An empirical analysis of intention of use for bike-sharing system in China through machine learning techniques

Tao Zhou [a]

[a]School of Engineering, Faculty of Science Engineering and Built Environment, Deakin University, Geelong, Australia;

Kris M. Y. Law [a,b]

[a]School of Engineering, Faculty of Science Engineering and Built Environment, Deakin University, Geelong, Australia;

[b]Department of Industrial Engineering Management, University of Oulu, Oulu, Finland

K. L. Yung[c]

[c]Department of Industrial and Systems Engineering, The Hong Kong Polytechnic University, Hong Kong

Abstract

Sharing bicycles, as boosted by the advanced mobile technologies, is expected to mitigate the traffic congestion and air pollution issues in China. A survey study was conducted with 335 valid samples to identify the key factors that influence the customers' intention of use for bike-sharing system and quantify the corre- sponding importance. Five machine learning techniques for classi- fication are applied and results are compared. The best performed technique is selected to prioritise and quantify the importance level of the influencing factors. The results indicate that the perceived ease of use is the most significant factor for the intention to use sharing bikes.

Introduction

The sharing economy can be defined as a new economic mode that provides consumers the access to idle social resources through information network technology (Ma and Zhang 2019). The booming of new technologies including mobile devices, big data and cloud computing have dramatically contributed to the prosperity of sharing economy which has led people to a low-carbon and sustainable lifestyle (Hew et al. 2017; McAfee et al. 2012). The global sharing economy is estimated to dramatically increase from 14 billion USD in 2014 to approximately 335 billion USD in 2025 (Yaraghi and Ravi 2017). Due to the rapid development of Internet and mobile payment technologies, China has become one of the major markets adopting the sharing economy (Ma and Zhang 2019). In 2016, the market value of sharing economy in China was estimated to be 253 billion USD accounting for 4.6% of the GDP and it is expected to reach 10% by 2020 (Ma and Zhang 2019). The sharing economy concept has been adopted and applied across various fields, such as house sharing (Airbnb), car sharing (Uber) and bike sharing (Mobike). Among these applications, bike sharing is one of the fastest expanding fields and the most representative of sharing economy in China (Liu and Yang 2018).

Due to the urbanisation in the major cities in China, traffic congestion and air pollution have become the critical issues concerned enormously by city inhabitants (Liu and Yang 2018). Bike-sharing system, a short-term public bicycle rental scheme that provides

customers with the temporary access to bicycles at self-service stations through online commercial platforms, is considered as an alternative to carbon-emission vehicles for short-distance travel (Yu et al. 2018). Comparing to the conventional use of bicycles, bike- sharing system provides customers the temporary use of bicycles instead of the owner- ship. Supported by the advanced technologies, such as Internet of Things (IoT), mobile online payment and big data processing system, cashless payment systems including Alipay and WeChat Wallet are extensively used by customers for renting a sharing bike (Liu and Yang 2018). The popularity of smart phones and mobile payment provide technical feasibility for sharing bicycle services and the potential for large-scale consump- tion. In this context, short-term and large-scale bicycle placement in dense residential areas with heavy traffics is the key marketing strategy which aims to improve accessibility and support flexible mobility for customers (Ricci 2015). According to China Electronic Commerce Research Centre, the active users for bike-sharing systems in China was about 298 million in 2017 ("China's Sharing Economy Development Report in 2016" 2017).

Despite of the advantages, there are also challenges. Firstly, the vicious completion has led to failure of many bike-sharing companies (Si et al. 2020). In 2017, Ofo, one of the major players in bike-sharing market in China, had placed approximately 8 million bikes in 170 cities around the world and processed over 25 million orders per day (Zhang and Mi 2018). However, because of irrational business expansion, Ofo was reported with serious financial crisis (Pham 2019). Another start-up company, WuKong bike, declared bankruptcy after only 90 days from its operation (Spinney and Lin 2018). Secondly, the bikes are mainly deployed in central business districts (CBD) and the accessibility of sharing bikes is heavily influenced by the tidal effects (Liu and Yang 2018). A considerable amount of bikes is abandoned which has led to the drastic declines in users satisfaction (Wang and Szeto 2018). In this case, bike-sharing companies should strive to attract more potential users while maintaining loyal customers to guarantee their market shares (Jia, Liu, and Liu 2018). Policies have been made by local governments in China to regulate the investment in bike sharing market as well as user behaviours (Yao et al. 2019). However, only a few studies have been carried out to investigate the factors that contribute to the intention of use (IU) for current bike-sharing system. Liu and Yang (2018) studied the factors including trust, subjective norm and the gender of users affecting the behavioural intention of sharing economy by employing technology acceptance model (TAM). Du and Cheng (2018) investigated the travel patterns of residents in Nanjing, China and found that the intention of using sharing-bike system increased when they were making short distance travels. In addition, Si et al. (2020) studied the sustainable usage intention of sharing-bike users and indicated that perceived behaviour control is the most important factor.

This paper is to identify the key factors that influence the customers' IU for bike-sharing system and to determine the importance of these factors. Perceived benefits (PBs) and perceived loss (PL) are considered as two main categories of factors in this study. The importance level of factors are testified and quantified utilising machine learning (ML) techniques. Distinct to the conventional statistical analysis methods including multiple discriminant analysis (MDA) (Brusch and Rappel 2019) and structural equation modelling (SEM) (Leong et al. 2015) which oversimplify the relationship between dependent and independent variables, ML techniques are developed for exploring complex non-liner relationships of variables with promising performance (Gerlein et al. 2016). We further employ five ML techniques to determine the importance of factors and the best

performed technique is selected to prioritise the importance level of influencing factors. The framework of the methodology adopted in this study consists of five steps: research model, survey study, preliminary analysis, ML analysis and sensitivity analysis (Figure 1).

The rest of this paper is structured as following: Section 2 presents an overview of theoretical background of the study as well as the overview of ML techniques. Section 3 presents the research model which incorporates the key influencing factors. Section 4 illustrates the survey study process. In Section 5, the preliminary data analysis and ML implementation are discussed. Section 6 presents the contributions and the limitations of this study as well as the future study plan. The conclusion of this paper is presented in Section 7.

Background

Theoretical models relating to intention of use

The word 'intention' is described as the subjective possibility of people to take certain actions while IU refers to the subjective tendency of customers to adopt certain service or purchase certain products (Fishbein and Ajzen 1977). Many models of IU were derived from or developed based on theory of reasoned action (TRA) proposed by Fishbein and Ajzen (1977) and technology acceptance model (TAM) proposed by Davis (1989).

TRA emphasises that an individual's execution of an action is determined by his attitude towards the action to be performed and subjective norms (Fishbein and Ajzen 1977; Ajzen 1991). Ajzen (1991) extended the TRA model by introducing the new influence factor of 'cognitive behaviour control' into the TRA model and formed the theory of planned behaviour (TPB) model. In TPB, human behaviour was not only influenced by individual attitudes and supervisory norms, as well external environment (Ajzen 1991). Based on the path of 'belief-attitude-behaviour intention' of the TRA model, Davis (1989) divided the concept of belief into perceived usefulness (PU) and perceived ease of use (PEU), and developed the TAM. PU and PEU are considered to be two prerequisites affecting technology adoption attitudes, which directly affect the intention and actual use of technology (Tan et al. 2014). PEU is also believed to indirectly affect behaviour intention and ultimately affect its user behaviour (Tan et al. 2014). Subsequently, Venkatesh and Davis (2000) found that attitudes had weaken the predictions of behavioural intentions and actual behaviours, and therefore excluded attitudinal variables from the model. In the following years, this model has been widely acknowledged by scholars to explore the impact of various new technologies and information systems on individual and organisational behaviour (Habibi, Davidson, and Laroche 2017; Wahyuni 2016).

Kim, Chan, and Gupta (2007) improved TAM in the new environment of mobile Internet, and put forward the theory of value-based adoption model (VAM). In this model, the perceived value (PV) of users is affected by the PB and the PL, and the PV of users determines their willingness to use. Empirical results have shown that VAM can better explain and predict user behaviour than other traditional models in mobile Internet environment (Belk 2014), therefore, VAM has been extensively implemented for analysing the customer behaviour towards Internet-related services or products. Kim, Park, and Choi (2017) investigated the influencing factors for the adoption of smart home devices by using VAM while Hsu and Lin (2018) explored the IU for IoT services.

Apart from the positive factors that influence the customer behaviours, perceived risk (PR) is another major perspective that many researchers have studied into (Lee 2009). The concept of PR was first introduced into marketing from psychology by Bauer in 1960 as he believed that human behaviours had to bear with risks because of their uncertainty. This concept was further extended to two aspects by Cox in 1967, namely uncertainty and severity. Uncertainty refers to the ambiguity of the performance, quality and other attributes of the product, while severity means the loss of time, money, psychology situation after purchasing the product (Pappas 2016). Over the last decades, the definition of PR varies. However, most researchers generally agree that the PR is composed of multiple dimensions. Jacoby and Kaplan (1972) divided PR into five risk factors including capital risk, personal risk, functional risk, psychological risk and social risk for measuring PR of 12 different products which indicated that the explaining variance of these five dimensions reached 61.5%. On the basis of this study, Peter and Tarpey Sr (1975) included time risk into the model which represented the uncertain loss of time and energy that may occur when purchasing a product. Stone and Grønhaug (1993) proposed a risk model that comprised financial risk (FR), performance risk, physical risk, psychological risk, social risk and time risk and demonstrated that their model could explain over 88% of the PR. With the rapid development of Internet technology, security risk (SR) and privacy risk (PR) are generally accepted as important considerations of PR (Forsythe and Shi 2003). For example, Hsieh (2015) defined the PR for physicians as their risk of suffering from data leakages while Marafon et al. (2018) investigated the influence of perceived security and PR to the intention of using Internet banking system.

