10 algorithms in data mining." Knowledge and Information Systems 14 (1): 1-37.

Yin, S., Nishi, T. and G. Zhang. 2016. "A game theoretic model for coordination of single manufacturer and multiple suppliers with quality variations under uncertain demands." International Journal of Systems Science: Operations & Logistics 3(2): 79-91.

Yu, W., R. Chavez, and M. Feng. 2017. "Green supply management and performance: a resource-based view." Production Planning & Control 28 (6-8): 659-670.

Yu, F., Y. Yang, and D. Chang. 2018. "Carbon footprint based green supplier selection under dynamic environment." Journal of Cleaner Production 170: 880-889.

Zadeh, L. 1983. "A computational approach to fuzzy quantifiers in natural languages." Computers & Mathematics with Applications 9 (1): 149-184.

Zhao, D., C. Huang, Y. Wei, F. Yu, M. Wang, and H. Chen. 2017. "An Effective Computational Model for Bankruptcy Prediction Using Kernel Extreme Learning Machine Approach." Computational Economics 49 (2): 325-341.

Zhou, F., X. Wang, M. K. Lim, Y. He, and L. Li. 2018. "Sustainable recycling partner selection using fuzzy DEMATEL-AEW-FVIKOR: A case study in small-and-medium enterprises (SMEs)." Journal of Cleaner Production 196: 489-504.

Zimmer, K., M. Froehling and F. Schultmann, 2016. "Sustainable supplier management - a review of models supporting sustainable supplier selection, monitoring and development." International Journal of Production Research 54 (5): 1412-1442.

Figures

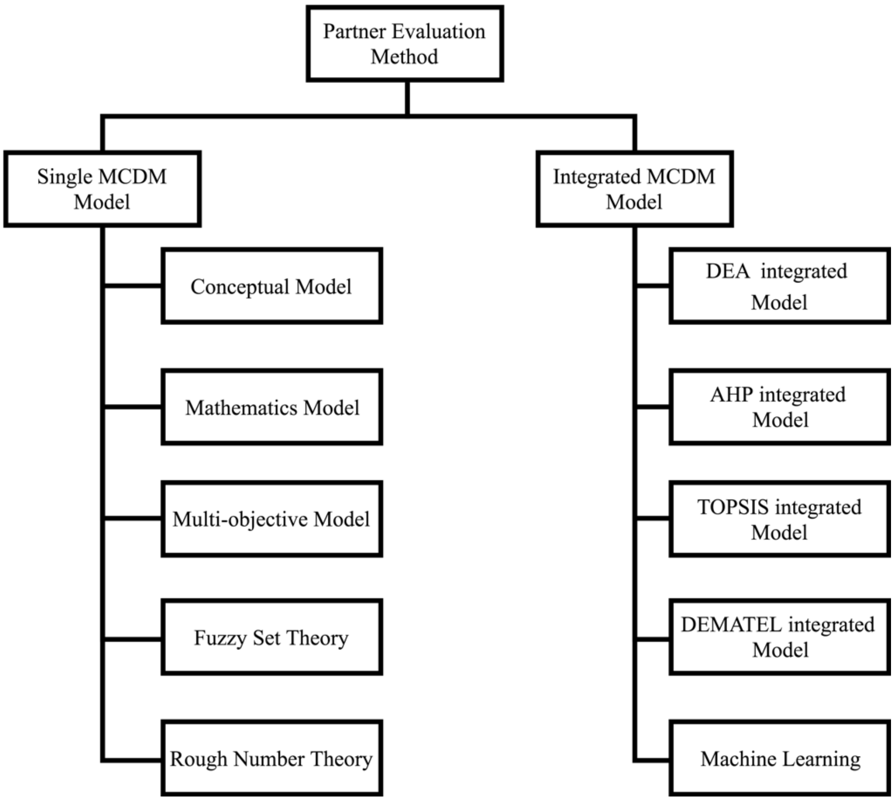


Figure 1: A basic classification of partner selection models/methods (based on Kannan et al. 2013)

