

Online-Purchasing Behavior Forecasting with a Firefly Algorithm-based SVM Model Considering Shopping Cart Use

Ling Tang¹, Anying Wang², Zhenjing Xu², Jian Li^{3*}

¹ School of Economics and Management, Beihang University, Beijing, CHINA

² Beijing University of Chemical Technology, Beijing, CHINA

³ Research Base of Beijing Modern Manufacturing Development, College of Economics and Management, Beijing University of Technology, Beijing, CHINA

Received 31 May 2017 • Revised 9 September 2017 • Accepted 25 September 2017

ABSTRACT

Due to the complexity of the e-commerce system, a hybrid model for online-purchasing behavior forecasting is developed to predict whether or not a customer makes a purchase during the next visit to the online store based on the previous behaviors, i.e., online-purchasing behavior. The proposed model makes contributions to literature from two perspectives: (1) a classification model is proposed based on the "hybrid modeling" concept, in which an effective artificial intelligence (AI) technique of support vector machine (SVM) is employed for classification forecasting and further extended by introducing the promising AI optimization tool of firefly algorithm (FA), to solve the crucial but tough task of parameters selection, i.e., the FA-based SVM model; (2) an appropriate predictor set is carefully designed especially considering online shopping cart use which was otherwise neglected in existing models, apart from other common online behaviors, e.g., clickstream behavior, previous purchase behavior and customer heterogeneity. To verify the superiority of the proposed model, an online furniture store is focused on as study sample, and the empirical results statistically support that the proposed FA-based SVM model considering online shopping cart use significantly beat all benchmarking models (with other popular classification methods and/or different predictor sets) in terms of prediction accuracy.

Keywords: e-commerce, online purchase behavior, support vector machine, firefly algorithm, shopping cart use

INTRODUCTION

With the development of e-commerce, the study on online-purchasing behavior has become an increasingly hot issue within the research fields of marketing, economic management, data mining and forecasting. Due to the unique economical virtue, online virtual store has been selected as one promising sales channel with a much lower cost (Bakos, 1997). In such competitive e-commerce market, enhancing online-purchasing conversion rates (defined as the probability of visits resulting in purchases) may be the main aim of online marketing research. For example, diverse recommendation mechanisms were designed (Lee and Kwon, 2008). Moreover, compared with physical sales, online commerce can provide much rich information about customer behaviors for further analyses such as clickstream data (Bucklin et al., 2002), customer heterogeneity (Moe and Fader, 2004), customer review (Salehan and Kim, 2016), etc. Under such a background, online-purchasing behavior analysis and prediction have attracted an increasingly wide interest from both theoretical and practical perspectives, for understanding online customer behavior through the rich information, thus enhancing online-purchasing conversion rates. Therefore, this study especially focuses on forecasting online-purchasing behavior whether or not a customer will make a purchase during the next visit to the website, in order to significantly enhance the prediction accuracy and further helpfully understand online customer behavior.

© **Authors.** Terms and conditions of Creative Commons Attribution 4.0 International (CC BY 4.0) apply.

✉ tangling_00@126.com ✉ 852131518@qq.com ✉ xuzhenjing2013@163.com

✉ lijian@amss.ac.cn (*Correspondence)

Contribution of this paper to the literature

- Using the model predictors (covering both common online customer behaviors and shopping cart uses) and the FA-based SVM model (based on the “hybrid modeling” concept), the proposed model can be used as a powerful tool for forecasting online-purchasing behavior, in terms of prediction accuracy, robustness and time saving.
- The prediction results can be used to helpfully facilitate design effective recommendation mechanisms, in order to enhance the conversion rate, which is another important issue in online marketing research.

Though still insufficient, an increasing number of studies have made great contributions to online-purchasing behavior prediction, in which two main factors are included in the existing models, i.e., prediction techniques and model variables (or predictors). In terms of prediction techniques, the existing studies on online-purchasing behavior prediction are quite insufficient compared with other fields of prediction researches, and the traditional statistical models may be the most popular forecasting tools, e.g., linear regression (LR), logit regression (LogR) and Markov chain models. The LogR model has been generally considered as the most typical traditional technique in e-commerce market prediction. For example, Van den Poel and Buckinx (2005) used the LogR to predict the online-purchasing behavior in an online wine store. Padmanabhan et al. (2001) employed several methods, e.g., the LR and LogR, to forecast whether a current visit results in a purchase or not based on clickstream data. Montgomery et al. (2004) applied the hidden Markov model to predict the purchase conversion rate in a popular online bookstore using clickstream data.

However, due to the complexity of online market, diverse powerful artificial intelligence (AI) models, with the powerful self-learning capabilities and machine learning, have been recently introduced and achieved much more satisfactory prediction results for online-purchasing behavior. For example, Gupta et al. (2014) employed sound machine learning algorithms to predict purchase by online customers based on dynamic pricing of a product. Boroujerdi et al. (2014) applied different (AI) classification algorithms, such as decision tree (DT), support vector machine (SVM) and rule-based method, to predict customer's purchase decision. Padmanabhan et al. (2001) employed the DT model to predict online-purchasing behavior based on clickstream data. Weng et al. (2011) employed Bayesian network to explore an online recommendation system for food online sales. Moe and Fader (2004) proposed a novel dynamic model based on Bayesian network to predict the online-purchasing conversion rate in the online book store of Amazon based on previous visiting and purchasing data. The empirical studies have observed that the AI tools were significantly powerful than traditional statistical ones in terms of higher prediction accuracy.

Among AI models, the SVM as an emerging AI technique has been proved to be one of promising tools for various classification problems (Wong and Hsu, 2006; Martens et al., 2007; Lessmann, 2009). The SVM model was proposed by Vapnik (1995), finely coupling solid theoretical foundation of statistical theory and powerful computer learning with the principle of structural risk minimization. According to existing studies, the SVM has been widely applied to various difficult prediction tasks and fully proved to possess excellent prediction capability, even for complex data samples (Tang et al., 2014; Tang et al., 2015). For example, Yu et al. (2006) proposed a novel method for crude oil price forecasting based on SVM. Yu et al. (2010) developed a four-stage SVM based multi-agent ensemble learning approach for credit risk evaluation. Chen et al. (2012) developed a novel approach called the hierarchical multiple kernel SVM for customer churn prediction directly using longitudinal behavioral data and Martin-Barragan et al. (2014) presented a new method, called interpretable SVMs for functional data, that provided an interpretable classifier with high predictive power. Basti et al. (2015) used SVMs to analyze initial public offerings' short-term performance. Sun et al. (2009) conducted a comparative study on the effectiveness of strategies in the context of imbalanced text classification using SVM classifier. Therefore, this study especially employs the promising technique of SVM to study online-purchasing behavior.

Though effective in classification, the SVM model has its own weakness—parameters sensitivity, and its performance is closely dependent on the parameter selection (Kim and Sohn, 2010). To address the essential but tough task of parameter selection, various optimizing methods have been introduced into SVM to formulate hybrid SVM variants based on the helpful concept of “hybrid modeling” (Yu et al., 2008; Tang et al., 2012). In particular, various AI optimization algorithms, e.g., genetic algorithm (GA), simulated annealing (SA) and particle swarm optimization (PSO), have been shown effective in addressing such backward of SVM (Lin et al., 2008). Apart from those parameter searching algorithms, another effective artificial optimization method, firefly algorithm (FA) recently proposed by Yang (2010) has recently become a promising technique in optimization programming and parameter selection. The FA is a modern heuristic algorithm based on an interesting idea that the fireflies with less light intensity (corresponding to fitness function) will be attracted by the firefly with greater ones. Existing researches have shown the superiority of FA over other AI optimization algorithms in term of optimal solution convergence (Kazem et al., 2013; Mandal et al., 2013). For example, Tang et al. (2015) introduced FA into least square

support vector regression (LSSVR) for parameter selection to predict hydropower market, and the empirical results confirmed the superiority of FA over other AI optimization tools of GA, SA and PSO. However, to the best of our knowledge, there is few study using FA as the parameters optimization algorithm of SVM. Therefore, based on the “hybrid modeling” concept, this paper especially incorporates FA into SVM to formulate a novel FA-based SVM model for online-purchasing behavior prediction.