The above-mentioned models have provided fundamental theories for studying the intention of using or accepting new technologies. Based on the theory of TAM, Hwang and Lee (2013) conducted an empirical study on users' willingness to sharing their social network information while playing mobile games. The results indicated that social norms, scenarios and perceived entertainment could have significant positive effects on users' IU. Liu (2015) constructed an IU model based on TAM and TPB for taxi software and pointed out that PU (such as time saving), PEU (such as convenience of calling), compatibility (such as habitual taxi driving) and subjective norms (such as influence of others) could have positive impacts on IU, while PR (such as privacy disclosure) and perceived price level (such as traffic cost) had negative effects on IU. Kim, Park, and Choi (2017) studied the influencing factors of consumer's mobile shopping intention by the integration of TAM and VAM, and research showed that emotional experience, PU and PEU could positively affect consumers' shopping intention.

Machine learning techniques

Support vector machine

Support vector machine (SVM) is a popular supervised ML algorithm for resolving multiclass classification problems with good performance by mapping the dataset into a high dimensional space and generating a hyperplane for separating the classes (Tehrany et al. 2015). SVM has been widely applied to practical problems, such as pattern recognition (Sahbi, Audibert, and Keriven 2010), text classification (Amayri and Bouguila 2010) and financial analysis (Chaudhuri and De 2011). SVM uses kernel functions to transfer the dimension of input data which enables the algorithm to deal with non-linear classification problems (Liu et al. 2013). The radial basis function is commonly selected as the kernel of SVM and the performance of SVM mainly depends on the parameters of kernel functions, namely Kernel coefficient (γ) and regularisation parameter (C) (Thanh Noi and Kappas 2018).

Ensemble learning

Ensemble learning is a method of training multiple weak ML learners to obtain a stronger learner with better generalisation ability (Singh, Gupta, and Rai 2013). Bootstrap aggregating ('Bagging') and boosting are the two main categories of ensemble learning. Bagging method performs the parallel training of several base learners on training datasets which are uniformly sampled from the original dataset and averages the predictions of these base learners (Thanh Noi and Kappas 2018; Xie et al. 2018). In contrast, base learners in boosting are trained in a sequential manner and the weights of misclassified samples are increased iteratively (Singh, Gupta, and Rai 2013). Random forest (RF) and gradient boosting decision tree (GBDT) are the two popular algorithms for bagging and boosting respectively.

RF, introduced by Breiman (2001), builds a group of decision trees by constructing the splits based on the random feature selection from a random sample of the datasets (Brown and Mues 2012), in which two parameters need to be pre-defined: number of trees and number of features in each split (Liu et al. 2013). RF has been successfully applied for data mining and classification in various areas including land cover mapping (Colditz 2015), remote sensing (Belgiu and Drăguţ 2016) and user intention (Qu et al. 2019). Initially proposed by Friedman (2001), GBDT introduces a new decision tree to the previous model at each step to minimise the loss function. By fitting new decision trees recursively to the residual, gradient boosting algorithm improves the performance in the misclassified samples (Touzani, Granderson, and Fernandes 2018). According to Touzani, Granderson, and Fernandes (2018), three parameters should be tuned for GBDT: the number of decision trees, learning rate and the depth of decision trees. GBDT has shown a strong performance through being implemented in fields including biological analysis (Babajide Mustapha and Saeed 2016), transportation accident analysis (Ma et al. 2017) and FR prediction (Zheng 2019).

Artificial neural network

Inspired by human brain, artificial neural network (ANN) is a computing system composed by parallel layers of processing units that are used for storing and transmitting knowledge or information from the inputs (Liébana-Cabanillas, Marinković, and Kalinić 2017). These processing units are called as neurons which are connected by weighted links for passing signals or information (Deb, Lee, and Santamouris 2018). ANN learns the relationship between input and output variables through data training process in which the feedback of training results will be used for updating the weights in order to achieve a better performance. Due to its unique structure and realisation of universal approximation theorem (Liang and Srikant 2016), ANN has the ability to identify complex and non-linear relationships. The typical architecture of ANN consists of several layers of connected neurons, which are one input layer, multiple hidden layers and one output layer (Liébana-Cabanillas et al. 2018). The goal of training neural network is to find out the optimal values of weights and biases.

As one of the important types of ML, ANN has been widely applied across multiple disciplines (Abiodun et al. 2018). For example, ANN has been applied in power electronics and motor drives area (Bose 2007), ultrasonic cleaning process (Wu et al. 2009), smart antenna arrays (Rawat, Yadav, and Shrivastava 2012) and order-cycle management (Sustrova 2016). Apart from the applications in engineering, ANN has also been proposed as a data analysis approach in social science. Priyadarshinee et al. (2017) predicted the determinants for cloud computing adoption using a hybrid approach that integrates SEM and ANN, in which SEM is applied for conducting hypothesis testing and selecting significant independent variables as the inputs of ANN. Similarly, Liébana-Cabanillas et al. (2018) investigated the determinants for mobile payment acceptance and Asadi et al. (2019) analysed the influencing factors for wearable healthcare devices adoption respectively by implementing hybrid SEM-ANN method.

Research model development

In this section, a research model that integrates TAM, VAM and perceived risk model (PRM) is proposed to identify the influencing factors for IU. According to Hsu and Lin (2018), PBs and PL are two major considerations that could influence users' decision of adopting a product. Studies have shown that PU and PEU could positively affect the PBs and ultimately the IU (Liu and Yang 2018). In addition, perceived discount (PD) could positively impact the PBs (Nusair et al. 2010). Therefore, in our study, PU, PEU and PD are regarded as the potential influencing factors in the category of PB. Since perceived loss or PRs usually impede customers from using a product (Belk 2014); therefore, PR, SR and FR are contributing to the perceived loss. Figure 2 depicts the research model that incorporates the influencing factors and the explanation of these factors is detailed as below.

Perceived benefits

Perceived usefulness

PU, which is defined as the degree of influence perceived by people that the adoption of a technology or service could have on their performance, is generally considered as one of the major factors used in TAM (Mathwick, Malhotra, and Rigdon 2002). The application of bike-sharing system is believed to mitigate the traffic congestion problem and improve the travel efficiency by encouraging city dwellers to use bicycles instead of driving and taking buses for short distance travelling (Yu et al. 2018). As indicated by Lai (2015), PU had positive influence on the intentions of use for their sample users of bike-sharing system in Taiwan and his results were supported by Chen and Lu (2016) and Wang et al. (2018). Therefore, as one of the widely studied variables, PU is considered in our study.

Perceived ease of use

PEU is another original variable adopted in TAM, which is defined as the 'degree to which an individual believes that using a new technology or system would be free of efforts' (Davis 1989). In many cases of analysing the acceptance of a new technology or system, PEU and PU were considered simultaneously. The importance of PEU has also confirmed in the sharing economy environment. Yu et al. (2018) and Ma and Zhang (2019) found that PEU and PU had significant positive effects on the users' attitude and intention towards using commercial bike-sharing system. In addition, the use of sharing-bike system is largely dependent on the mobile applications and the accessibility of the bicycles and these applications could affect the PEU (Zhang, Shaheen, and Chen 2014).

Perceived discount

As indicated by Grewal et al. (1998), price discount is an effective economic incentives that could attract potential customers to increase their willingness to buy the products. Nusair et al. (2010) and Iranmanesh et al. (2017) demonstrated that discounts or promotions had significant influence on customers' perceptions on the value of the discount as well as their intention of purchasing or use. Bike-sharing system is operated based on the mobile applications, such as Mobike, where customers could interact with the operators and a great deal of information is shared and exchanged between customers and companies (Yu et al. 2018; Si et al. 2020). In order to expand their market, the major sharing bike enterprises have issued a series of policies to subsidise potential customers directly or indirectly. For example, Mobike has cooperated with Wechat to introduce holiday specials and these advertisements and promotion information quickly attracted the attention of customers (Zhang, Shaheen, and Chen 2014). Hence, in this study, PD, which is defined as the degree of influence perceived by users that discounts or promotions could have on their intention to use (Iranmanesh et al. 2017), is considered as an influencing factor.