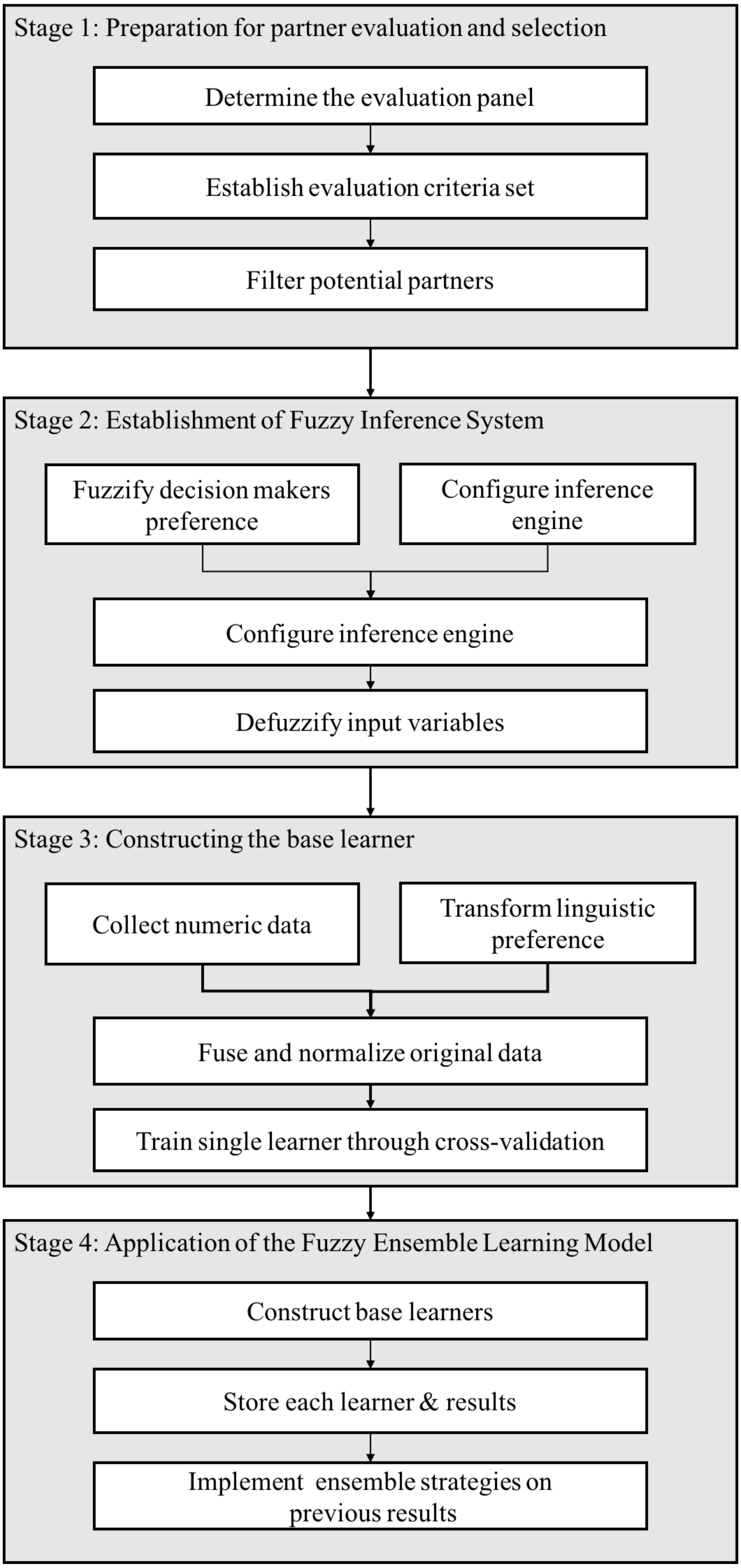


Figure 2: Flowchart of the proposed model

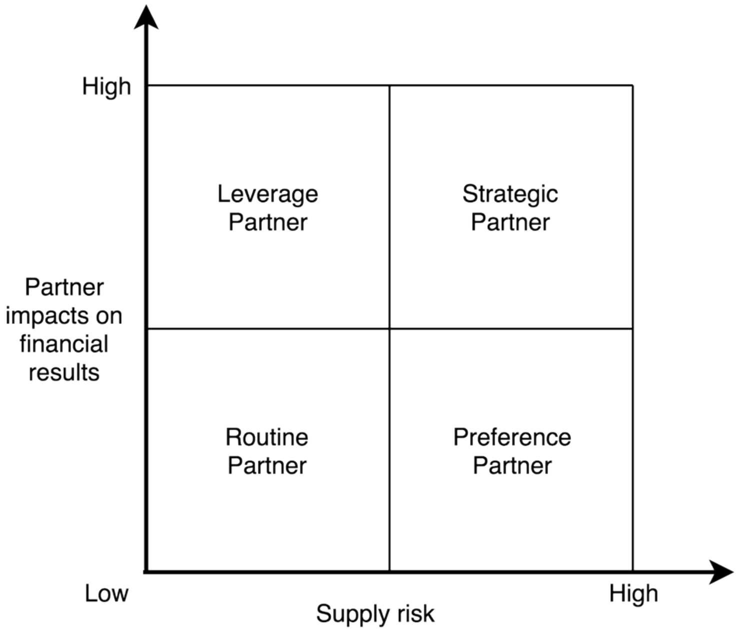


Figure 3: Category matrix for partner management (Kraljic 1983)

