mathematics

MDPI

Article

A Fuzzy Association Rules Mining Analysis of the Influencing Factors on the Failure of oBike in Taiwan

Shianghau Wu

School of Business, Macau University of Science and Technology, Macau 999079, China; shwu@must.edu.mo; Tel.: +853-8897-2399

Received: 9 September 2020; Accepted: 29 October 2020; Published: 31 October 2020



Featured Application: This study attempted to combine the text mining method and the fuzzy association rules mining method to explore the main points of web posts about the reasons of bike-sharing company's failure in Taiwan. This research can offer a new way of data mining research.

Abstract: The sharing economy has become an important issue in recent years. Many researchers have paid attention to its application around the world. The sharing of bikes, as one of the major applications of the sharing economy, has shown its advantage in the realm of environmental protection and low energy consumption. However, bike-sharing system has encountered problems in certain regions. This arouses the concern about the sustainable development of the bike-sharing system. This research focused on the failure case of oBike in Taiwan. This research used text mining and fuzzy association rules mining methods to evaluate Taiwan's public opinion about the oBike in order to verify the reasons for oBike's failure in Taiwan. This study also made a comparison between the bike-sharing system in Mainland China and Taiwan. The research results explored the factors of oBike's failure in Taiwan and showcased the problems of bike-sharing systems in different regions. The research results also offer useful information for bike-sharing companies and the authorities concerned in order to develop a sustainable bike-sharing system.

Keywords: bike-sharing system; public opinion; oBike

1. Introduction

The bike-sharing system is an important application of sharing economy, and it is also considered an effective method of reducing air pollution, greenhouse effects, and energy consumption. The bike-sharing system has become a non-polluting public travel means. Many cities have proposed the usage of bike-sharing system as a means of connection between residential areas and public transportation hubs [1]. The bike-sharing system as a transportation means can be traced back to 2008, and is referred to as the way to solve the last mile problem [2]. In China, Chile, Brazil, New Zealand, South Korea, Taiwan, and the United States, public bike-sharing system began to take shape and became an important form of infrastructure in major cities of these countries [3]. Many related policies were introduced in order to promote the bike-sharing system as an environmentally friendly way of urban transportation [4]. However, previous literature indicates that the barriers to the bike-sharing system include the convenience and safety concerns, mandatory helmet legislation, and lack of an immediate sign-up process [5].

The bike-sharing system includes the public bike system (PBS) and the commercial bike-sharing system. The public bike system (PBS) is a bike-sharing system wherein one can lend and return bicycles freely at any station located in the city [6].

The bike-sharing system in each city has its special characteristics due to regional differences. Since the research around the development of the bike-sharing system in East Asia is quite limited [7], this study aimed to analyze one of the commercial bike-sharing systems in Taiwan. Taiwan is a small island with an area of 36,193 square kilometers and over 23 million residents. According to the census data in 2018, 69.29% Taiwanese people resided in six major municipalities (Taipei City, New Taipei City, Taichung City, Tainan City, and Kaohsiung City) [8]. As a densely populated island, Taiwan, like other Asian countries and regions, has faced environmental problems. The bike-sharing system has become a better choice as a clean and energy-saving means of public transportation in the six Taiwanese major municipalities, inspiring many academic researchers to design an ideal system [9]. The criticism of the defects of bike-sharing system is the mess of illegal parking of sharing bikes on the public areas, the waste of metal and other materials from the deserted sharing bikes, and the poor maintenance and mobilization of sharing bikes [10]. Other reasons for the OfO bike and oBike's market failure included bike vandalism, insufficient shared bike locations, privacy concerns of leaking credit card information, and unavailability of mandatory safety helmets [11].