Perceived loss

According to Featherman and Pavlou (2003), PR can be defined as the uncertainty regarding to the negative consequences of adopting a product or service. Generally, PR is comprised of various components such as SR (Arpaci 2016), PR (Wang and Lin 2017), FR (Al-Jabri 2015), time risk (Nayak, Nath, and Singhal 2018) and performance risk (Wang et al. 2016). Yuan et al. (2016) analysed PR as an antecedent factor and identified its significant negative relationship with behaviour intention in mobile banking. In the context of bike-sharing system, users need to scan the QR codes and pay through mobile payment system to unlock the bicycles (Si et al. 2020). PR and SR are constantly concerned by the customers when using the online services or mobile applications (Marafon et al. 2018). Additionally, due to the rapid market expansion, sharing bike companies are encountering enormous pressure from their capital chain which increases the FR of their users (Spinney and Lin 2018). Therefore, FR, PR and SR are considered as the components of perceived loss in our proposed model.

Data collection

Survey design

A survey study was conducted to collect the opinions towards the intention of using Mobike system among residents in Beijing and the questionnaire is designed based on our research model (Figure 2). The research model consists of six independent (input) variables and one dependent (output) variable. These independent variables are PU, PEU, SR, FR, PR and PD. The dependent variable is the IU. Each of the independent variable is measured by three items and the IU is measured by four items as shown in Table 1. A 5-point Likert scale is adopted where 1 means 'strongly disagree' and 5 represents 'strongly agree'.

Sampling and data collection

Direct data collection on the street and survey online was adopted. After initial data filtering and cleansing, 355 out of 400 received questionnaires (89%) are valid for further data analysis. The demographic information of samples is shown in Table 2, it consists of

Table 1. Measurement of variables.

Variable	Label	Measurement items	Reference
PU	PU1	You will consider the usefulness of sharing bike	Wang et al. (2015)
	PU2	Sharing bike can improve travel efficiency	
	PU3	Sharing bike is necessary for your daily life	
PEU	PEU1	It is easy for you to use sharing bike	Kim, Chan, and Gupta (2007)
	PEU2	You spend less energy in learning	
	PEU3	The steps to use sharing bike are simple	
FR	FR1	You will worry about of the failure of deposit	Al-Jabri (2015)
	FR2	You will worry about the illegal use of transaction information	
	FR3	You will worry about property loss caused by information embezzlement	
PR	PR1	You will worry about personal and trading information being leaked	Wang and Lin (2017)
	PR2	It's easy to get spam messages after using sharing bike	
	PR3	You will worry about personal data being illegal used	
SR	SR1	Personal safety problems can be generated by malfunction of sharing bike during driving	Li. et al. (2017)
	SR2	Positioning function of sharing bike can make you feel insecure	
	SR3	Bike design is vulnerable to damage and then poses safety risk	
PD	PD1	It is more economical for the use of sharing bike	Ritonga and Astuti (2019)
	PD2	Discount activities will promote your use of sharing bike	
	PD3	Sharing bike can bring you high-quality travel experience	
IU	IU1	You will think it a good decision to use sharing bike	Hwang and Lee (2013)
	IU2	You are insisting on using sharing bike	
	IU3	You are satisfied with the use of sharing bike	
	IU4	You will recommend your friends in using sharing bike	

Table 2. Demographic information of samples.

Description variable	Classification	Frequency (355)	Percentage (%)
Gender	Male	146	41.1
	Female	209	58.9
Age	18–25	144	40.6
	26–32	113	31.8
	33–39	74	20.8
	40–50	24	6.8
Education	Below high school	1	0.3
	High school	24	6.8
	University	191	53.8
	Master	81	22.8
	Doctor	58	16.3
Job	Student	25	7.1
	Teacher and researcher	65	18.3
	Company employee	156	43.9
	Executive staff	65	18.3
	Liberal profession	44	12.4
Monthly frequency	Below 3 times	71	20.0
	3 times to 6 times	99	27.9
	7 times to 10 times	85	23.9
	11 times to 14times	67	18.9
	Above 14 times	33	9.3
Average duration	Below 10 minutes	82	23.1
	10 minutes to 20 minutes	134	37.7
	20 minutes to 30 minutes	83	23.5
	30 minutes to 40 minutes	35	9.8
	Above 40 minutes	21	5.9
Primary travel scenario	Commuting	52	14.6
	Transfer	113	31.8
	Short distance travel	114	32.1
	Exercise	76	21.4

41.1% male respondents and 58.9% female respondents. The majority of the respondents are between 18 to 32 years old (72.4%) and only 6.8% are over 40 years old. In regard to the education level, almost 93% have university degree or above. The average duration of using a sharing bike for the majority of respondents with a percentage of 61.2% falls between 10 and 30 minutes. There are 63.9% of respondents choosing the sharing bikes for short distance travel and 21.4% for cycling exercise. The respondents who had never used Mobike system are classified as 'Below 3 times' in Monthly Frequency and 'Below 10 minutes' in Average Duration. The respondents are relatively young people with decent education qualifications and most of the respondents tend to use sharing bikes for short-distance travel.

Data analysis and results

Preliminary analysis

Descriptive statistical analysis

Descriptive analysis is conducted for testing the variability and distribution properties of the samples. In this section, average value, standard deviation, skewness and kurtosis of each measurement item are obtained, as shown in Table 3. The mean values of the items measured within the same variable are relatively close, for example, the mean values of PU, PEU and SR fluctuate around 4.40 and the values for PR and PD are around 4.30. The mean values for items in IU are 4.17, 3.97, 4.03 and 4.02 respectively. The mean values for FR varies between 2.5 and 3, which is lower than other factors. The standard deviation is comparatively small with the values oscillate around 1. In regard to skewness and kurtosis, which are used to measure the symmetry and combined size of tails of data distribution, the absolute values of skewness vary within 1.5 while the figure for kurtosis is less than 2. This suggests that the research samples tend to follow a normal distribution.

Table 3. Descriptive statistics analysis of variables.

Variable	Measurement item	Mean value	Standard deviation	Skewness	Kurtosis
PU	PU1	4.48	0.674	−1.003	0.033
	PU2	4.42	0.734	−0.886	−0.452
	PU3	4.56	0.646	−1.293	1.034
PEU	PEU1	4.47	0.779	−1.410	1.288
	PEU2	4.45	0.760	−1.284	0.996
	PEU2	4.28	0.911	−1.004	0.036
FR	FR1	2.41	1.412	0.630	−1.021
	FR2	2.49	1.353	−0.529	−0.968
	FR3	3.09	1.521	−0.144	−1.470
SR	SR1	4.45	0.729	−1.247	1.357
	SR2	4.38	0.789	−1.072	0.502
	SR3	4.53	0.682	−1.357	1.328
PR	PR1	4.33	0.852	−1.385	1.855
	PR2	4.32	0.874	−1.352	1.562
	PR3	4.37	0.843	−1.369	1.552
PD	PD1	4.44	0.712	−0.975	−0.033
	PD2	4.39	0.715	−0.838	−0.244
	PD3	4.38	0.747	−0.823	−0.456
IU	IU1	4.17	0.923	−0.853	−0.068
	IU2	3.97	1.270	−1.143	0.203
	IU3	4.03	1.134	−1.197	0.720
	IU4	4.02	1.146	−0.998	0.109

Reliability and validity

To measure the reliability of the sample data, the Cronbach's alpha and composite reliability (CR) are calculated. The average variance extracted (AVE) is analysed for verifying the convergent and discriminant validity. In addition, exploratory factor analysis with varimax rotation is performed to further validate the reliability and validity of the research samples. In Table 4, most values of Cronbach's alpha exceed 0.7 except for PD with 0.681. The high CR values (>0.7) indicate the reliability of survey design (Hew et al. 2017). Additionally, the AVE values fall in the range of 0.6 to 0.8 implying that the convergent and discriminant validity are supported. Results from factor analysis with all factor loadings greater than 0.7 with a significant level of 0.05 confirm the validity (Table 5).

Modelling of intention of using machine learning techniques

Five ML techniques are employed to construct classification models for the IU of bike-sharing system. These ML models are trained in a supervised learning framework with labelled data presented in Section 3. IU is represented by the nearest integer of the mean value of IU1, IU2, IU3, IU4 and IU5. As shown by Figure 3, four distinct classes represented by 2–5 are identified in our sample data and these four label classes are defined as the output of classification models. In order to facilitate the learning process, the input data is

Table 4. Cronbach's α, CR and AVE for variables.

Variable	Cronbach's alpha	CR	AVE
PU	0.789	0.849	0.807
PEU	0.841	0.852	0.811
FR	0.846	0.911	0.875
SR	0.763	0.837	0.794
PR	0.940	0.946	0.923
PD	0.681	0.781	0.732
IU	0.798	0.739	0.633

Table 5. Results of the reliability and validity analysis.

Variable	Measurement item	Factor loading	Communalities extraction	KMO value
PU	PU1	0.850	0.725	0.706
	PU2	0.842	0.710	
	PU3	0.827	0.682	
PEU	PEU1	0.888	0.691	0.699
	PEU2	0.912	0.786	
	PEU2	0.827	0.724	
FR	FR1	0.938	0.882	0.602
	FR2	0.958	0.914	
	FR3	0.736	0.571	
SR	SR1	0.814	0.676	0.696
	SR2	0.837	0.694	
	SR3	0.823	0.681	
PR	PR1	0.968	0.935	0.719
	PR2	0.962	0.932	
	PR3	0.904	0.814	
PD	PD1	0.762	0.581	0.601
	PD2	0.819	0.692	
	PD3	0.815	0.693	
IU	IU1	0.749	0.635	0.719
	IU2	0.794	0.784	
	IU3	0.830	0.690	
	IU4	0.790	0.599	

normalised in the range of [0, 1] through cross channel normalisation method (Wang et al. 2015).