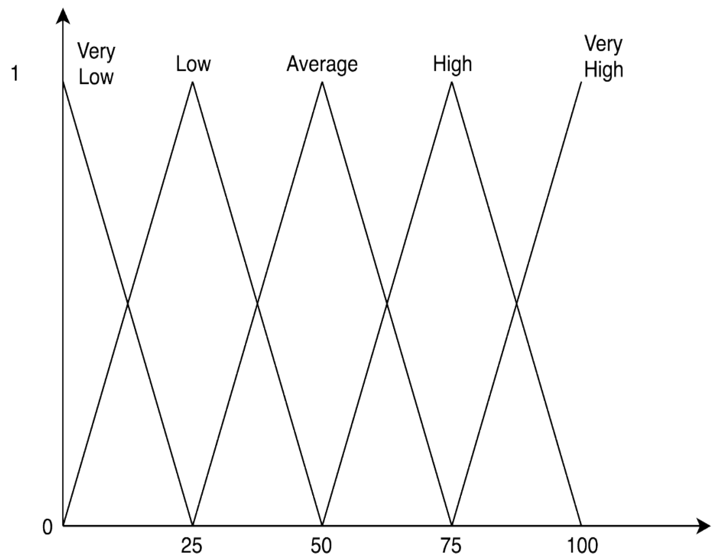


Figure 4: The triangular membership function

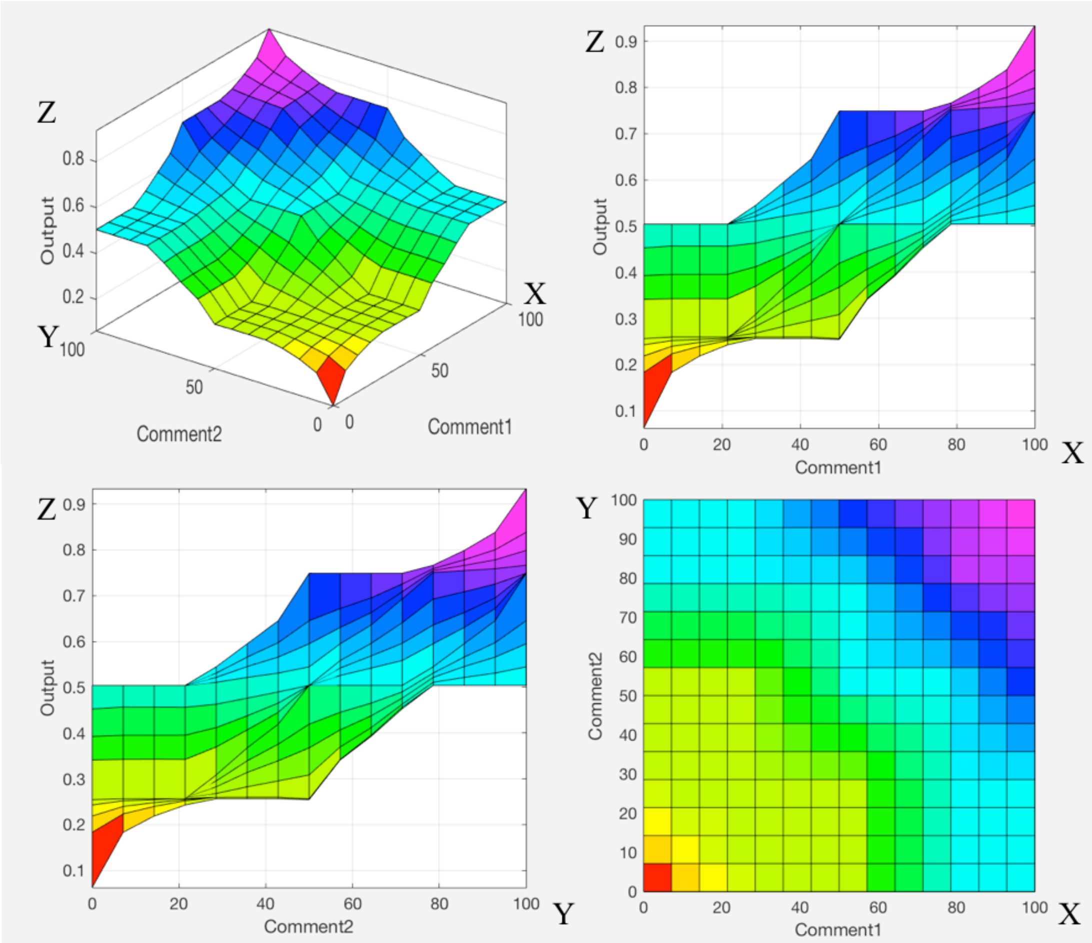


Figure 5: Fuzzy reasoning surface

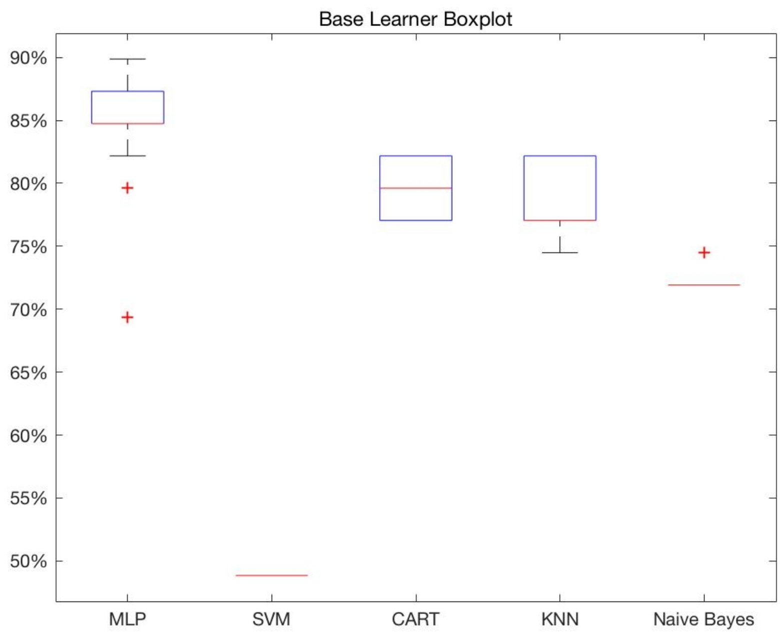