This study attempted to analyze the failure case of commercial bike-sharing company oBike in Taiwan. oBike is a Singapore-based company that uses the dockless bike-sharing system. oBike users can utilize the service via a mobile phone application system (mobile app). oBike, which claimed to offer a "first and last mile" transportation solution, attracted a reported 100,000 users in 2017 [12]. oBike expanded quickly in 40 cities across 24 countries, including Malaysia, Taiwan, South Korea, Thailand, Germany, Italy, United Kingdom, Australia, the Netherlands, France, and Switzerland [13]. oBike began its business in Taiwan in 2017 and quickly expanded its service in several Taiwanese counties and cities. However, oBike filed for insolvency and ceased operations in June of 2018. The company's subsidiary in Taiwan, Aozhi Network Technology Co., announced that it laid off all its employees in October of 2018, leaving thousands of bicycles scattered around Taipei City [14]. Since there are fewer related studies about the failure of oBike in Taiwan, this study intended to analyze the public opinion change about the operation of oBike in Taiwan in order to search for the time when oBike's business began to decline, concluding the problem of oBike's business model. In contrast to the previous literature that defined the public bike-sharing system's failure as low ridership and insufficient governmental subsidization [15], this study defined the failure of oBike as a malfunction due to its managerial problems. The originality of this study is the research of public opinion via the social website discussion boards. Due to lack of operation and financial data of the private company oBike, this study could only conduct analysis through social websites in order to comprehend the public opinion about oBike in order to search for the reasons for oBike's failure. The research limitation of this research is that the research results can only reflect some social website users' opinions. It should also adjoin other related information such as news information data or statistical data in order to make a thorough assessment.

Previous research of the oBike system were restrained by the limited data availability. In order to fill with the gap of the literature, this study intended to analyze the public opinion related to oBike on web discussion boards in order to comprehend the following research questions:

- What kinds of factors caused oBike's failure according to public opinion on the web?
- Does public opinion on the web envisage the failure of oBike and offer any advice on the bike-sharing system?

The structure of this article is as follows: Section 2 contains the literature review of the previous research about oBike, Section 3 indicates the methodology, Section 4 is the discussion of the results, and Section 5 contains the conclusion.

2. Literature Review

There is little previous research about the process of oBike's business booming and wane. The oBike's business expansion in Vienna was also an example of quick expansion and decay.

oBike's bike-sharing system in Vienna was like that of OfO bike, wherein users downloaded the oBike app and registered with their credit cards. In 2018, the oBikes in Vienna encountered vandalism and the regulatory challenge of the free-floating bike-sharing. The Vienna city government announced its operation as illegal after 1st August 2018. Due to this regulation, oBike terminated its service in Vienna [16]. oBike in Singapore also faced insufficient parking lot problems. Singapore's Land Transport Authority established the "Parking Places Amendment Bill" to offer parking lots for sharing bikes operated by legal bike-sharing companies. The Amendment Bill was passed by the Singapore Parliament [17]. oBike in Australia also had the same development process. oBike began its operation in Sydney and Melbourne in 2017. Then, bike vandalism took place, and the Victoria Environment Protection Authority threatened a large number of fines if oBike did not clean up the damaged bikes. It resulted in oBike's termination of operation in June 2018 [18].

Previous research about oBike's development process in Taiwan is also limited. Hsu et al. (2018) indicate that oBike's business faced criticism owing to lack of comfort and functionality. oBike users were unable to change gears to adapt to different conditions of terrain. The only advantages of oBike were its lower setup cost and lower rental price [19].

oBike's business in Taiwan faced the challenge of its competitor "Ubike" (the official brand name is "Youbike"). Ubike was established in 2012 and has become Taipei's public bike-sharing system with more than 400 stations and 13,072 sharing bikes and 30 million users in 2019 [20]. Ubike intends to connect public transportation and the destination. Since Ubike was established several years before oBike, it has the advantage of network externalities. Farrell and Saloner (1986) indicate that when more people use the same product, that product will be more valuable [21]. This points out the existence of network externalities. Katz and Shapiro (1985) argue that network externalities would enhance consumers' expectations and the demand-side economics of scale [22]. Ubike has successfully created the a platform to meet the demand and supply of sharing bikes and has attracted consumer demand by means of network externalities in the same way as other sharing economy platforms [23]. In Taiwan, oBike faced the fierce competition from Ubike because Ubike had successfully enabled the brand attachment of the consumers, and thus it was difficult for oBike to attract consumers [24].

3. Methodology

3.1. Research Process

This study attempted to analyze the problem of oBike in Taiwan by exploring web posts and obtaining keywords. The data type of this research belongs to the category of Big Data. Big Data include the data extracted from social media; traditional databases of organizations; unstructured data from new communication technologies; and user editing platforms, such as text, images, and videos [25]. In order to analyze Big Data, researchers have applied new mathematical and statistical techniques in order to obtain the relationship among different variables [26].