To avoid overfitting of our model, K-fold cross validation is implemented with 80% of data is used for training purpose and 20% is used for testing (Azadi and Karimi-Jashni 2016). Five ML learners are adopted and compared in this study: SVM, RF, GBDT, ANN) and k-nearest neighbours (KNN). Python 3.4 with Keras, Scikit-learn and Tensorflow packages is employed for the modelling, training and evaluating processes. Sensitivity analysis is then conducted on the ML model with best performance in order to explore the relationship between the input factors and the output (IU) as well as the relative importance level for these factors.

Parameters tuning

Every ML algorithms or models have pre-defined parameters that could greatly affect the performance (Maxwell, Warner, and Fang 2018). To select the appropriate parameters for each ML model, a range of combinations for different parameter values are tested through a grid search algorithm with five-fold cross validation, and this process is repeated for five times for each model (Heung et al. 2016). Table 6 below shows the parameters that need tuning for each ML model. For ANN, the number of neurons in hidden layers is mainly determined based on the rule of thumb, and one of the most commonly acknowledged rules is that the number of hidden neurons is between the number of inputs and outputs (Liu and Yoon 2019). To avoid over-fitting caused by too many epochs, early stopping method is applied in ANN to stop training when the loss is no longer improved (Chi et al. 2019). For RF, the number of trees are assessed for all values

Table 6. Parameters for tuning.

Classifier	Parameters	Reference	Best value
ANN	Neurons per layer	(Liu and Yoon 2019)	16
	Number of layers		2
Random forest	Number of trees	(Dao et al. 2017)	150
	Number of features for splitting		8
GBDT	Learning rate	(Touzani, Granderson, and Fernandes 2018)	0.025
	Number of trees		100
	Depth of trees		5
SVM	Kernel coefficient	(Thanh Noi and Kappas 2018)	0.001
	Regularisation parameter		10
KNN	Number of neighbours	(Xie et al. 2018)	8

in ½10; 50; 100; 150; 200; 250; 500; 1000] and the search range for number of features for splitting is from 1 to the number of inputs (Xie et al. 2018). For GBDT, a large depth of trees might result in the over-fitting (Touzani, Granderson, and Fernandes 2018). Through grid search process, the best values for the parameters of every model are obtained and Table 6 indicates the parameters for tuning as well as their selected values.

Performance evaluation and comparison

In order to assess the classification performance of each ML model, evaluation metrics including accuracy, precision, recall, F1-score and AUC (Area under the curve) score are introduced in this section. The calculation of these evaluation metrics is shown as below.

$$accuracy = \frac{TP}{TP + TN + FP + FN}$$

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F1 - score = 2 * \frac{precision * recall}{precision + recall}$$

where *TP*, *FP*, *TN* and *FN* represent true positive, false positive, true negative and false negative respectively. Precision and recall measure the ratio of positive predicted values and the sensitivity of the ML model, and F1-score is the harmonic mean of precision and recall. AUC score is the value of area under the receiver operating characteristic (ROC) curve with the best value of 1 and ROC curve is commonly used for identifying the relationship between true positive rate (sensitivity) and false positive rate (specificity) of a classifier (Dao et al. 2017).

Taking the randomness of data partitioning into consideration, the 10-fold cross validation process for each ML model is conducted repeatedly and the average value of each performance metrics are recorded (Austin et al. 2013). Table 7 shows results obtained from the five different ML models after 5 and 10 repetition of 10-fold cross validation. In terms of accuracy, GBDT is highest in both 5 repetition and 10 repetition cases with values of 0.8016 and 0.8169 while SVM and KNN are the two less accurate models with only 0.7025 and

Table 7. Comparison of machine learning models.

10-fold	Classifier	Performance matrices				
		ACC	Precision	Recall	AUC	F1
5	ANN	0.7704	0.7275	0.5833	0.8757	0.6475
	Random forest	0.7767	0.7508	0.7183	0.8874	0.7342
	Gradient boosting	0.8016	0.8592	0.7465	0.9265	0.7989
	SVM	0.7025	0.6873	0.5493	0.8635	0.6105
	KNN	0.6480	0.6101	0.4807	0.8479	0.5377
10	ANN	0.8046	0.7389	0.5988	0.8964	0.6615
	Random forest	0.7903	0.7863	0.7524	0.9026	0.7689
	Gradient boosting	0.8169	0.8179	0.7606	0.9366	0.7882
	SVM	0.7339	0.7088	0.5634	0.8656	0.6278
	KNN	0.6788	0.6156	0.5352	0.8193	0.5726

0.6480 in 5 repetition case and 0.7339 and 0.6788 in 10 repetition case respectively. ANN and RF provide almost the same accuracy with 0.7704 and 0.7767 in 5 repetition case and 0.8046 and 0.7903 in 10 repetition case. Regarding the precision, recall and F1-score, it is obvious that GBDT model outperforms other ML models while KNN has the lowest F1-score of 0.5377. The precision value of ANN is slightly smaller than the value of RF in both cases while the recall values of ANN are only 0.5833 (5 repetitions) and 0.5988 (10 repetitions) which result in low values of F1-score for ANN. Regarding the AUC score, GBDT outperforms the others with the values of 0.9265 in 5 repetition case and 0.9366 in 10 repetition case followed by the RF and the ANN. Figure 4 shows the ROC curves and micro-average values for AUC score of each model. The GBDT performs the best among all five ML models in regard of evaluation metrics. Therefore, GBDT model is selected to further explore the relationship between the input factors and the output of the proposed research model determining the IU for bike-sharing system.

Sensitivity Analysis

Sensitivity analysis is a means of identifying how the uncertainty of output variable can be influenced by its dependent input variables (Leong et al. 2015). Through conducting a sensitivity analysis, the relative importance of inputs on the output can be quantified. One-factor-at-a-time algorithm is one of the most widely used sensitivity methods which analyses the sensitivity by changing the values of one variable at a time while keeping the other variables fixed (Jung et al. 2010). However, it is not applicable to nonlinear models. Therefore, in this section, Sobol's sensitivity analysis is adopted. As a popular method of variance based sensitivity analysis, Sobol's sensitivity analysis aims to measure the variance of the contribution to the output that each input variable has made, which is capable of measuring the sensitivity across the whole input space and dealing with nonlinear relationships (Rosolem et al. 2012).

Sobol's sensitivity analysis is conducted for the selected GBDT model to explore and quantify the importance of input factors affecting the IU. The relative and normalised importance values of each input variable are obtained and shown in Table 8. The PEU is found to be the most importance factor influencing the IU with 100% normalised importance level followed by PU with importance of 0.256. PR and FR are at the third and fourth positions with relative importance of 0.196 and 0.175 respectively. PPD is the fourth most significant factor with 0.126 importance level where as the SR is the least important influencing factor with 0.051. Comparing to PBs, perceived loss does not have significant contribution to the IU for sharing bike.

To further verify the impacts of factors on IU, one-factor-at-a-time (OFAT) sensitivity analysis is applied to explore the variances in accuracy of GBDT by excluding factors with small importance level, SR and PD. A 10-fold cross validation is repeated for 10 iterations and the average accuracy scores for removing SR and PD are shown in Table 9. The average accuracy of GBDT model after removing SR is 0.8292 which is slightly higher than 0.8169, the accuracy of GBDT without removing any factors. On the contrary, applying the GBDT without PD results in a small decrease in the accuracy. By excluding both PD and SR, the result shows the worst accuracy score with 0.8044. Therefore, the OFAT results indicate that SR is the only factor creating noise to the model training and it can be removed from the proposed model.

Table 8. Normalised variable importance.

Variable	Normalised importance
PU	0.256
PEU	1.000
FR	0.175
SR	0.051
PR	0.196
PD	0.126

Table 9. Cross-validation results by excluding least important factors.

10-fold CV	Drop SR	Drop PD	Drop ALL
1	0.7822	0.8421	0.8125
2	0.8421	0.8147	0.7500
3	0.8347	0.8421	0.8375
4	0.8174	0.7868	0.7368
5	0.7668	0.7522	0.8421
6	0.8421	0.8889	0.7668
7	0.8889	0.7667	0.8189
8	0.8525	0.7530	0.8347
9	0.8125	0.8375	0.8421
10	0.8530	0.8125	0.8025
Average	0.8292	0.8097	0.8044
Std	0.0358	0.0448	0.0395

Discussion

Contributions and practical implications

The contributions of this study can be two-fold. On one hand, this research has made several theoretical and methodological contributions. Firstly, instead of adopting a single research model such as TAM or PRM, the study extends the research model by taking both PBs and PR into consideration. This allows us to comprehensively analyse potential factors for IU in the context of bike-sharing system. Secondly, the adoption of ML techniques, rather than other conventional analysis methodologies such as SEM, addresses the limitations of making assumptions on variable distributions and hypothesis testing. The ML techniques introduced in this study are capable of exploring non-linear relationships between variables. By selecting the best ML model, the importance values of influencing factors could be prioritised and quantified in an accurate manner.