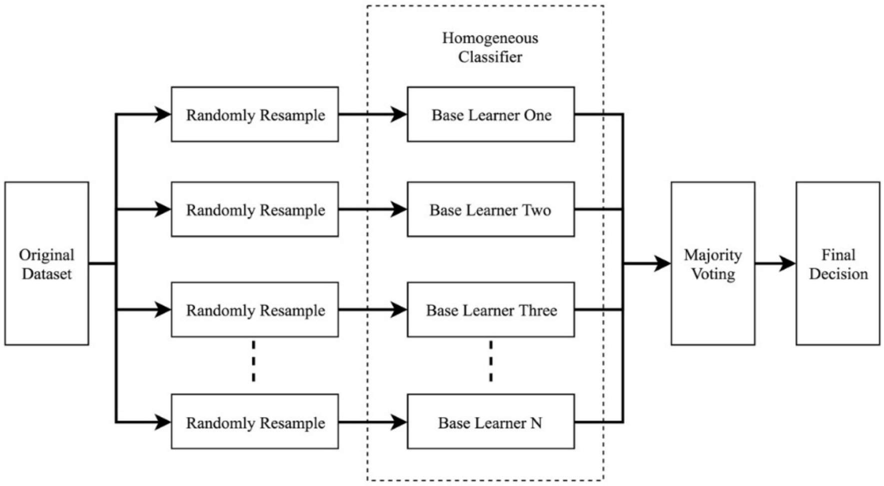
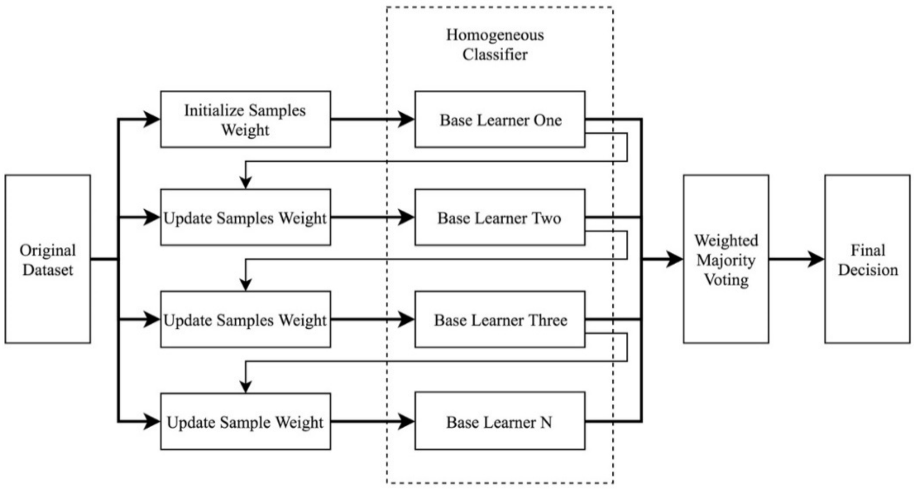


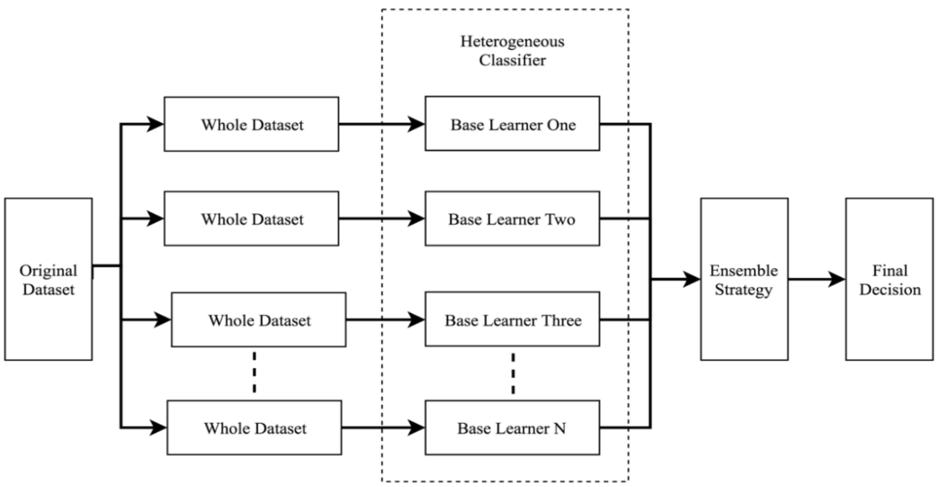
Figure 7: Boxplot of base learners' performance