Regarding the other important part of online-purchasing behavior prediction models—predictors, various information concerning online customers and their behaviors has been employed, such as clickstream behavior, historical purchase behavior and customer demographics (Van den Poel and Buckinx, 2005). As for clickstream measures, the information at the level of sessions, the total number of past visits to the store, the period elapsed since the last visit, the total period spent at the site during the entire period of observation, etc. have been usually considered as effective indicators for evaluating purchase potential (Moe and Fader, 2004; Van den Poel and Buckinx, 2005). Besides, the total number of viewed page, the average time spent per page, the percentage of pages viewed at category and product levels have been also taken into consideration as clickstream behaviors (Moe and Fader, 2004). As for historical purchase behavior, the popular predictors are the frequency of past purchases (Van den Poel and Buckinx, 2005; Lemon et al., 2002), the total expenditure amount (Van den Poel and Buckinx, 2005; Baensens et al., 2002), and the average expenditure amount (Van den Poel and Buckinx, 2005). Customer demographics (and heterogeneity), which help identify the online customers as buyers or non-buyers according to the motivation of entering the online stores (Moe and Fader, 2004; Ansari et al., 2000; Liu et al., 2016), mainly refer to customers’ gender, age, income and education level (Van den Poel and Buckinx, 2005; Padmanabhan et al., 2001).

However, another important online consumer behavior, i.e., online shopping cart use, may be also helpful for online-purchasing behavior prediction, which was otherwise neglected in existing studies. Inspired by on-ground shopping cart, most online shops have provided virtual shopping carts to customers as convenient tools to collect the items of interest. Moreover, the online shopping cart can also offer an ongoing search function (Bloch et al., 1986). Besides the convenience for customers, the online shopping cart can also provide helpful information for understanding online customer behaviors (Wu and Perng, 2016). For example, Close and Kukar-Kinney (2010) analyzed the virtue of virtual shopping carts and argued providing online shopping cart use can lead to a greater possibility of online purchasing. However, in the existing forecasting models for online-purchasing behavior, such an interesting behavior of online shopping cart use was usually neglected. Therefore, this study especially fills in such a literature gap by considering the behavior of online shopping cart use as an important predictor for online-purchasing behavior.

Generally speaking, this paper aims to improve online-purchasing behavior forecasting from the following two perspectives. First, a hybrid classification model, i.e., the FA-based SVM, is formulated based on the “hybrid modeling” concept, by coupling the AI prediction technique of SVM and the emerging AI optimization tool of FA for selecting parameters of SVM. Second, the customer behavior of online shopping cart use is especially considered as one important predictor, which was otherwise neglected in existing models for online-purchasing behavior prediction. To verify the effectiveness of the proposed model, an online furniture store is selected as study sample, and a series of benchmarking models with other popular classification methods and/or diverse predictor sets are formulated for comparison purpose.

The main motivation of this paper is to propose a novel model, i.e., the hybrid FA-based SVM considering online shopping cart use, for online-purchasing behavior prediction, and to verify its superiority by comparing with other benchmarking models (with other popular classification techniques and/or various sets of predictive variables). The rest of this paper is structured as follows. Section 2 describes the proposed model in detail, the experimental study is designed in Section 3, Section 4 reports and discusses the corresponding results, and Section 5 concludes the paper and outlines the main directions of future research.

MODEL FORMULATION

This section presents the overall formulation of the proposed classification model for online-purchasing behavior prediction. Subsection 2.1 gives the overall framework of the proposed model. The detailed model designs, including the predictors (or features) and the classification model, are described in Subsections 2.2 and 2.3, respectively.

Overall Framework

According to existing forecasting models for online-purchasing behavior, two key parts are included, i.e., model variables (or predictors) and classification mechanism. Regarding model predictors, a rich of information concerning online customer behaviors should be carefully analyzed. Besides some common features (e.g., clickstream behavior, previous purchase behavior and customer heterogeneity), behavior of online shopping car use which was rarely mentioned in existing prediction models is especially considered in this study. As for

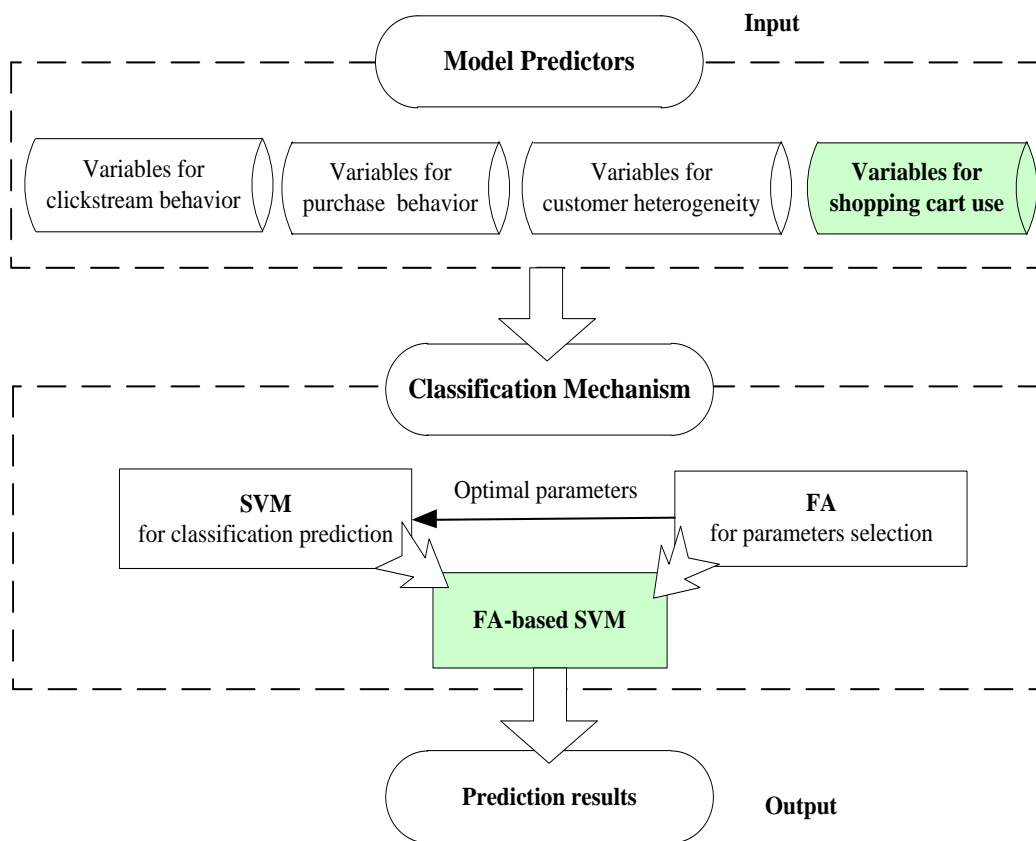


Figure 1. The framework of the proposed model for online-purchasing behavior prediction

classification mechanism, a hybrid model coupling SVM and FA is formulated to predict the online-purchasing behavior whether or not a customer will make a purchase during the next visit to the online store based on the previous behaviors. Accordingly, the framework of the novel model can be constructed, as shown in **Figure 1**.

Generally speaking, the novel model contributes to literature from two main perspectives. First, based on the “hybrid modeling” concept, a novel hybrid classification model, i.e., FA-based SVM, is proposed. In particular, the effective AI technique of SVM is employed for classification, while the promising AI optimization tool of FA is introduced to address the crucial but tough task of parameters selection in SVM. To the best of our knowledge, the studies on online-purchasing behavior prediction by using SVM are quite insufficient, not to mention the hybrid SVMs with the emerging AI optimization technique of FA. Second, the novel model especially considers the customer behavior of online shopping cart use as an important predictive feature. Since the online shopping cart use was usually neglected in existing online-purchasing behavior prediction models, this novel model helpfully fills in such a literature gap by employing this interesting and useful information.