Our findings also provide a practical reference for bike-sharing companies as well as the concerned stakeholders. PEU is found as the most important contributor to IU for bike-sharing system. Therefore, the design of online platform should provide an 'ease-of-use' interface for better user experience and companies should also optimise the deployment of bicycles with an effective bikes recycling strategy. Compared to PEU and PU, FR and PR are considered less important for IU. This is probably because of the low cost of renting sharing bikes and the lack of awareness of potential privacy and FRs, it is still important for both enterprises and local governments to better protect the information and financial benefits of customers. As PD is regarded as an influencing factor, sharing bike companies could put forward promotion strategies to attract more potential users and to retain the loyal customers by improving their services. The analytics approach proposed in this study can be further extended to other enterprises for various management applications.

Research limitations and future works

The research has few limitations as follows. Firstly, this research mainly focuses on the users in Beijing, China where bike-sharing system has been widely deployed in. However, bike-sharing system is still a novel application to people living in smaller cities and their opinions towards bike-sharing might be different. Secondly, the limit size and imbalanced output classes could also impede the performance of ML algorithms. This study can be further extended to others cities in China as well as in other countries, such as Australia. Comparisons can be performed to find out how these factors could influence the IU of customers of different backgrounds, and as well for better ML performance.

Conclusion

This study proposes a research model to identify potential influencing factors of IU for bike-sharing system in China and to quantify the importance level of these factors by using ML techniques. Survey study was carried out in Beijing, China with 355 valid responses. Results from the comparison of five different ML techniques show that GBDT outperforms others in terms of accuracy, precision, recall, F1-score and AUC score. Our findings offer practical references for bike-sharing companies as well as concern stakeholders. PEU is found as the most important contributor to IU for bike-sharing system. Therefore, the online platform should be of 'ease-of-use' for better user experience and companies should establish more effective strategies for bicycle deployment and recycling.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Abiodun, O. I., A. Jantan, A. E. Omolara, K. V. Dada, N. A. Mohamed, and H. Arshad. 2018. "State-of-the-art in Artificial Neural Network Applications: A Survey." *Heliyon* 4 (11): e00938. doi:10.1016/j.heliyon.2018.e00938.

Ajzen, I. 1991. "The Theory of Planned Behavior." *Organizational Behavior and Human Decision Processes* 50 (2): 179–211. doi:10.1016/0749-5978(91)90020-T.

Al-Jabri, I. M. 2015. "The Intention to Use Mobile Banking: Further Evidence from Saudi Arabia." *South African Journal of Business Management* 46 (1): 23–34. doi:10.4102/sajbm.v46i1.80.

Amayri, O., and N. Bouguila. 2010. "A Study of Spam Filtering Using Support Vector Machines." *Artificial Intelligence Review* 34 (1): 73–108. doi:10.1007/s10462-010-9166-x.

Arpaci, I. 2016. "Understanding and Predicting Students' Intention to Use Mobile Cloud Storage Services." *Computers in Human Behavior* 58: 150–157. doi:10.1016/j.chb.2015.12.067.

Asadi, S., R. Abdullah, M. Safaei, and S. Nazir. 2019. "An Integrated SEM-Neural Network Approach for Predicting Determinants of Adoption of Wearable Healthcare Devices." *Mobile Information Systems* 2019.

Austin, P. C., J. V. Tu, J. E. Ho, D. Levy, and D. S. Lee. 2013. "Using Methods from the Data-mining and Machine-learning Literature for Disease Classification and Prediction: A Case Study Examining Classification of Heart Failure Subtypes." *Journal of Clinical Epidemiology* 66 (4): 398–407. doi:10.1016/j.jclinepi.2012.11.008.

Azadi, S., and A. Karimi-Jashni. 2016. "Verifying the Performance of Artificial Neural Network and Multiple Linear Regression in Predicting the Mean Seasonal Municipal Solid Waste Generation

Rate: A Case Study of Fars Province, Iran." *Waste Management* 48: 14–23. doi:10.1016/j.wasman.2015.09.034.

Babajide Mustapha, I., and F. Saeed. 2016. "Bioactive Molecule Prediction Using Extreme Gradient Boosting." *Molecules* 21 (8): 983. doi:10.3390/molecules21080983.

Belgiu, M., and L. Drăguţ. 2016. "Random Forest in Remote Sensing: A Review of Applications and Future Directions." *ISPRS Journal of Photogrammetry and Remote Sensing* 114: 24–31. doi:10.1016/j.isprsjprs.2016.01.011.

Belk, R. 2014. "You are What You Can Access: Sharing and Collaborative Consumption Online." *Journal of Business Research* 67 (8): 1595–1600. doi:10.1016/j.jbusres.2013.10.001.

Bose, B. K. 2007. "Neural Network Applications in Power Electronics and Motor drives—An Introduction and Perspective." *IEEE Transactions on Industrial Electronics* 54 (1): 14–33. doi:10.1109/TIE.2006.888683.

Breiman, L. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32. doi:10.1023/A:1010933404324.

Brown, I., and C. Mues. 2012. "An Experimental Comparison of Classification Algorithms for Imbalanced Credit Scoring Data Sets." *Expert Systems with Applications* 39 (3): 3446–3453. doi:10.1016/j.eswa.2011.09.033.

Brusch, I., and N. Rappel. 2019. "Exploring the Acceptance of Instant shopping–An Empirical Analysis of the Determinants of User Intention." *Journal of Retailing and Consumer Services* 54: 101936.

Chaudhuri, A., and K. De. 2011. "Fuzzy Support Vector Machine for Bankruptcy Prediction." *Applied Soft Computing* 11 (2): 2472–2486. doi:10.1016/j.asoc.2010.10.003.

Chen, S.-Y., and C. -C. Lu. 2016. "A Model of Green Acceptance and Intentions to Use Bike-sharing: YouBike Users in Taiwan." *Networks and Spatial Economics* 16 (4): 1103–1124. doi:10.1007/s11067-015-9312-8.

Chi, J., L. Xiaolei, H. Wang, D. Gao, and P. Gerstoft. 2019. "Sound Source Ranging Using a Feed-forward Neural Network Trained with Fitting-based Early Stopping." *The Journal of the Acoustical Society of America* 146 (3): EL258–EL64. doi:10.1121/1.5126115.

China Electornic Commence Research Centre, 2017. "China's Sharing Economy Development Report in 2016." Hangzhou, China: China Electronic Commence Research Centre.

Colditz, R. R. 2015. "An Evaluation of Different Training Sample Allocation Schemes for Discrete and Continuous Land Cover Classification Using Decision Tree-based Algorithms." *Remote Sensing* 7 (8): 9655–9681. doi:10.3390/rs70809655.

Dao, F.-Y., H. Yang, S. Zhen-Dong, W. Yang, W. Yun, D. Hui, W. Chen, H. Tang, and H. Lin. 2017. "Recent Advances in Conotoxin Classification by Using Machine Learning Methods." *Molecules* 22 (7): 1057. doi:10.3390/molecules22071057.

Davis, F. D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology." *MIS Quarterly* 13 (3): 319–340. doi:10.2307/249008.

Deb, C., S. E. Lee, and M. Santamouris. 2018. "Using Artificial Neural Networks to Assess HVAC Related Energy Saving in Retrofitted Office Buildings." *Solar Energy* 163: 32–44. doi:10.1016/j.solener.2018.01.075.

Du, M., and L. Cheng. 2018. "Better Understanding the Characteristics and Influential Factors of Different Travel Patterns in Free-floating Bike Sharing: Evidence from Nanjing, China." *Sustainability* 10 (4): 1244. doi:10.3390/su10041244.

Featherman, M. S., and P. A. Pavlou. 2003. "Predicting E-services Adoption: A Perceived Risk Facets Perspective." *International Journal of Human-computer Studies* 59 (4): 451–474. doi:10.1016/S1071-5819(03)00111-3.

Fishbein, M., and I. Ajzen. 1977. "Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research." *Journal of Business Venturing* 5: 177-189.

Forsythe, S. M., and B. Shi. 2003. "Consumer Patronage and Risk Perceptions in Internet Shopping." *Journal of Business Research* 56 (11): 867–875. doi:10.1016/S0148-2963(01)00273-9.

Friedman, J. H. 2001. "Greedy Function Approximation: A Gradient Boosting Machine." *Annals of Statistics* 1189–1232. doi:10.1214/aos/1013203451.

Gerlein, E. A., M. McGinnity, A. Belatreche, and S. Coleman. 2016. "Evaluating Machine Learning Classification for Financial Trading: An Empirical Approach." *Expert Systems with Applications* 54: 193–207. doi:10.1016/j.eswa.2016.01.018.

Grewal, D., R. Krishnan, J. Baker, and N. Borin. 1998. "The Effect of Store Name, Brand Name and Price Discounts on Consumers' Evaluations and Purchase Intentions." *Journal of Retailing* 74 (3): 331–352. doi:10.1016/S0022-4359(99)80099-2.

Habibi, M. R., A. Davidson, and M. Laroche. 2017. "What Managers Should Know about the Sharing Economy." *Business Horizons* 60 (1): 113–121. doi:10.1016/j.bushor.2016.09.007.