(a) Procedures of Bagging algorithms

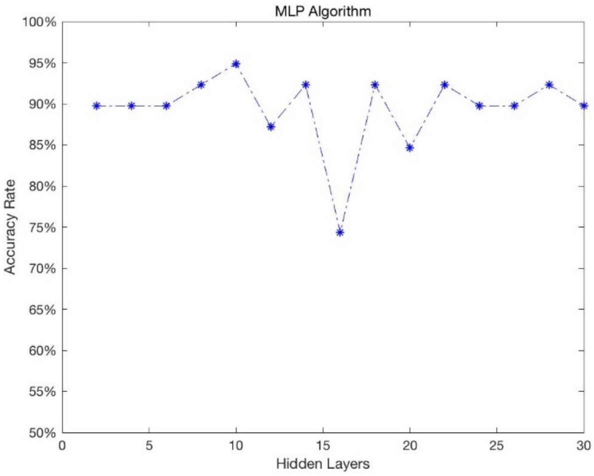


(b) Procedures of AdaBoosting algorithms

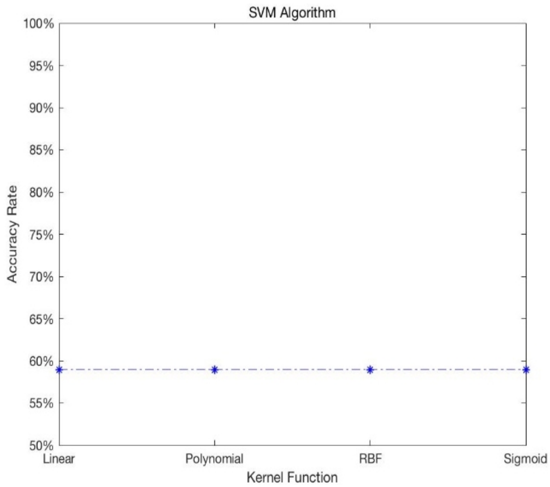


(c) Procedures of Stacking algorithms

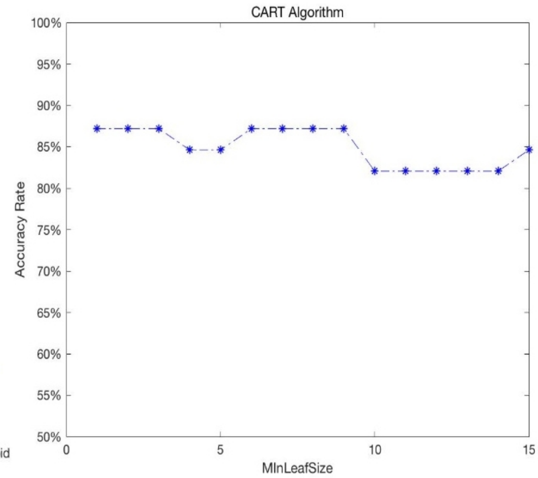
Figure 6: Procedures of ensemble learning algorithms