Actually, the proposed model is a classification model, identifying whether the customer i will make a purchase during the next visit ($\hat{y}_i = 1$) or not ($\hat{y}_i = 0$) based on the historical online behaviors $x_i = \{x_{i,l}\}$, ($l = 1, \dots, d$), where x_i is the input vector with d predictive features (or predictors), and $\hat{y}_i = \{0,1\}$ is the prediction results.

The two key parts of the novel model, i.e., model predictors and classification mechanism, are respectively depicted in Subsections 2.2 and 2.3.

Model Predictors

To effectively predict whether or not a purchase will be made during the next visit, the useful information covering various online customer behaviors may be taken into consideration. According to existing studies, the online customer behaviors can be generally summarized into three main categories: clickstream measures, historical purchase behavior and customer demographics (or heterogeneity) (Moe and Fader, 2004; Van den Poel and Buckinx, 2005; Ansari et al., 2000). Besides these popular features, the online shopping cart use behavior is especially considered in the proposed model to enhance the prediction accuracy. Therefore, four categories of model predictors are included for online-purchasing behavior prediction.

Table 1. Predictive features (or predictors) in the proposed model

Category	Predictors	Description	References
Clickstream measures	<i>FrequencyVisit</i>	The total number of past visits to the store	Moe and Fader (2004), Van den Poel and Buckinx (2005), Iwanaga (2016)
	<i>RecencyVisit</i>	The time (e.g., the number of days) elapsed since last visit	Van den Poel and Buckinx (2005), Iwanaga (2016)
	<i>PriorVisit</i>	The number of visits without purchasing since the last purchase	Moe and Fader (2004)
Purchase behavior	<i>TotPurchases</i>	The total number of purchases	Lemon et al. (2002), Moe and Fader (2004), Van den Poel and Buckinx (2005)
Customer heterogeneity	<i>CustomerType</i>	Identifying the registrant as a visitor or a customer	Janiszewski (2016), Moe et al. (2002), Moe and Fader (2004), Ansari et al. (2000)
Shopping Cart use	<i>ShoppingCartPuts</i>	The number of items placed in the online shopping cart	Close and Kukar-Kinney (2010)

To design specific predictors, two basic principles are followed. First, to capture online-purchasing behavior, all available information covering the above four online customer behaviors, clickstream measures, previous purchasing behavior, customer heterogeneity and the shopping cart uses, are considered. Second, a correlation analysis is conducted to ensure the selected predictors in no relationship with each other. Finally, a total of six effective predictors covering the four types of online customer behaviors can be selected, as listed in [Table 1](#).

Clickstream measures

Based on session data, rich information about clickstream behavior can be obtained, in which a session represents a visit to the website. First, the total number of past visits to the store (labelled as *FrequencyVisit*), has been repeatedly shown be positively related to the possibility of online purchase. For example, Moe and Fader (2004b) argued that the accumulated visits can be used as an effective indicator for estimating potential for purchasing. Moe and Fader (2004b) showed that a higher frequency of visits leads to a higher conversion rate. Second, the period (e.g., in terms of days or months) elapsed since the last visit (*RecencyVisit*) has been also regarded as one of the most important features in online-purchasing behavior studies. For example, Van den Poel and Buckinx (2005) argued that it often has a positive impact on the purchasing possibility during the next visit. Moreover, visits without purchasing can accumulate a strong effect on purchasing (Weng et al., 2011). Accordingly, the number of visits without purchasing since the last purchase (*PriorVisit*) is also introduced into the model. Therefore, a total of three important model variables for clickstream behavior are considered, *FrequencyVisit*, *RecencyVisit* and *PriorVisit*.

Purchase behavior

Similar to offline world, the relationship between previous online-purchasing history and future online-purchasing behavior has been repeatedly observed. For example, Moe and Fader (2004) and Wu and Chen (2000) argued that the historical purchasing behavior can strongly influence the future purchasing behavior in e-commerce. Lemon et al. (2002) argued that the frequency of past purchases is positively correlated to the possibility of future purchasing. Van den Poel and Buckinx (2005) proved that the number of past purchases is one of the most important variables for online-purchasing behavior prediction. Accordingly, the frequency of past purchases, i.e., the total number of purchases (*TotPurchases*), is included in the proposed model.

Customer heterogeneity

Besides behavioral data, customer heterogeneity can also provide important information for forecasting online-purchasing behavior, since different types of customers with different characteristics may have different behavior rules and preferences. Therefore, customer classification becomes another interesting issue for online customer behavior study. For example, Janiszewski (2016) divided online customers into two main categories: exploratory searching group and directed searching group. The former visit the online store for only collecting information with a low possibility of purchasing, while the latter for purchasing with a much higher possibility of purchasing. Moe and Fader (2004) proposed a novel dynamic approach where two categories of online customers are individually modeled, i.e., hard-core never-buyers and common customers. Accordingly, the novel model similarly divides the customers into two main type (*CustomerType*)—visitors having not made a purchase yet (like exploratory searching group or hard-core never-buyers) and buyers having already made one or more purchases (directed searching group or common customers).

Shopping Cart use

Since most online stores have offered the useful tool of virtual shopping cart to assist customers to collect items of interest, such helpful information on online shopping cart use may provide an interesting and helpful perspective for online-purchasing behavior research. For example, Close and Kukar-Kinney (2010) argued that the online behavior of shopping cart use can significantly enhance the possibility of purchasing. Moreover, different from traditional on-ground shopping carts (e.g., grocery carts), the data for the use behavior of electronic carts if offered in the online stores can be fortunately collected for analyses. However, to the best of our knowledge, online shopping cart use behavior was somehow neglected in existing models. Therefore, this study especially fills in such a literature gap by considering the online shopping use behavior as useful information. In particular, the number of items placed in the shopping-cart (*ShoppingCartPuts*) which might have a positive relationship with purchasing possibility is used as an important model variable in the proposed model.

Classification Techniques

This section formulates a hybrid classification mechanism for online-purchasing behavior prediction, by incorporating FA into SVM for parameter selection. First, brief introductions into SVM and FA are given in Subsections 2.3.1 and 2.3.2, respectively. Second, the hybrid FA-based SVM algorithm can be formulated, as presented in subsection 2.3.3.

Support vector machine (SVM)

The SVM, an effective AI technique, was proposed by Vapnik (1995) based on the principle of structural risk minimization. Due to both powerful intelligent learning capability and solid statistical theoretical foundation, the SVM has been repeatedly shown to possess excellent prediction performance, even for complex problems. The generic idea of SVM is to first map the original data into a high-dimension feature space based on nonlinear mapping function and further to make regression by maximizing the margin hyperplane. This study implements the SVM as the classification technique for the online-purchasing behavior—whether or not a purchase will be made during the next visit to the website.

Given the training data with n observations, i.e., $\{(x_1, y_1), \dots, (x_n, y_n)\}$, where x_i ($i=1, \dots, n$) is the input and y_i ($i=1, \dots, n$) is the output, the SVM classification can be described as follows:

$$\begin{aligned} \min J(a, b, \xi) &= \left(\frac{1}{2}\right) a^T a + \gamma \sum_{i=1}^n \xi_i \\ \text{s.t. } y_i[\varphi(x_i) \cdot a + b] &\geq 1 - \xi_i, \\ \xi_i &\geq 0, i = 1, 2, \dots, n, \end{aligned} \quad (1)$$

where $a = \{a_1, \dots, a_n\}$ is the hyperplane vector, b is the bias, and $\varphi(x_i)$ is the nonlinear mapping function for transforming the input data into the high-dimension space. $\xi = \{\xi_1, \xi_2, \dots, \xi_n\}$ is the tolerable misclassification errors, and γ is the regularization parameter for the trade-off between the maximal margin and the tolerable misclassification errors.

In practice, kernel functions can be employed as the nonlinear mapping functions, helpfully simplifying the mapping process. Any asymmetric kernel function satisfying Mercer's condition can be introduced. The most popular kernel functions are Gaussian (RBF) kernel $K(x_i, x_j) = \exp(-\|x_i - x_j\|/2\sigma^2)$ with parameter σ^2 and polynomial kernel $K(x_i, x_j) = (a_1 x_i x_j + a_2)^d$ with an order of d and constants a_1 and a_2 . In this study, the RBF kernel is employed.