Heung, B., H. C. Ho, J. Zhang, A. Knudby, C. E. Bulmer, and M. G. Schmidt. 2016. "An Overview and Comparison of Machine-learning Techniques for Classification Purposes in Digital Soil Mapping." *Geoderma* 265: 62–77. doi:10.1016/j.geoderma.2015.11.014.

Hew, J.-J., L.-Y. Leong, G. Wei-Han Tan, K.-B. Ooi, and V.-H. Lee. 2017. "The Age of Mobile Social Commerce: An Artificial Neural Network Analysis on Its Resistances." *Technological Forecasting and Social Change* 144: 311-324.

Hsieh, P.-J. 2015. "Physicians' Acceptance of Electronic Medical Records Exchange: An Extension of the Decomposed TPB Model with Institutional Trust and Perceived Risk." *International Journal of Medical Informatics* 84 (1): 1–14. doi:10.1016/j.ijmedinf.2014.08.008.

Hsu, C.-L., and J.-C.-C. Lin. 2018. "Exploring Factors Affecting the Adoption of Internet of Things Services." *Journal of Computer Information Systems* 58 (1): 49–57. doi:10.1080/08874417.2016.1186524.

Hwang, K., and E. Lee. 2013. "The Role of Entertainment and External Values in Consumer's Intention to Use Social Network Mobile." *International Proceedings of Economics Development and Research* 63: 33.

Iranmanesh, M., K. Jayaraman, S. Zailani, and S. M. Ghadiri. 2017. "The Effects of Consumer Perception of Volume Discount Benefits on Intention to Purchase Grocery Products: Deal Proneness as a Moderator." *Asia Pacific Journal of Marketing and Logistics* 29 (5): 1017–1035. doi:10.1108/APJML-07-2016-0135.

Jacoby, J., and L. B. Kaplan. 1972. "The Components of Perceived Risk." Paper presented at the Proceedings of the Third Annual Conference of the Association for Consumer Research, 382-393. Chicago, Illinois, USA.

Jia, L., X. Liu, and Y. Liu. 2018. "Impact of Different Stakeholders of Bike-sharing Industry on Users' Intention of Civilized Use of Bike-sharing." *Sustainability* 10 (5): 1437. doi:10.3390/su10051437.

Jung, Y. W., O. Dong-Shig, M. Kim, and J.-W. Park. 2010. "Calibration of LEACHN Model Using LH-OAT Sensitivity Analysis." *Nutrient Cycling in Agroecosystems* 87 (2): 261–275. doi:10.1007/s10705-009-9337-9.

Kim, H.-W., H. C. Chan, and S. Gupta. 2007. "Value-based Adoption of Mobile Internet: An Empirical Investigation." *Decision Support Systems* 43 (1): 111–126. doi:10.1016/j.dss.2005.05.009.

Kim, Y., Y. Park, and J. Choi. 2017. "A Study on the Adoption of IoT Smart Home Service: Using Value-based Adoption Model." *Total Quality Management & Business Excellence* 28 (9–10): 1149–1165. doi:10.1080/14783363.2017.1310708.

Lai, W.-T. 2015. "Exploring Use Intention of a Smart Bike-sharing System-extending Technology Acceptance Model with Trust." Paper presented at the *LISS 2014*, 1597–1603. Springer, Berlin, Heidelberg.

Lee, M.-C. 2009. "Factors Influencing the Adoption of Internet Banking: An Integration of TAM and TPB with Perceived Risk and Perceived Benefit." *Electronic Commerce Research and Applications* 8 (3): 130–141. doi:10.1016/j.elerap.2008.11.006.

Leong, L.-Y., T.-S. Hew, V.-H. Lee, and K.-B. Ooi. 2015. "An SEM–artificial-neural-network Analysis of the Relationships between SERVPERF, Customer Satisfaction and Loyalty among Low-cost and Full-service Airline." *Expert Systems with Applications* 42 (19): 6620–6634. doi:10.1016/j.eswa.2015.04.043.

Li, Z., W. Wang, C. Yang, and H. Ding. 2017. "Bicycle Mode Share in China: A City-level Analysis of Long Term Trends." *Transportation* 44 (4): 773-88.

Liang, S., and R. Srikant. 2016. "Why Deep Neural Networks for Function Approximation?" *arXiv preprintarXiv:1610.04161*.

Liébana-Cabanillas, F., V. Marinković, and Z. Kalinić. 2017. "A SEM-neural Network Approach for Predicting Antecedents of M-commerce Acceptance." *International Journal of Information Management* 37 (2): 14–24. doi:10.1016/j.ijinfomgt.2016.10.008.

Liébana-Cabanillas, F., V. Marinkovic, I. R. de Luna, and Z. Kalinic. 2018. "Predicting the Determinants of Mobile Payment Acceptance: A Hybrid SEM-neural Network Approach." *Technological Forecasting and Social Change* 129: 117–130. doi:10.1016/j.techfore.2017.12.015.

Liu, M., M. Wang, J. Wang, and L. Duo. 2013. "Comparison of Random Forest, Support Vector Machine and Back Propagation Neural Network for Electronic Tongue Data Classification: Application to the Recognition of Orange Beverage and Chinese Vinegar." *Sensors and Actuators. B, Chemical* 177: 970–980. doi:10.1016/j.snb.2012.11.071.

Liu, T.-H., and Y. Yoon. 2019. "Development of Enhanced Emission Factor Equations for Paved and Unpaved Roads Using Artificial Neural Network." *Transportation Research Part D: Transport and Environment* 69: 196–208. doi:10.1016/j.trd.2019.01.033.

Liu, Y., and Y. Yang. 2018. "Empirical Examination of Users' Adoption of the Sharing Economy in China Using an Expanded Technology Acceptance Model." *Sustainability* 10 (4): 1262. doi:10.3390/su10041262.

Liu, Z. Y. 2015. "An Analysis of Technology Acceptance Model-Exploring user acceptance and intension of taxi-hailing app in Shanghai." University of Gothenburg.

Ma, X., C. Ding, S. Luan, Y. Wang, and Y. Wang. 2017. "Prioritizing Influential Factors for Freeway Incident Clearance Time Prediction Using the Gradient Boosting Decision Trees Method." *IEEE Transactions on Intelligent Transportation Systems* 18 (9): 2303–2310. doi:10.1109/TITS.2016.2635719.

Ma, Y., and H. Zhang. 2019. "Development of the Sharing Economy in China: Challenges and Lessons." In *Innovation, Economic Development and Intellectual Property in India and China*, 467–484. Springer, Singapore.

Marafon, D. L., K. Basso, L. B. Espartel, M. D. de Barcellos, and E. Rech. 2018. "Perceived Risk and Intention to Use Internet Banking: The Effects of Self-confidence and Risk Acceptance." *International Journal of Bank Marketing* 36 (2): 277–289. doi:10.1108/IJBM-11-2016-0166.

Mathwick, C., N. K. Malhotra, and E. Rigdon. 2002. "The Effect of Dynamic Retail Experiences on Experiential Perceptions of Value: An Internet and Catalog Comparison." *Journal of Retailing* 78 (1): 51–60. doi:10.1016/S0022-4359(01)00066-5.

Maxwell, A. E., T. A. Warner, and F. Fang. 2018. "Implementation of Machine-learning Classification in Remote Sensing: An Applied Review." *International Journal of Remote Sensing* 39 (9): 2784–2817. doi:10.1080/01431161.2018.1433343.

McAfee, A., E. Brynjolfsson, T. H. Davenport, D. J. Patil, and D. Barton. 2012. "Big Data: The Management Revolution." *Harvard Business Review* 90 (10): 60–68.

Nayak, N., V. Nath, and N. Singhal. 2018. "Predicting the Risk Factors Influencing the Behavioral Intention to Adopt Mobile Banking Services: An Exploratory Analysis." Paper presented at the International Conference on Futuristic Trends in Network and Communication Technologies,176-191. Springer, Singapore.

Nusair, K., H. J. Yoon, S. Naipaul, and H. G. Parsa. 2010. "Effect of Price Discount Frames and Levels on Consumers' Perceptions in Low-end Service Industries." *International Journal of Contemporary Hospitality Management* 22 (6): 814–835. doi:10.1108/09596111011063106.

Pappas, N. 2016. "Marketing Strategies, Perceived Risks, and Consumer Trust in Online Buying Behaviour." *Journal of Retailing and Consumer Services* 29: 92–103. doi:10.1016/j.jretconser.2015.11.007.

Peter, J. P., and L. X. Tarpey Sr. 1975. "A Comparative Analysis of Three Consumer Decision Strategies." *Journal of Consumer Research* 2 (1): 29–37. doi:10.1086/208613.

Pham, S. 2019. "Chinese Bike-sharing Startup Ofo Went Gloabl. Now It May Go Bust." Accessed October 26. https://edition.cnn.com/2018/12/21/tech/ofo-china-bike-sharing-crisis/index.html

Priyadarshinee, P., R. D. Raut, M. K. Jha, and B. B. Gardas. 2017. "Understanding and Predicting the Determinants of Cloud Computing Adoption: A Two Staged Hybrid SEM-Neural Networks Approach." *Computers in Human Behavior* 76: 341–362. doi:10.1016/j.chb.2017.07.027.