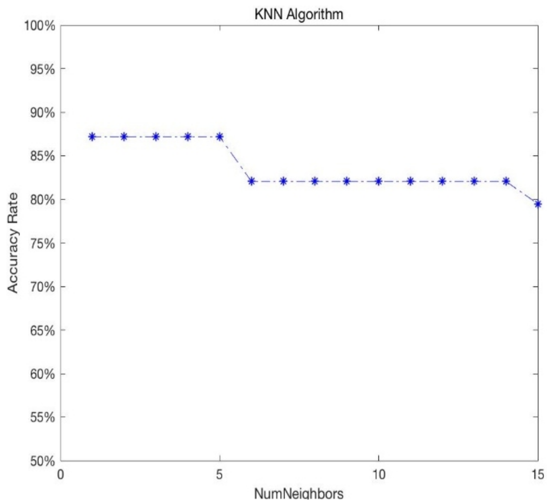
(a) Multi-layer Perceptron



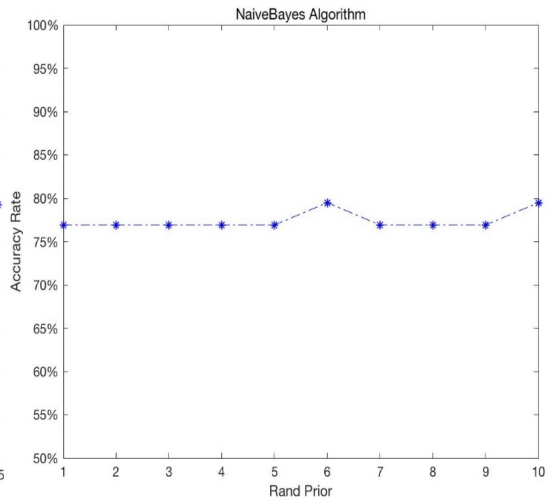
(b) SVM



(c) Classification & Regression Tree



(d) KNN



(e) Naïve Bayes

Figure 8: Comparison of base learners with respect to various model settings

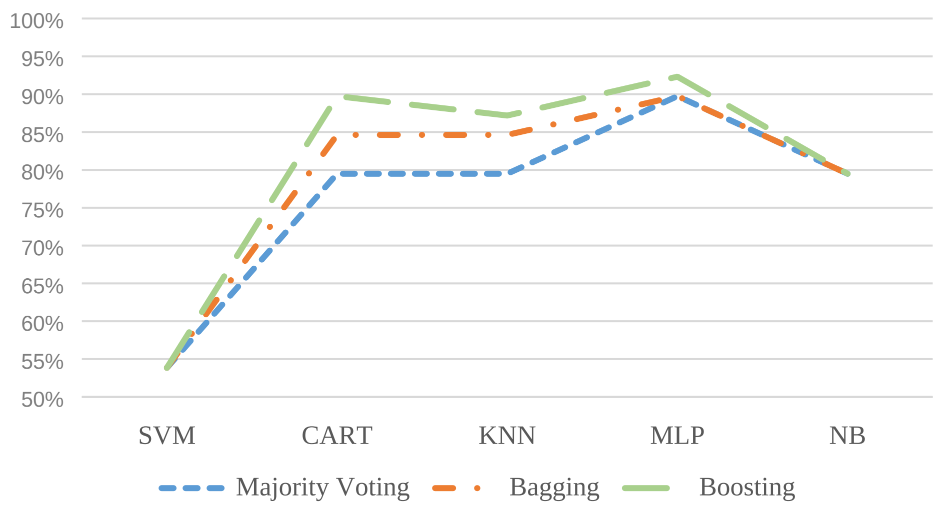


Figure 9: The accuracy of base learners with different strategies

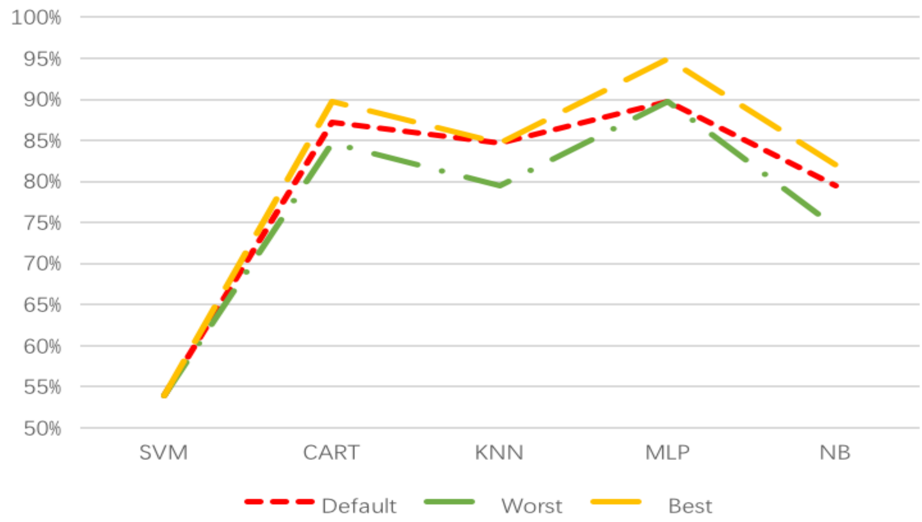


Figure 10: The accuracy of base learners under different scenarios in Bagging strategy

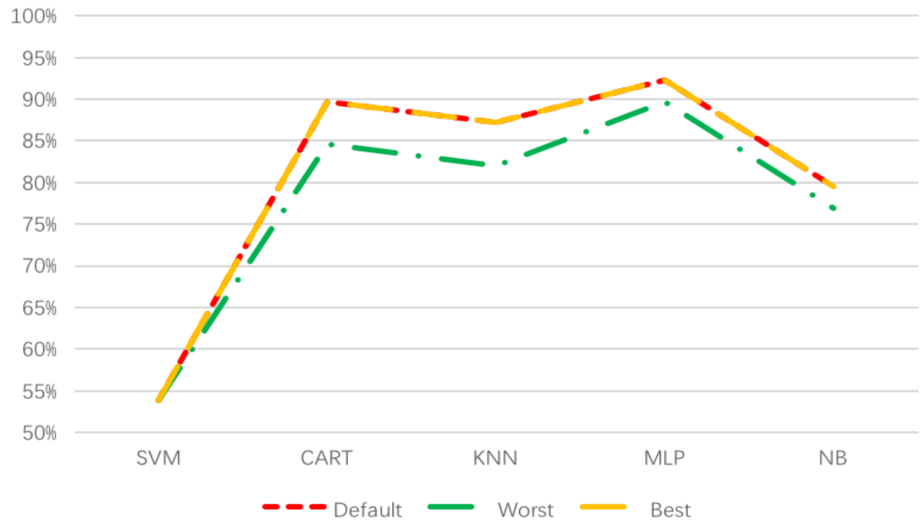


Figure 11: The accuracy of base learners under different scenarios in Boosting strategy

Tables

Table 1. The comparison of existing methods with the proposed one in partner selection