Accordingly, in the SVM model, two parameters, i.e., the penalty parameter γ and the RBF kernel function parameter σ^2 , need to be carefully determined to ensure the effectiveness of the model. The regularization parameter γ determines the trade-off cost between classification errors and model complexity, while the parameter σ^2 in RBF kernel function defines the non-linear mapping from the input space to high-dimension feature space. In this study, a promising AI optimization tool of FA is especially introduced in this study to optimize the two parameters in the SVM.

Firefly algorithm (FA)

Though effective in classification prediction, SVM has its own intrinsic weakness, i.e., parameter sensitiveness. Thus, this study introduces an emerging AI optimization tool, the FA, to select the optimal parameters in SVM (Tang et al., 2015). The FA, proposed by Yang (2010), is a modern heuristic algorithm. In the FA, each firefly represents a potential solution to the problem, and its attractiveness to the potential prey and absolute brightness are determined by fitness function (e.g., the classification errors in this study). At each iteration, a firefly with

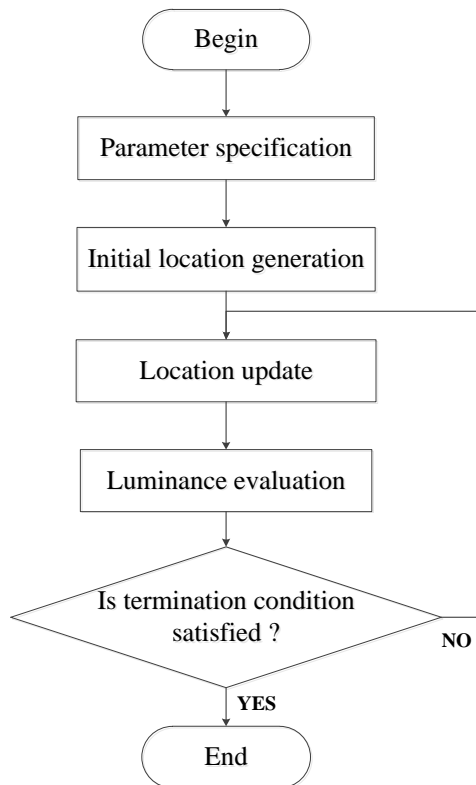


Figure 2. Searching process of FA

greater brightness attracts other fireflies with relatively less brightness (Moe and Fader, 2004b). Moreover, the brightest firefly moves randomly within a certain range. Relative brightness (or light intensity) between the firefly i and its prey j can be defined as:

$$I_{i,j} = I_i e^{-\gamma r_{i,j}^2} \tag{2}$$

where I_i represents the absolute light intensity of firefly i , i.e., $I_i = I_{i,j}$, ($i = j$) at the distance $r_{i,j}$ ($i = j$) = 0. γ is the light absorption coefficient ranging between 1.0 and 10.0. $r_{i,j}$ denotes the distance between the two fireflies i and j , which can be evaluated via Cartesian distance:

$$r_{i,j} = \|\vec{x}_i - \vec{x}_j\| = \sqrt{\sum_{l=1}^d (x_{i,l} - x_{j,l})^2} \tag{3}$$

where $\vec{x}_i = \{x_{i,l}\}$, ($l = 1, \dots, d$) represents the position of firefly i in a d -dimension future space.

Similarly, the attractiveness $\beta_{i,j}$ of firefly i to its prey j is proportional to its light intensity:

$$\beta_{i,j} = \beta_0 e^{-\gamma r_{i,j}^2} \tag{4}$$

where β_0 represents the maximum attractiveness, usually set to 1.

At iteration k , firefly j moves toward to its counterpart i with a higher brightness, and updates its position into:

$$\vec{x}_j(k+1) = \vec{x}_j(k) + \beta_{i,j}(k)(\vec{x}_i(k) - \vec{x}_j(k)) + \alpha(k)\vec{\epsilon}_j \tag{5}$$

where $\vec{\epsilon}_j$ is a random term following Gaussian distribution or uniform distribution, and $\alpha(k)$ is the randomization parameter ranging between 0 and 1 and decreases gradually:

$$\alpha(k+1) = \text{delta} \times \alpha(k) \tag{6}$$

where delta is a predetermined parameter.

Generally, five main steps are included in the FA, as shown in Figure 2.

- (a) Parameter specification: The user-defined parameters in the FA are specified according to existing researches, i.e., firefly population I , absorption coefficient γ , maximum attractiveness β_0 and delta .
- (b) Initial location generation: Randomly locate each firefly within the searching space, corresponding a feasible solution to the problem, and let iteration $k=0$.

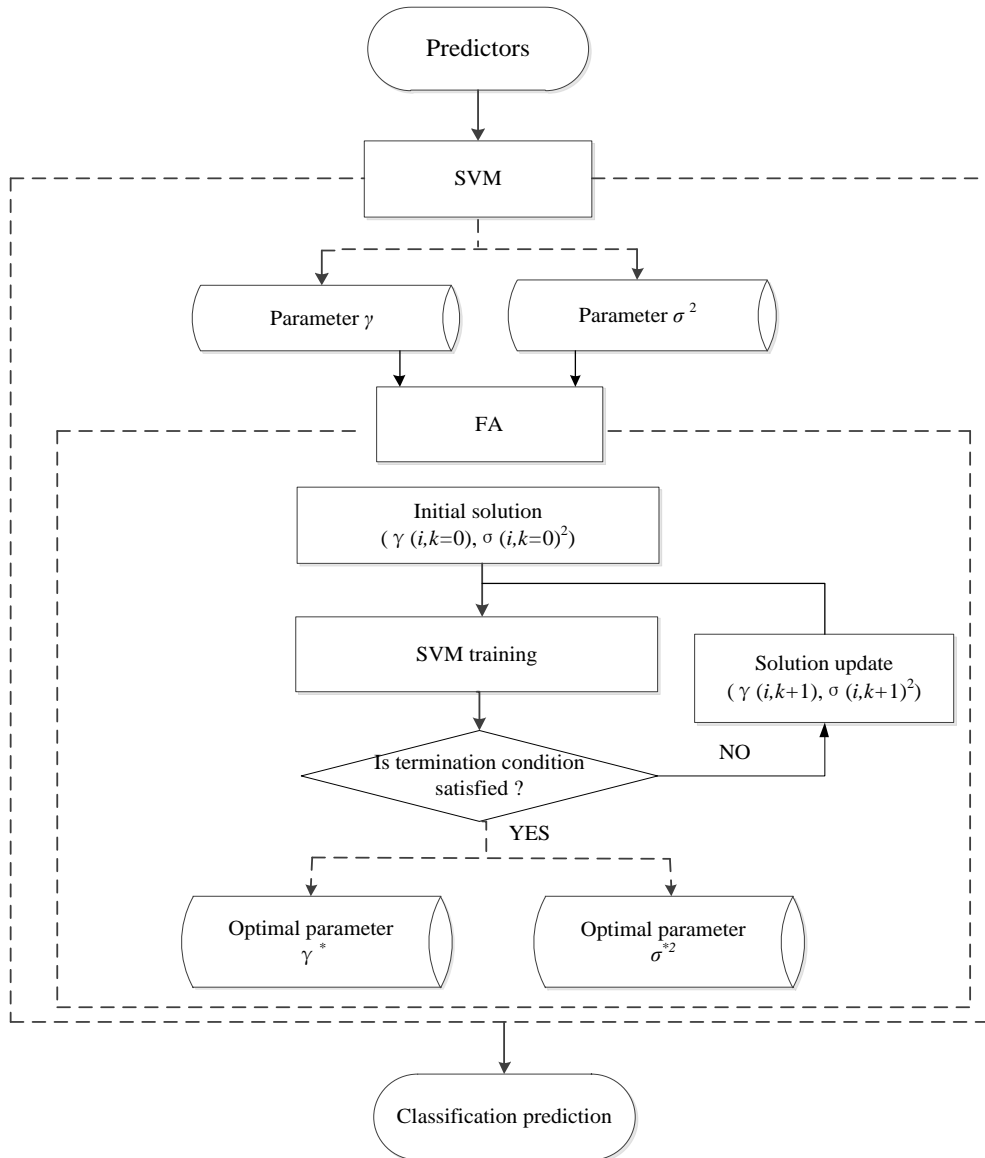


Figure 3. Main process of the hybrid FA-based SVM model

- (c) Location update: Each firefly moves toward the ones with higher brightness according to Eq. (5) or moves randomly when no brighter one can be found.
- (d) Luminance evaluation: Each firefly is evaluated in terms of fitness function, e.g., classification errors in this study.
- (e) Termination condition check: Return to Step (c) and let $k=k+1$, or stop when the termination conditions are met that iteration k reaches the maximum number of generations or the prediction error can be controlled within a tolerance level.