Qu, C., L. Yang, W. Bruce Croft, Y. Zhang, J. R. Trippas, and M. Qiu. 2019. "User Intent Prediction in Information-seeking Conversations." Paper presented at the Proceedings of the 2019 Conference on Human Information Interaction and Retrieval, 25-33. Glasgow, Scotland, UK.

Rawat, A., R. N. Yadav, and S. C. Shrivastava. 2012. "Neural Network Applications in Smart Antenna Arrays: A Review." *AEU-International Journal of Electronics and Communications* 66 (11): 903–912. doi:10.1016/j.aeue.2012.03.012.

Ricci, M. 2015. "Bike Sharing: A Review of Evidence on Impacts and Processes of Implementation and Operation." *Research in Transportation Business & Management* 15: 28–38. doi:10.1016/j.rtbm.2015.03.003.

Ritonga, N., and E. S. Astuti. 2019. "The Influence of Perceived Ease of Use, Discount, and Perceived Usefulness on Intention to Use Grab Application and Its Impact on Purchase Decision of Grab Services." *The International Journal of Accounting and Business Society* 27 (2): 55-70.

Rosolem, R., H. V. Gupta, W. James Shuttleworth, X. Zeng, and L. G. G. De Gonçalves. 2012. "A Fully Multiple-criteria Implementation of the Sobol' Method for Parameter Sensitivity Analysis." *Journal of Geophysical Research: Atmospheres* 117 (D7). doi:10.1029/2011JD016355.

Sahbi, H., J.-Y. Audibert, and R. Keriven. 2010. "Context-dependent Kernels for Object Classification." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33 (4): 699–708. doi:10.1109/TPAMI.2010.198.

Si, H., J.-G. Shi, D. Tang, W. Guangdong, and J. Lan. 2020. "Understanding Intention and Behavior toward Sustainable Usage of Bike Sharing by Extending the Theory of Planned Behavior." *Resources, Conservation and Recycling* 152: 104513. doi:10.1016/j.resconrec.2019.104513.

Singh, K. P., S. Gupta, and P. Rai. 2013. "Identifying Pollution Sources and Predicting Urban Air Quality Using Ensemble Learning Methods." *Atmospheric Environment* 80: 426–437. doi:10.1016/j.atmosenv.2013.08.023.

Spinney, J., and W.-I. Lin. 2018. "Are You Being Shared? Mobility, Data and Social Relations in Shanghai's Public Bike Sharing 2.0 Sector." *Applied Mobilities* 3 (1): 66–83. doi:10.1080/23800127.2018.1437656.

Stone, R. N., and K. Grønhaug. 1993. "Perceived Risk: Further Considerations for the Marketing Discipline." *European Journal of Marketing* 27 (3): 39–50. doi:10.1108/03090569310026637.

Sustrova, T. 2016. "An Artificial Neural Network Model for a Wholesale Company's Order-Cycle Management." *International Journal of Engineering Business Management* 8: 2. doi:10.5772/63727.

Tan, G. W.-H., K.-B. Ooi, L.-Y. Leong, and B. Lin. 2014. "Predicting the Drivers of Behavioral Intention to Use Mobile Learning: A Hybrid SEM-Neural Networks Approach." *Computers in Human Behavior* 36: 198–213. doi:10.1016/j.chb.2014.03.052.

Tehrany, M. S., B. Pradhan, S. Mansor, and N. Ahmad. 2015. "Flood Susceptibility Assessment Using GIS-based Support Vector Machine Model with Different Kernel Types." *Catena* 125: 91–101. doi:10.1016/j.catena.2014.10.017.

Thanh Noi, P., and M. Kappas. 2018. "Comparison of Random Forest, K-nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery." *Sensors* 18 (1): 18.

Touzani, S., J. Granderson, and S. Fernandes. 2018. "Gradient Boosting Machine for Modeling the Energy Consumption of Commercial Buildings." *Energy and Buildings* 158: 1533–1543. doi:10.1016/j.enbuild.2017.11.039.

Venkatesh, V., and F. D. Davis. 2000. "A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies." *Management Science* 46 (2): 186–204. doi:10.1287/mnsc.46.2.186.11926.

Wahyuni, T. 2016. "The Influence of Technology Acceptance Model (TAM) on the Users' Behavior of Sikesya Application in IAIN Surakarta." *Shirkah: Journal of Economics and Business* 1 (1): 47–72. doi:10.22515/shirkah.v1i1.15.

Wang, E. S.-T., and R.-L. Lin. 2017. "Perceived Quality Factors of Location-based Apps on Trust, Perceived Privacy Risk, and Continuous Usage Intention." *Behaviour & Information Technology* 36 (1): 2–10.

Wang, L.-C., H.-M. Chuang, Y.-S. Chen, C.-K. Lin, and Y.-S. Wang. 2016. "A Study on the Acceptance Intention of "My Health Bank" from the Perspective of Health Belief Model." Paper presented at the International Conference on Frontier Computing, Singapore, 905-910.

Wang, Y., L. Jian, G. Junzhong, Z. Zhou, and Z. Wang. 2015. "Artificial Neural Networks for Infectious Diarrhea Prediction Using Meteorological Factors in Shanghai (China)." *Applied Soft Computing* 35: 280–290. doi:10.1016/j.asoc.2015.05.047.

Wang, Y., S. Wang, J. Wang, J. Wei, and C. Wang. 2018. "An Empirical Study of Consumers' Intention to Use Ride-sharing Services: Using an Extended Technology Acceptance Model." *Transportation 47:* 1–19.

Wang, Y., and W. Y. Szeto. 2018. "Static Green Repositioning in Bike Sharing Systems with Broken Bikes." *Transportation Research Part D: Transport and Environment* 65: 438–457. doi:10.1016/j.trd.2018.09.016.

Wu, C. H., Y. S. Wong, W. H. Ip, H. C. W. Lau, C. K. M. Lee, and G. T. S. Ho. 2009. "Modeling the Cleanliness Level of an Ultrasonic Cleaning System by Using Design of Experiments and Artificial Neural Networks." *The International Journal of Advanced Manufacturing Technology* 41 (3–4): 287–300. doi:10.1007/s00170-008-1471-z.

Xie, Y., C. Zhu, W. Zhou, L. Zhongdong, X. Liu, and T. Mei. 2018. "Evaluation of Machine Learning Methods for Formation Lithology Identification: A Comparison of Tuning Processes and Model Performances." *Journal of Petroleum Science and Engineering* 160: 182–193. doi:10.1016/j.petrol.2017.10.028.

Yao, Y., L. Liu, Z. Guo, Z. Liu, and H. Zhou. 2019. "Experimental Study on Shared Bike Use Behavior under Bounded Rational Theory and Credit Supervision Mechanism." *Sustainability* 11 (1): 127. doi:10.3390/su11010127.

Yaraghi, N., and S. Ravi. 2017. "The Current and Future State of the Sharing Economy." *Available at SSRN 3041207*.

Yu, Y., Y. Wenjie, Y. Feng, and J. Liu. 2018. "Understanding the Intention to Use Commercial Bike-sharing Systems: An Integration of TAM and TPB." Paper presented at the Proceedings of the 51st Hawaii International Conference on System Sciences, 646-655. Honolulu, Hawaii, USA.

Yuan, S., Y. Liu, R. Yao, and J. Liu. 2016. "An Investigation of Users' Continuance Intention Towards Mobile Banking in China." *Information Development* 32 (1): 20–34. doi:10.1177/0266666914522140.

Zhang, H., S. A. Shaheen, and X. Chen. 2014. "Bicycle Evolution in China: From the 1900s to the Present." *International Journal of Sustainable Transportation* 8 (5): 317–335. doi:10.1080/15568318.2012.699999.

Zhang, Y., and Z. Mi. 2018. "Environmental Benefits of Bike Sharing: A Big Data-based Analysis." *Applied Energy* 220: 296–301. doi:10.1016/j.apenergy.2018.03.101.

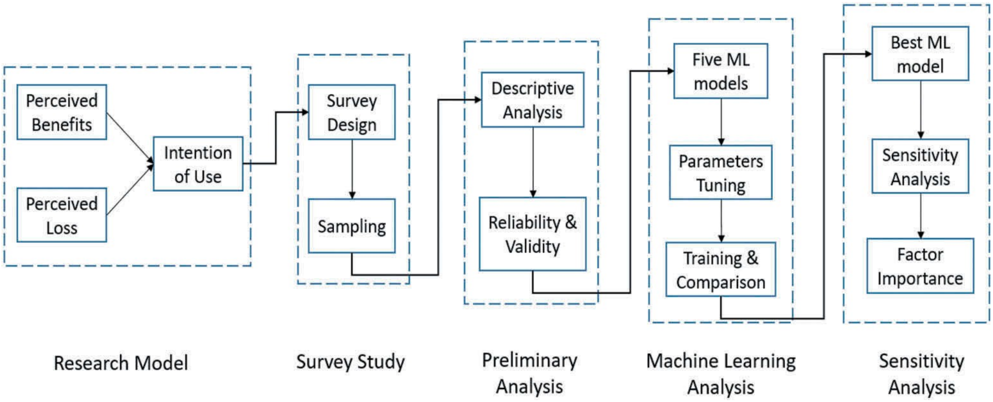
Zheng, B. 2019. "Financial Default Payment Predictions Using a Hybrid of Simulated Annealing Heuristics and Extreme Gradient Boosting Machines." *International Journal of Internet Technology and Secured Transactions* 9 (4): 404–425. doi:10.1504/IJITST.2019.102796.