Publication	Method	Types of criteria	Category of criteria	Premise	Industry	Features
Kuo et al. (2011)	DEA + ANN	Quantitative	Economic & Environmental	Precise & certain	Production	Design ANN-DEA hybrid model in consideration of traditional economic criteria and extended environmental indicators, and then test the capacity of discrimination and noise insensitivity.
Govindan et al. (2013)	TOPSIS + FST	Quantitative & Qualitative	Economic, Environmental & Social	Vague & uncertain	No specific	Identify fuzzy multi-criteria decision-making model based on triple bottom line, apply FST to modify experts' preference and use TOPSIS for final ranking.
Chen et al. (2017)	Conceptual model	Quantitative	Economic	Precise & certain	Retail	Give insight to the corporation value creation in supply chain by examine supplier-retailer model where each component maximizes its profit while promise mutual commitments.
Karaer et al. (2017)	Mathematics model	Quantitative	Economic	Precise & certain	Chemical	Study the effect of wholesale price premium or buyer-supplier cost sharing on the improvement of partners' environmental performance.
Nurjanni et al. (2017)	Multiple obj. programming	Quantitative	Economic & Environmental	Precise & certain	Production	Construct a closed-loop network to achieve a multi-objective optimization considering the environmental and financial issues.
Ahmadi et al. (2017)	AHP + FST	Quantitative & Qualitative	Economic & Environmental	Vague & uncertain	Telecom	Employ FST to handle with the vagueness and uncertainty in decision making, and AHP for accessing the indicator weights and apply grey relational analysis for partners ranking.
Halteh et al. (2018)	Machine learning	Quantitative	Economic	Precise & certain	Finance	Propose three type decision tree algorithms to predict corporation financial stress and forecast the probability of bankrupt from an 18-criteria set.
Rabbani et al. (2019)	FST + GDM	Qualitative	Economic, Environmental & Social	Vague & uncertain	Manufacture	Introduce interval-valued fuzzy group decision-making model to process the linguistic preference from decision-makers, then evaluate the sustainability of potential suppliers
Proposed model	ELT & FST	Quantitative & Qualitative	Economic, Environmental & Social	Vague & uncertain	Electrical	Design customized FST to remedy the vagueness and inherent subjective in human reasoning progress, extend typical quantities indicators to both qualitative and quantitative ones.

Table 2: Linguistic variables and fuzzy number for evaluation indicator

Linguistic variable	Triangular Fuzzy numbers
Very Low	(0, 0, 25)
Low	(0, 25, 50)
Average	(25, 50, 75)
High	(50, 75, 100)
Very High	(75, 100, 100)

Table 3: Rules for fuzzy inference system (based on Wu and Barnes 2014)

Input rules						Output	Cases
3	Very High					Very High	1
3	High					High	1
3	Average					Average	1
3	Low					Low	1
3	Very Low					Very Low	1
2	Very High	1	High			Very High	3
2	Very High	1	Average			Very High	3
2	Very High	1	Low			High	3
2	Very High	1	Very Low			Average	3
2	High	1	Very High			Very High	3
2	High	1	Average			High	3
2	High	1	Low			Average	3
2	High	1	Very Low			Average	3
2	Average	1	Very High			High	3
2	Average	1	High			Average	3
2	Average	1	Low			Low	3
2	Average	1	Very Low			Low	3
2	Low	1	Very High			Average	3
2	Low	1	High			Average	3
2	Low	1	Average			Low	3
2	Low	1	Very Low			Very Low	3
2	Very Low	1	Very High			Low	3
2	Very Low	1	High			Low	3
2	Very Low	1	Average			Very Low	3
2	Very Low	1	Low			Very Low	3
1	Very High	1	High	1	Average	High	6
1	Very High	1	High	1	Low	High	6
1	Very High	1	High	1	Very Low	Average	6
1	Very High	1	Average	1	Low	Average	6
1	Very High	1	Average	1	Very Low	Average	6
1	Very High	1	Low	1	Very Low	Average	6
1	High	1	Average	1	Low	Average	6
1	High	1	Average	1	Very Low	Low	6
1	High	1	Low	1	Very Low	Low	6
1	Average	1	Low	1	Very Low	Low	6

Note: The total number of rules is 125.

Table 4: Economic criteria set

Economic attributes	Economic attributes	Source
Liquidity Ratio	Operating profit	Liang et al. (2018); Geng et al (2015); Luo et al. (2009) Bai and Sarkis (2010) Wu and Barnes (2010)
Operating cash flow per share	Net profit	
Gross profit margin	Net cash flow from operations	
Return on equity	Net cash flow from investing activities	
Total net assert interest rate	Net cash flow from financing activities	
Total asset turnover rate (times)	Net increase in cash and cash equivalents	
Assets and liabilities (%)	Total assets	
Equity multiplier	Total liabilities	
Current assets/total assets (%)	Total shareholders' equity	
Current liabilities/liabilities total	Current assets	
Total operating costs	Current liabilities	

Table 5: Social criteria set

Social attribute	Source
Job opportunities	Bai and Sarkis (2010); Yu et al. (2017)
Salary	Bai and Sarkis (2010); Ghadimi et al. (2016)
Health & Safety	Bai and Sarkis (2010); Azadnia et al. (2015)
Economic Welfare & Development	Bai and Sarkis (2010); Govindan et al. (2013)
Stockholder right	Rabbani et al. (2018) ; Awasthi and Omrani (2019)

Table 6: Environmental criteria set

Environmental attribute	Source
Green design	Kuo et al. (2010); Luthra et al. (2018)
ISO Standards	Bai and Sarkis (2010); Azadnia et al. (2015)
Energy consumption	Luthra et al. (2017); Kafa et al. (2015)
Waste	Govindan et al. (2013); Galo et al. (2018)
Polluting prevention	Gharaei et al. (2019c); Kazemi et al. (2018)

Table 7: Original data of potential partners

Company Name	Liquidity Ratio	Operating cash flow per share (RMB)	Gross profit margin	Return on equity	Total net assert interest rate
SHF	1.127	0.041	9.339	0.304	0.154
SKJ	1.078	0.489	6.398	9.740	4.053
SFZ	5.760	-0.056	11.924	2.228	1.772
DSD	1.115	-2.035	8.801	23.474	5.700
STM	1.104	1.371	20.849	5.718	3.145
….	…	…	…	…	…

Table 8: Decision-makers' preference on social indicators.