FA-based SVM

As mentioned above, the two parameters, the penalty parameter γ and the kernel function parameter (e.g., σ^2 in the RBF) need to be carefully predetermined to ensure the effectiveness of SVM. To solve such essential but tough task, this paper especially introduces the emerging AI optimization tool of FA into SVM and formulates a hybrid FA-based SVM model for online-purchasing behavior classification. The main steps of the proposed FA-based SVM can be described as follows, with the model framework shown in **Figure 3**.

- (a) Let iteration $k=0$, and randomly initialize a population of feasible solutions $(\gamma(i, k = 0), \sigma(i, k = 0)^2)$, $(i=1, 2, \dots, I)$, in terms of the position of fireflies in the FA, where I is the population.

Table 2. Descriptives for the online customer behaviors in the furniture site (from April 1st, 2013 to September 14th, 2013)

Number of visits	3,006,524
Number of purchases	1,267,757
Number of shopping cart puts	3,630,459
Number of visitors	2,685,043
Number of buyers	1,227,609
Conversion rate (purchase/visits)	42.17%
Visits/visitor	1.12
Cart-puts/ visitor	1.35
Purchases/visitor	47.22%
Purchases/cart-put	34.92%

- (b) Train the SVM model with each pair of parameter candidates $(\gamma(i, k), \sigma(i, k)^2)$ and calculate the classification errors as the fitness function in FA.
- (c) Update the location of each firefly into $(\gamma(i, k + 1), \sigma(i, k + 1)^2)$ according to Eq. (5) and let $k=k+1$.
- (d) Return to Step (b); or otherwise go to Step (e) when meeting the termination conditions that a satisfactory solution with tolerable errors is obtained or that k reaches the maximum iteration K .
- (e) Select the optimal solution (γ^*, σ^{*2}) with the best fitness function, among final generation of fireflies $(\gamma(i, k = K), \sigma(i, k = K)^2), (i=1, 2, \dots, J)$.
- (f) Finally, the SVM model with the optimal parameters (γ^*, σ^{*2}) can be formulated, which can be further applied to online-purchasing behavior prediction.

EXPERIMENTAL DESIGN

To verify the effectiveness of the proposed model, the online-purchasing behavior in an online furniture store is studied in this paper. Section 3.1 first gives a description to the sample data. Section 3.2 presents the evaluation criterion of classification performance. Besides, some popular classification algorithms with different predictors and/or other popular classification techniques, are introduced as benchmarking models for comparison purpose, as depicted in Section 3.3.

Data Descriptions

An anonymous online store selling furniture and the related products is selected as study sample. On the homepage, the main categories of products are listed, including lounge chairs, office chairs, desks, coffee tables, table lamps, desk lamps and accessories. Moreover, the top hot items are listed on the homepage. For convenience, a search function is also provided. The detailed information about each product can be viewed when registering with username and password. In particular, the store provides virtue shopping cart to customers, for keeping a record of the products of interest. Accordingly, the data involve all detailed log files, purchase data and shopping cart use data. The collected data cover from April 1st, 2013 to September 14th, 2013 (about five months). **Table 2** summarizes the online customer behaviors at this site for the sampling period.

According to **Table 2**, some interesting findings can be roughly obtained. First, with 3,006,524 visits and 1,267,757 purchases within the sampling period, this online furniture store is quite a hot website. Second, its conversion rate (about 42.17%) is much higher than other samples (e.g., 5% in Moe and Fader, 2004), implying this online store is quite successful. The main reason can be attributed to the relatively cheap price and good quality in this online furniture store. Finally but the most importantly, one interesting implication can be obtained that, the online shopping cart use behavior may be an effective predictor for purchasing possibility, since about 34.92% of online shopping cart uses result in purchases.

Due to the large scale of the sample data, we randomly withdraw 4,312 observations (i.e., 3% of total 143,674 customers) for each run, and the selected samples are further randomly divided into two subsets with equal samples: training subset for model training and testing subset for performance evaluation (Polat and Güneş, 2007). It is worth noticing that due to the randomness stemming from model initialization and data sampling, a total of 50 experiments with different initial solutions in the FA-based SVM model and different sampling datasets are performed, and the average values are calculated as the final results.

Evaluation Criteria

To evaluate the prediction performance of the model, the most well-established criterion for classification accuracy is selected, i.e., percentage correctly classified (*PCC*) accuracy (Edwards et al., 2006):

$$PCC = \frac{\sum_{i=1}^M a_i}{M}, \quad a_i = \begin{cases} 1, & \hat{y}_i = y_i \\ 0, & \hat{y}_i \neq y_i \end{cases} \quad (7)$$

where $\hat{y}_i = \{0,1\}$ and $y_i = \{0,1\}$ are respectively the predicted value and actual value of case i . In particular, $\hat{y}_i = 1$ (or $\hat{y}_i = 0$) stands for that the customer i is predicted to make a (or no) purchase during the next visit. $y_i = \{0,1\}$ follows a similar way. M is the size of the testing dataset. Obviously, a higher *PCC* indicates a higher prediction accuracy of the tested model.

Furthermore, to statistically test the difference across different models in classification accuracy, one-tailed t test is performed, with the null hypothesis that the *PCC* value of the target model is no more than that of its benchmarking model. Accordingly, the t statistic can be defined as:

$$t = \frac{\overline{PCC}_1 - \overline{PCC}_2}{s_{12} \cdot \sqrt{\frac{2}{N}}}, \quad s_{12} = \sqrt{\frac{1}{2}(s_1^2 + s_2^2)} \quad (8)$$

where $\overline{PCC}_1 = \sum_{n=1}^N PCC_{1,n}$ and $\overline{PCC}_2 = \sum_{n=1}^N PCC_{2,n}$ are the mean *PCC* of N experiments respectively produced by the target model and the benchmarking model, s_{12} is the grand standard deviation (or pooled standard deviation) of the two results groups, s_1^2 and s_2^2 are the unbiased variances of the two result groups, and $N=50$ is the total number of experiments.

Benchmarking Models

In order to test the performance of the proposed classification model, a series of benchmarking models, with different model variables (see Subsection 3.3.1) and/or other popular classification techniques (see Subsection 3.3.2), are formulated for comparison.

Benchmarking predictor design

To testify the effectiveness of our predictor design, the existing designs in other studies are especially introduced for comparison purpose, as listed in **Tables 3** and **4**. It is worth noticing that different from our design (A0), all the existing designs (A1-3) somehow neglected the important information of shopping cart use. Furthermore, they otherwise employed some other predictors, which we go without due to the strong relationship with similar features. For example, the designs A1 and A2 considered a much rich of model variable; however, many features fall into one category and might have strong relationships with each other, while some importing behaviors are otherwise neglected (e.g., customer heterogeneity and especially shopping cart use). In contrast, even with a relatively small size of predictor set, our design (A0) covers all the four main categories of online customer behaviors.