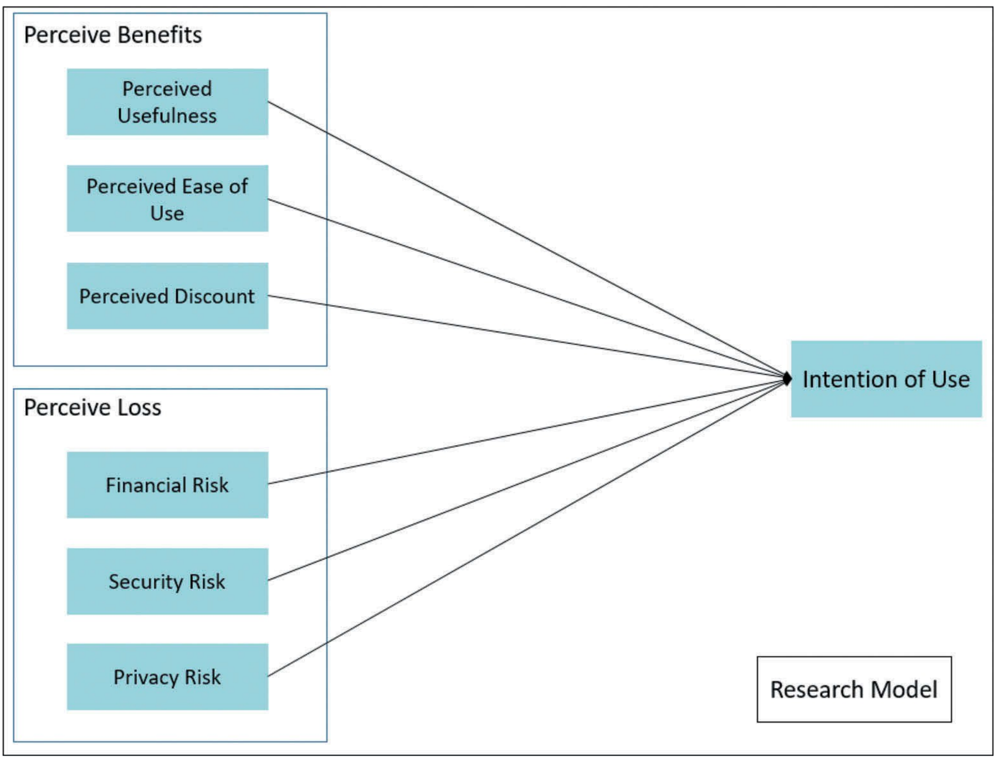
Figure 1. Framework of analysis

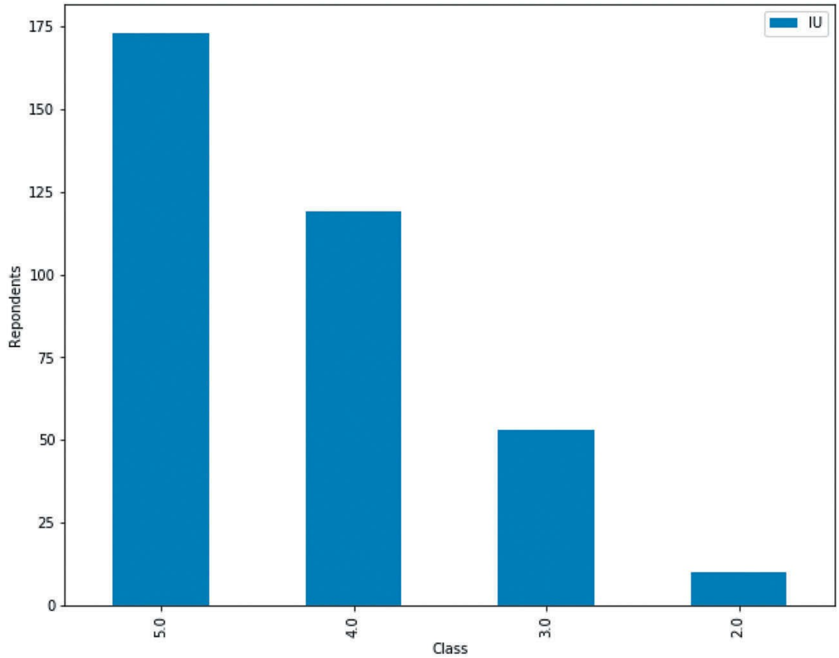
Figure 2. Framework of research model

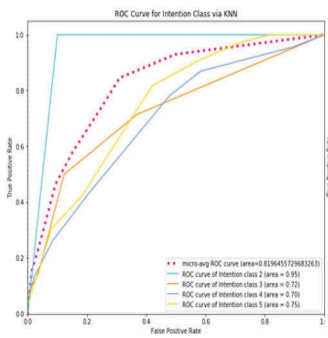
Figure 3. Number of respondents for each class

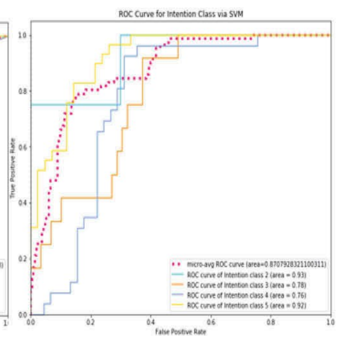
Figure 4. ROC curves for machine learning models

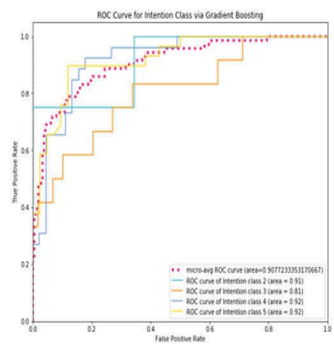

Perceived Benefits
Perceived Loss
Intention of Use
Survey Design
Sampling
Descriptive Analysis
Reliability & Validity
Five ML models
Parameters Tuning
Training & Comparison
Best ML model
Sensitivity Analysis
Factor Importance
Research Model
Survey Study
Preliminary Analysis
Machine Learning Analysis
Sensitivity Analysis

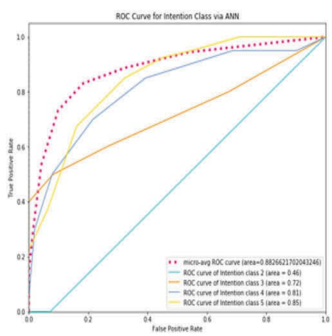

Perceive Benefits
Perceived Usefulness
Perceived Ease of Use
Perceived Discount
Perceive Loss
Financial Risk
Security Risk
Privacy Risk
Intention of Use
Research Model

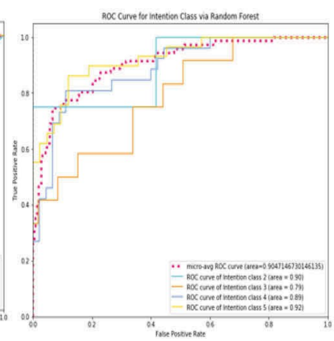

IU
Repondents
Class
5.0
4.0
3.0
2.0


ROC Curve for Intention Class via KNN
True Positive Rate
False Positive Rate
micro-avg ROC curve (area=0.8196455729683263)
ROC curve of Intention class 2 (area = 0.95)
ROC curve of Intention class 3 (area = 0.72)
ROC curve of Intention class 4 (area = 0.70)
ROC curve of Intention class 5 (area = 0.75)


ROC Curve for Intention Class via SVM
True Positive Rate
False Positive Rate
micro-avg ROC curve (area=0.8707928321100311)
ROC curve of Intention class 2 (area = 0.93)
ROC curve of Intention class 3 (area = 0.78)
ROC curve of Intention class 4 (area = 0.76)
ROC curve of Intention class 5 (area = 0.92)


ROC Curve for Intention Class via Gradient Boosting
True Positive Rate
False Positive Rate
micro-avg ROC curve (area=0.9077233353170667)
ROC curve of Intention class 2 (area = 0.91)
ROC curve of Intention class 3 (area = 0.81)
ROC curve of Intention class 4 (area = 0.92)
ROC curve of Intention class 5 (area = 0.92)


ROC Curve for Intention Class via ANN
True Positive Rate
False Positive Rate
micro-avg ROC curve (area=0.8826621702043246)
ROC curve of Intention class 2 (area = 0.46)
ROC curve of Intention class 3 (area = 0.72)
ROC curve of Intention class 4 (area = 0.81)
ROC curve of Intention class 5 (area = 0.85)


ROC Curve for Intention Class via Random Forest
True Positive Rate
False Positive Rate
micro-avg ROC curve (area=0.9047146730146135)
ROC curve of Intention class 2 (area = 0.90)
ROC curve of Intention class 3 (area = 0.79)
ROC curve of Intention class 4 (area = 0.89)
ROC curve of Intention class 5 (area = 0.92)