Company Name	Job opportunity			Salary		
	E1	E2	E3	E1	E2	E3
SHF	Low	Very Low	Low	Very High	High	Very High
SKJ	Low	Low	Low	Very High	Very High	Very High
SFZ	Average	High	Very High	Very Low	Low	Average
DSD	Very Low	Low	Low	High	Very High	Very High
STM	Very Low	Low	Low	Average	Very High	Very High
….	…	…	…	…	…	…

Table 9: Decision makes' preference on environmental indicators

Company Name	Green Design			ISO		
	E1	E2	E3	E1	E2	E3
SHF	Average	Low	Average	Average	Low	Average
SKJ	High	High	High	Very High	Very High	Very High
SFZ	Very Low	Low	Average	Average	High	Very High
DSD	Low	Average	Average	Low	Average	Average
STM	Low	High	High	Average	Very High	Very High
….	…	…	…	…	…	…

Table 10. Normalized evaluation data (partial)

Company Name	Liquidity Ratio	Operating cash flow per share	Gross profit margin	Return on equity	Total net assert interest rate
SHF	0.0231	0.3219	0.2830	0.5546	0.4850
SKJ	0.0222	0.3892	0.2492	0.6072	0.5248
SFZ	0.1155	0.3073	0.3127	0.5653	0.5015
DSD	0.0229	0.0100	0.2768	0.6837	0.5416
STM	0.0227	0.5217	0.4154	0.5848	0.5155
…	…	…	…	…	…

Table 11: Defuzzification of qualitative indicators (social & environmental criteria)

Company Name	Job opportunities	Salary	Green Design	ISO
SHF	0.062	0.254	0.934	0.934
SKJ	0.254	0.504	0.934	0.934
SFZ	0.749	0.504	0.504	0.749
DSD	0.062	0.504	0.934	0.934
STM	0.062	0.254	0.934	0.934
….	…	…	…	…

Table 12: The settings of base learners

Base learner	Parameters
SVM	Kernel function: RBF, other parameters set as Matlab default
CART	MinLeafSize:1, other parameters set as Matlab default
KNN	NumNeighbors:2, other parameters set as Matlab default
MLP	Default setting in initial LIBSVM toolbox
Naïve Bayes	Default setting in initial Matlab toolbox

Table 13: Confusion metrics for the model accuracy

Actual\Predicted	Strategic Supplier	Competitive Supplier	Influential Supplier	Common Supplier
Strategic partner	A_{11}	A_{12}	A_{13}	A_{14}
Competitive partner	A_{21}	A_{22}	A_{23}	A_{24}
Influential partner	A_{31}	A_{32}	A_{33}	A_{34}
Common partner	A_{41}	A_{42}	A_{43}	A_{44}

Table 14: The Accuracy of each base learners (Unit: %)

Name	SVM	CART	KNN	MLP	Naïve Bayes
Minimum	53.85	82.05	79.49	76.92	76.92
Maximum	53.85	87.18	87.18	94.87	79.49
Average	53.85	84.97	83.59	89.23	77.44
St. D.	0.00	2.35	2.69	5.13	1.08

Table 15: The settings of ensemble machine learning models

Combination Strategy	Parameters
Majority Voting	10-fold cross-validation
Bagging	Bagging MLP/SVMCART/KNN/NB; Number of bootstraps: 15
Boosting	Boosting MLP/SVMCART/KNN/NB; Number of iterations: 20
Stacking	Base learner: MLP/SVMCART/KNN/NB; Meta learner: CART

Table 16: The comparisons of accuracy with respect to different ensemble strategies (Unit: %)

Classifier	Majority voting	Bagging	Boosting	Stacking
SVM	53.85	53.85	53.85	
CART	79.49	84.62	89.74	
KNN	79.49	84.62	87.18	
MLP	89.74	89.74	92.31	
NB	79.49	79.49	79.49	
				89.74

Table 17: Key parameters for base learners

Base learner	Key parameter	Values
SVM	Kernel function	Line, Polynomial, RBF, Sigmoid
CART	Minimum leaf size	(1:1:15)
KNN	Number of nearest neighbours	(1:1:15)
MLP	Number of hidden layers	(2:2:30)
NB	Prior probability	15 Groups rand prior

Table 18: The comparisons of accuracy in different scenarios (Unit: %)

Classifier	Bagging	Boosting	Stacking
SVM	53.85/53.85	53.85/53.85	
CART	84.62/89.74	84.62/89.74	
KNN	79.49/84.62	82.05/87.18	
MLP	89.74/94.87	89.74/92.31	
NB	74.36/82.05	76.92/79.49	
			87.18/92.31

Note: accuracy under worst setting/accuracy under best setting