Table 3. Benchmarking designs of model predictors

Type	Predictors	Description	Mark
I-Clickstream measures	<i>FrequencyVisit</i>	The total number of past visits	V1
	<i>RecencyVisit</i>	Number of days since last visit	V2
	<i>PriorVisit</i>	Number of visits since last purchase	V3
	<i>TotPage</i>	The total number of viewed pages	V4
	<i>TotPageSearch</i>	The total number of times one made use of the search engine of the site	V5
	<i>TotPageProduct</i>	The total number of pages viewed concerning the product	V6
	<i>TotPageCat</i>	The total number of pages viewed which are at the category level	V7
	<i>MaxRep</i>	The maximum number of repeat page views per product	V8
	<i>TotUniqCat</i>	The total number of all category level pages viewed that are unique	V9
	<i>TotUniqProd</i>	The total number of all product level pages viewed that are unique	V10
II-Customer heterogeneity	<i>CustomerType</i>	Whether the registrant is a visitor or a customer.	V11
III-Purchase Behavior	<i>Totpurchases</i>	The total number of past purchases	V12
	<i>PurchaseRecency</i>	Number of days since last purchase	V13
IV-Shopping cart use	<i>ShoppingCartPuts</i>	The total number of items placed in shopping-cart	V14

Table 4. Model predictor designs in the proposed model A0 and benchmarking models A1-3

Type	I														II	III	IV	Reference
Mark	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14				
A0	√	√	√	-	-	-	-	-	-	-	√	√	-	√	This study			
A1	√	-	-	√	√	√	√	√	√	√	-	√	-	-	Moe and Fader (2004b)			
A2	√	√	-	√	-	√	√	-	-	-	-	√	√	-	Buckinx and Van (2007)			
A3	√	-	√	-	-	-	-	-	-	-	√	√	-	-	Moe and Fader (2004)			

Benchmarking classification techniques

To enhance prediction accuracy, a hybrid FA-based SVM classification mechanism is proposed in this study. For comparison purpose, some popular classification techniques and similar hybrid learning paradigms coupling SVM and other AI optimizations, are formulated as benchmarking classification models.

According to existing researches, the most popular classification techniques are DT, LogR, naive Bayesian classifier (NBC), Bayesian network (BN), back propagation (BP), generalized regression neural networks (GRNN) and SVM (Yu et al., 2010; Quinlan, 1986). Accordingly, these techniques are introduced as benchmarks for the proposed FA-based SVM model. For similar hybrid SVM variants, three popular AI optimization algorithms, i.e., PSO, SA and GA, are also introduced into SVM to formulate three AI-based SVMs (Tang et al., 2015; Huang and Dun, 2008; Lin et al., 2008).

To sum up, a total of ten benchmarking classification methods are employed for the proposed FA-based SVM (i.e., FA-SVM) model, including seven popular classification tools (DT, LR, NBC, BN, BP, GRNN and single SVM) and three hybrid SVM variants (PSO-SVM, SA-SVM and GA-SVM). For these hybrid paradigms, the first part in the abbreviations is the parameters optimization algorithms in SVM.

Table 5. Parameter specification of AI optimization algorithms

FA	Population size	20	PSO	Population size	20
	Generations	50		Generations	50
	alpha	0.90		personal learning factor	1.60
	beta	0.97		social learning factor	1.50
	gamma	1		inertia weight	[0.40, 0.90]
SA	Generations	50	GA	Population size	20
	Initial temperature	500		Generations	50
	Final temperature	0.01		Crossover rate	0.60
				Mutation rate	0.50

EXPERIMENTAL RESULTS

According to the model steps (see Section 2) and the experimental design (see Section 3), the proposed FA-based SVM model considering online shopping cart use is employed to predict the online-purchasing behavior in an anonymous furniture e-commerce website. Subsection 4.1 gives the parameter settings for the proposed model and its ten benchmarking models, and Subsection 4.2 discusses the corresponding results.

Parameter Settings

First, the seven single benchmarking models, i.e., DT, LogR, NBC, BN, BP, GRNN and single SVM, are performed to predict the online-purchasing behavior in the furniture electronic store. As for DT, iterative dichotomiser 3 (ID3) algorithm is employed (Quinlan, 1986). In BN, the K2 algorithm is introduced to learn the structure, the Bayesian estimation (BE) is used as the parameter learning method, and the junction tree algorithm is applied for inference (Cooper and Herskovits, 1992). In BP, the number of input units is set equal to the number of model variables, the number of hidden units is determined according to a classic mathematical result of Kolmogorov, i.e., $2^n + 1$, where n is the number of input units, and the number of output unit is one (Tang et al., 2015). As for single SVM, the RBF kernel function is used and the parameters γ and σ^2 are set by the trial-and-error method (Tang et al., 2014).

For hybrid SVMs, the RBF kernel function is also used, and various AI optimization methods of GA, PSO, SA and FA are individually introduced to search the optimal parameters γ and σ^2 on the range from 0.01 to 10,000.00 in the SVM. In particular, the user-defined parameters in these AI optimization algorithms are set according to the related literature and our modification (Tang et al., 2015; Lin et al., 2008), as listed in Table 5. It is worth noticing that for consistency, the common parameters in different algorithms are set the same, e.g., population size and generations.

Furthermore, randomness is inevitable in view of initial solutions and some random parameters in the AI optimization tools and data sampling (see the experimental design in Section 3.1). Therefore, each model is run fifty times, with the mean value as the final result. All above models are performed via Matlab software with the version number of 7.14 (R2012a).

Results Analysis

For clearly understanding, the experimental results are discussed from three perspectives. Subsection 4.2.1 focuses on the effectiveness of our predictor design. Subsection 4.2.2 compares the proposed model with other classification techniques. Subsection 4.2.3 summarizes the main findings in the empirical study.

Effectiveness of considering online shopping cart

To verify the effectiveness of the proposed predictor design, the existing designs are introduced. Table 6 presents the corresponding results of different FA-SVM models with different variable sets (designs A0-3), in terms of the mean PCC values of 50 experiments (*PCC*), the standard deviations (*Std.*), the maximum values (*Max.*) and the minimum values (*Min.*). Furthermore, the average computational time is calculated. To statistically test the

Table 6. Comparison results of the proposed predictor design A0 and its benchmarking designs A1-3

Design		A0	A1	A2	A3
PCC		0.7673	0.7026	0.7152	0.7606
Std.		0.0067	0.0113	0.0105	0.0065
Max.		0.7824	0.7220	0.7290	0.7745
Min.		0.7536	0.6600	0.6863	0.7462
Time Consuming (seconds)		5612.7295	7465.8408	9735.2771	679.4646
t-test result	t-statistic		27.1181	23.4118	3.9295
	p-value		(<0.0001)	(<0.0001)	(0.0002)

superiority of the proposed model, one-tailed t test is performed, with the null hypothesis that the classification accuracy (in terms of PCC) of the proposed model A0 is no higher than the benchmarking models A1-3, as the t -statistic and p -value (in bracket) listed in the last two rows of **Table 6**.

From the comparison results, one important conclusion can be obviously found that our design A0 appears to be the best model in classification prediction for online-purchasing behavior, given that its mean PCC value, together with the maximum and minimum values, are the highest without exception, with a relatively low standard deviation ranking the second. Among benchmarks, design A3, which is the most similar to A0 but without online shopping cart use, ranks the best in terms of classification accuracy. However, design A2, even with the most features, performs the worst. Furthermore, the t -test statistically confirms the superiority of our design over all the other model variable designs, under a confident level of 90%.

From the above results, two interesting implications can be obtained for online-purchasing behavior research. First, as online shopping cart becomes an increasingly important tool in electronic stores, the information on shopping cart uses cannot be neglected in online customer behavior study. Therefore, our design considering online shopping cart uses significantly defeats all the benchmarking designs neglecting such important information, in terms of prediction accuracy. Second, even with a large number of model variables, designs A1 and A2 appear to be ineffective in capturing online-purchasing behavior, mainly due to two important reasons. On the one hand, some importing behaviors (e.g., customer heterogeneity) are neglected in designs A1 and A2. On the other hand, many features in designs A1 and A2 fall into one category responding to one type of customer behavior, between which there exists a strong relationship. Therefore, such multicollinearity leads to modeling inefficiency. In contrast, even with a small feature size, models A0 and A3 covering diverse customer behaviors (clickstream measures, previous purchase behavior and customer heterogeneity) appear much more effective in in terms of both prediction accuracy and time saving.

Effectiveness of hybrid FA-SVM model

To verify the effectiveness of the proposed hybrid FA-SVM method, seven single popular classification techniques (i.e., LogR, GRNN, BP, NBC, DT, single SVM and BN) and three hybrid SVMs (i.e., GA-SVM, PSO-SVM and SA-SVM) are also performed. **Table 7** presents the comparison results of different classification methods with the design A0.

From the comparison results, four main conclusions can be easily observed. First, the proposed FA-SVM model significantly outperforms all the benchmarking models without exception, given that its mean value of PCC is the highest. Second, the four hybrid SVM models of FA-SVM, GA-SVM, SA-SVM and PSO-SVM perform much better than single SVM model in terms of much higher PCC values, confirming the effectiveness of the "hybrid modeling" concept in model improvement. Third, most AI algorithms (i.e., BP, NBC, DT, SVM and BN) defeat the most traditional statistical model LogR, indicating the effectiveness of AI techniques for modeling the complex online e-commerce system. It is worth noticing that the LogR model may be the most popularly used technique for online customer behavior research (e.g., Van den Poel and Buckinx, 2005; Padmanabhan et al., 2001). From the results of t -test, the proposed FA-based SVM classification model can be statistically proved to be superior to all the considered benchmarks listed here, at the confidence level of 90%.

Given the randomness and arbitrariness in the AI searching methods of FA, GA, SA and PSO and data sampling, the model robustness should be also taken into consideration as another important criterion for performance evaluation. In particular, all the single models (i.e., LogR, GRNN, BP, NBC, DT, SVM and BN) and hybrid SVM models (i.e., FA-SVM, GA-SVM, PSO-SVM and SA-SVM) are run fifty times (as mentioned in Section 3.1), and the standard deviations ($std.$) of PCC are calculated, as the results reported in **Table 7**. According to the results, BN is

Table 7. Comparison results of the proposed FA-SVM model and its benchmarking classification techniques

	<i>PCC</i>	<i>Std.</i>	<i>Max.</i>	<i>Min.</i>	<i>t-test result</i>	
					<i>t-statistic</i>	<i>p-value</i>
FA-SVM	0.7673	0.0068	0.7823	0.7531		
GA-SVM	0.7649	0.0066	0.7777	0.7517	1.7569	0.0420
PSO-SVM	0.7647	0.0066	0.7809	0.7452	1.7405	0.0440
SA-SVM	0.7653	0.0078	0.7837	0.748	1.5295	0.0660
LogR	0.5869	0.0470	0.7174	0.5457	27.1857	(<0.0001)
GRNN	0.5664	0.0091	0.586	0.5443	136.4710	(<0.0001)
BP	0.7482	0.0559	0.7842	0.5707	2.3948	0.0100
NBC	0.7024	0.0274	0.7689	0.6403	16.1515	(<0.0001)
DT	0.7640	0.0083	0.7758	0.7359	2.2190	0.0150
SVM	0.7592	0.0066	0.7749	0.7489	5.7269	(<0.0001)
BN	0.7505	0.0016	0.7619	0.7503	18.0624	(<0.0001)

shown to be the most stable model with the smallest standard deviation, but its prediction accuracy is at a relatively low level. Apart from BN, all the hybrid SVM variants (i.e., FA-SVM, GA-SVM, PSO-SVM and SA-SVM) are much more stable than the single models in terms of much lower standard deviations, testifying to the effectiveness of both SVM tool and the “hybrid modelling” concept.

Generally, with the highest *PCC* value and relatively low standard deviation, the proposed hybrid FA-SVM model can be considered as an effective and robust classification approach for online customer behavior prediction.

Summarizations

According to above analyses, seven conclusions can be obtained from the experimental study.

Regarding model predictors: (1) the proposed design A0 considering online shopping cart uses significantly outperforms all the benchmarking designs neglecting such an important feature, in terms of prediction accuracy; (2) focusing on the number of features, large-scaled variable sets (e.g., designs A1 and A2) still be ineffective, due to multicollinearity between the similar predictors and neglecting certain type(s) of customer behavior; and (3) in contrast, even with a small feature size, the proposed design A0 covering the main customer behaviors performs the best, in terms of both prediction accuracy and time saving.

For classification techniques: (4) the proposed hybrid FA-SVM model significantly outperforms all the benchmark models without exception, in terms of prediction accuracy; (5) the four hybrid models, i.e., FA-SVM, GA-SVM, SA-SVM and PSO-SVM, perform much better than single SVM model, confirming the effectiveness of the “hybrid modeling” concept in model improvement; and (6) the AI optimization algorithms (i.e., BP, NBC, DT, SVM and BN) significantly outperform the traditional statistical LogR model, indicating the effectiveness of AI techniques for modeling the complex online e-commerce system.

Therefore, (7) using the model predictors (covering both common online customer behaviors and shopping cart uses) and the FA-based SVM model (based on the “hybrid modeling” concept), the proposed model can be used as a powerful tool for forecasting online-purchasing behavior, in terms of prediction accuracy, robustness and time saving.

CONCLUSIONS

To improve online-purchasing behavior prediction, a hybrid FA-based SVM model considering online shopping cart use is proposed in this study. In particular, a hybrid classification model, i.e., FA-based SVM, is formulated based on the “hybrid modeling” concept, by coupling the effective AI technique of SVM for classification and the emerging AI optimization tool of FA for parameters selection in SVM. Furthermore, an appropriate set of predictors is carefully structured considering online shopping cart use, which was otherwise neglected in existing models.

For illustration and verification purposes, the proposed model is utilized to predict the online-purchasing behavior in an anonymous online furniture store. Furthermore, a series of benchmarking models with different model variable designs and other popular classification techniques are employed for comparison purpose. The empirical results statistically confirm that using the model predictors (covering both common online customer behaviors and shopping cart uses) and the FA-based SVM model (based on the “hybrid modeling” concept), the proposed model can be used as a powerful tool for forecasting online-purchasing behavior, in terms of prediction accuracy, robustness and time saving.

However, besides the sample data used here, the proposed model should be extended to other e-commerce sites to verify the generality and universality. Furthermore, the prediction results can be used to helpfully facilitate design effective recommendation mechanisms, in order to enhance the conversion rate, which is another important issue in online marketing research. We will look into these issues in the near future.

ACKNOWLEDGEMENTS

This work was supported by grants from the National Natural Science Foundation of China (NSFC Nos. 71433001, 71301006, 71571010, and 71622011), the National Program for Support of Top Notch Young Professionals, and Beijing Advanced Innovation Center for Soft Matter Science and Engineering.

REFERENCES

- Ansari, A., Essegai, S., & Kohli, R. (2000). Internet recommendation systems. *Journal of Marketing research*, 37(3), 363-375.
- Baesens, B., Viaene, S., & Van den Poel, D. (2002). Bayesian neural network learning for repeat purchase modelling in direct marketing. *European Journal of Operational Research*, 138(1), 191-211.
- Bakos, J. Y. (1997). Reducing buyer search costs: Implications for electronic marketplaces. *Management science*, 43(12), 1676-1692.
- Basti, E., Kuzey, C., & Delen, D. (2015). Analyzing initial public offerings' short-term performance using decision trees and SVMs. *Decision Support Systems*, 73, 15-27.
- Bloch, P. H., Sherrell, D. L., & Ridgway, N. M. (1986). Consumer search: An extended framework. *Journal of Consumer Research*, 13(1), 119-126.
- Boroujerdi, E. G., Mehri, S., Garmaroudi, S. S., Pezeshki, M., Mehrabadi, F. R., Malakouti, S., & Khadivi, S. (2014). A study on prediction of user's tendency toward purchases in websites based on behavior models. In *Proceedings of the 6th IEEE Conference on Information and Knowledge Technology*, 61-66.
- Buckinx, W., Verstraeten, G., & Van den Poel, D. (2007). Predicting customer loyalty using the internal transactional database. *Expert systems with applications*, 32(1), 125-134.
- Bucklin, R. E., Lattin, J. M., Ansari, A., Gupta, S., Bell, D., Coupey, E., ..., & Steckel, J. (2002). Choice and the Internet: From clickstream to research stream. *Marketing Letters*, 13(3), 245-258.
- Chen, Z. Y., Fan, Z. P., & Sun, M. (2012). A hierarchical multiple kernel support vector machine for customer churn prediction using longitudinal behavioral data. *European Journal of operational research*, 223(2), 461-472.
- Close, A. G., & Kukar-Kinney, M. (2010). Beyond buying: Motivations behind consumers' online shopping cart use. *Journal of Business Research*, 63(9), 986-992.
- Cooper, G. F., & Herskovits, E. (1992). A Bayesian method for the induction of probabilistic networks from data. *Machine learning*, 9(4), 309-347.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
- Edwards, T. C., Cutler, D. R., & Zimmermann, N. E. (2006). Effects of sample survey design on the accuracy of classification tree models in species distribution models. *Ecological Modelling*, 199(2), 132-141.
- Gupta, R., & Pathak, C. (2014). A machine learning framework for predicting purchase by online customers based on dynamic pricing. *Procedia Computer Science*, 36, 599-605.

- Huang, C. L., & Dun, J. F. (2008). A distributed PSO-SVM hybrid system with feature selection and parameter optimization. *Applied Soft Computing, 8*(4), 1381-1391.
- Iwanaga, J., Nishimura, N., Sukegawa, N., & Takano, Y. (2016). Estimating product-choice probabilities from recency and frequency of page views. *Knowledge-Based Systems, 99*, 157-167.
- Janiszewski, C. (1998). The influence of display characteristics on visual exploratory search behavior. *Journal of Consumer Research, 25*(3), 290-301.
- Kazem, A., Sharifi, E., & Hussain, F. K. (2013). Support vector regression with chaos-based firefly algorithm for stock market price forecasting. *Applied soft computing, 13*(2), 947-958.
- Kim, H. S., & Sohn, S. Y. (2010). Support vector machines for default prediction of SMEs based on technology credit. *European Journal of Operational Research, 201*(3), 838-846.
- Lee, K. C., & Kwon, S. (2008). Online shopping recommendation mechanism and its influence on consumer decisions and behaviors: A causal map approach. *Expert Systems with Applications, 35*(4), 1567-1574.
- Lemon, K. N., White, T. B., & Winer, R. S. (2002). Dynamic customer relationship management: Incorporating future considerations into the service retention decision. *Journal of marketing, 66*(1), 1-14.
- Lessmann, S., Sung, M. C., & Johnson, J. E. V. (2009). Identifying winners of competitive events: A SVM-based classification model for horserace prediction. *European Journal of Operational Research, 196*(2), 569-577.
- Lin, S. W., Lee, Z. J., & Chen, S. C. (2008). Parameter determination of support vector machine and feature selection using simulated annealing approach. *Applied Soft Computing, 8*(4), 1505-1512.
- Lin, S. W., Lee, Z. J., Chen, S. C., & Tseng, T. Y. (2008). Parameter determination of support vector machine and feature selection using simulated annealing approach. *Applied Soft Computing, 8*(4), 1505-1512.
- Lin, S. W., Ying, K. C., Chen, S. C., & Lee, Z. J. (2008). Particle swarm optimization for parameter determination and feature selection of support vector machines. *Expert systems with applications, 35*(4), 1817-1824.
- Liu, L. W., Chang, C. M., Huang, H. C., & Chang, Y. L. (2016). Verification of social network site use behavior of the university physical education students. *Eurasia Journal of Mathematics, Science & Technology Education, 12*(4), 793-805.
- Łukasik, S., & Żak, S. (2009). Firefly algorithm for continuous constrained optimization tasks. *International Conference on Computational Collective Intelligence*. Springer Berlin Heidelberg, 97-106.
- Mandal, P., Haque, A. U., Meng, J., Srivastana, A. K., & Martinez, R. (2013). A novel hybrid approach using wavelet, firefly algorithm, and fuzzy ARTMAP for day-ahead electricity price forecasting. *IEEE Transactions on Power Systems, 28*(2), 1041-1051.
- Martens, D., Baesens, B., Van Gestel, T., & Vanthienen, J. (2007). Comprehensible credit scoring models using rule extraction from support vector machines. *European journal of operational research, 183*(3), 1466-1476.
- Martin-Barragan, B., Lillo, R., & Romo, J. (2014). Interpretable support vector machines for functional data. *European Journal of Operational Research, 232*(1), 146-155.
- Moe, W. W., Chipman, H., George, E. I., & McCulloch, R. (2002). A Bayesian treed model of online purchasing behavior using in-store navigational clickstream. Revising for 2nd review at *Journal of Marketing Research*.
- Moe, W. W., & Fader, P. S. (2004). Dynamic conversion behavior at e-commerce sites. *Management Science, 50*(3), 326-335.
- Moe, W. W., & Fader, P. S. (2004b). Capturing evolving visit behavior in clickstream data. *Journal of Interactive Marketing, 18*(1), 5-19.
- Montgomery, A. L., Li, S., Srinivasan, K., Liechty, J. C. (2004). Modeling online browsing and path analysis using clickstream data. *Marketing Science, 23*(4), 579-595.
- Padmanabhan, B., Zheng, Z., & Kimbrough, S. O. (2001). Personalization from incomplete data: What you don't know can hurt. In *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, 154-163.
- Polat, K., & Güneş, S. (2007). Breast cancer diagnosis using least square support vector machine. *Digital Signal Processing, 17*(4), 694-701.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine learning, 1*(1), 81-106.
- Salehan, M., & Kim, D. J. (2016). Predicting the performance of online consumer reviews: A sentiment mining approach to big data analytics. *Decision Support Systems, 81*, 30-40.
- Sun, A., Lim, E. P., & Liu, Y. (2009). On strategies for imbalanced text classification using SVM: A comparative study. *Decision Support Systems, 48*(1), 191-201.

- Tang, L., Wang, S., & He, K. (2015). A novel mode-characteristic-based decomposition ensemble model for nuclear energy consumption forecasting. *Annals of Operations Research*, 234(1), 111-132.
- Tang, L., Wang, Z., Li, X., Yu, L., & Zhang, G. (2015). A novel hybrid FA-Based LSSVR learning paradigm for hydropower consumption forecasting. *Journal of Systems Science and Complexity*, 28(5), 1080-1101.
- Tang, L., Yu, L., & He, K. (2014). A novel data-characteristic-driven modeling methodology for nuclear energy consumption forecasting. *Applied Energy*, 128, 1-14.
- Tang, L., Yu, L., Wang, S., Li, J., & Wang, S. (2012). A novel hybrid ensemble learning paradigm for nuclear energy consumption forecasting. *Applied Energy*, 93, 432-443.
- Van den Poel, D., & Buckinx, W. (2005). Predicting online-purchasing behavior. *European Journal of Operational Research*, 166(2), 557-575.
- Weng, S. S., Liu, S. C., & Wu, T. H. (2011). Applying Bayesian network and association rule analysis for product recommendation. *International Journal of Electronic Business Management*, 9(2), 149.
- Wong, W. T., & Hsu, S. H. (2006). Application of SVM and ANN for image retrieval. *European Journal of Operational Research*, 173(3), 938-950.
- Wu, C., & Chen, H. L. (2000). Counting your customers: Compounding customer's in-store decisions, interpurchase time and repurchasing behavior. *European Journal of Operational Research*, 127(1), 109-119.
- Wu, W. C., & Perng, Y. H. (2016). Research on the Correlations among Mobile Learning Perception, Study Habits, and Continuous Learning. *Eurasia Journal of Mathematics, Science & Technology Education*, 12(6).
- Xie, W., Yu, L., & Xu, S. (2006). A new method for crude oil price forecasting based on support vector machines. *Lecture Notes in Computer Science*, 3994, 444-451.
- Yang, X. S. (2010). Firefly algorithm, stochastic test functions and design optimization. *International Journal of Bio-Inspired Computation*, 2(2), 78-84.
- Yu, L., Wang, S., & Lai, K. K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. *Energy Economics*, 30(5), 2623-2635.
- Yu, L., Yue, W., Wang, S., Lai, K. K. (2010). Support vector machine based multiagent ensemble learning for credit risk evaluation. *Expert Systems with Applications*, 37(2), 1351-1360.

<http://www.ejmste.com>