This study then utilized text mining analysis to obtain the keyword frequency data. Text mining analysis is widely applied as the text classification method in several areas, such as healthcare and marketing [27]. Text mining analysis is also applied to the design of the early warning system for adverse drug reactions [28]. As far as public opinion analysis is concerned, text mining is used to facilitate knowledge discovery, reducing the cost as compared with the survey method [29]. With regard to the text mining application to business research, text mining can be used to analyze financial reports of companies in order to predict their future performance [30]. The text mining method includes text prepossessing and knowledge extraction. The text prepossessing process removes stop words, enables tokenization and stemming, and then converts unstructured text data into the document-term matrix [31]. The document-term matrix is constructed by the term weight. The presentation type of document-term matrix includes binary term (BT), term frequency (TF), and term frequency inverse document (TF-IDF) [32]. This study used the TF-IDF to obtain the frequencies of keywords in each

document because it is the most popular term weighting model. Let q be the term weighting scheme, the weight of each word (w) in the document (d) is calculated as follows [33]:

$$q(w) = f_d(w) * log\frac{|D|}{f_D(w)}$$

In the equation above, $|D|$ is the number of documents in the collection D.

As shown in Figure 1, the study analyzed social media web post text data. This study obtained major keywords through the text mining process and utilized the TF-IDF value of each keyword to put into the fuzzy association rules mining model in order to comprehend the relations of keywords in different time periods. The fuzzy association rule mining results would indicate different level thresholds (low, medium, high) of keyword frequency of each keyword in each month. The research process flowchart is presented in Figure 1.

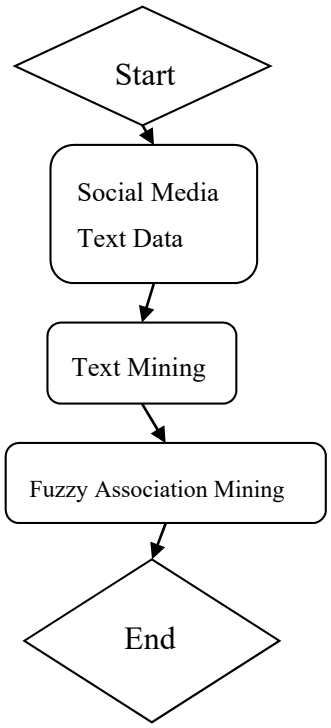

Figure 1. Flowchart of the research process.

3.2. Fuzzy Set Theory

Fuzzy set theory relates to the quantifying and reasoning process utilizing the natural language. In natural language, words can have ambiguous meanings. In the discrimination function of fuzzy set, the individuals from a universal set X are determined to be members or non-members of a crisp set. For a crisp set, this function assigns a value $\mu_A(x)$ to each $x \in X$ as [34]

$$\mu_A(x) = \begin{cases} 1 \; if \; and \; only \; if \; x \in A \\ 0 \; if \; and \; only \; if \; x \notin A \end{cases}$$

The function depicts the elements of the universal set including 0 and 1. It can be generalized such that the values assigned to the elements of the universal set fall within specified ranges.

3.3. Fuzzy Association Mining

Association rules indicate the dependent relations among items in the dataset [35]. The association rules are represented as $X \rightarrow Y$. X and Y are item sets and are independent. The association rules

imply that *X* appears in a condition that Y also appears with higher probability. In order to explore the associations in the real world, researchers apply fuzzy theory in order to avoid unnatural boundaries in partitioning, improving interpretability in language of the association rules [36]. Many methods have been developed in order to analyze fuzzy association rules from quantitative datasets [37]. Suppose there is a database with two attributes (B_1, B_2) and three linguistic levels (LOW, MEDIUM, and HIGH)—a possible mined fuzzy association rule would be (B_1 is LOW→B_2 is HIGH) [38].

Support and confidence are features that are commonly used to obtain association rules. Suppose we have the rule $X \rightarrow Y$, the two measurements of the XY relations, support and confidence, can be listed as follows [36]:

$$Support(R) = \frac{\sum_{X_p \in D} \mu_R\left(X_p\right)}{|D|} \tag{1}$$

$$Confidence(R) = \frac{\sum_{X_p \in D} \mu_R\left(X_p\right)}{\sum_{X_p \in D} \mu_X\left(X_p\right)} \tag{2}$$

In Equation (2), $\mu_X\left(X_p\right)$ is the matching degree of the rule pattern Xp with the rule antecedents. $\mu_R\left(X_p\right)$ is the matching degree of the rule pattern Xp with the rule antecedents and consequent. $|D|$ is the cardinality of the dataset. The fuzzy association rule mining also uses the measurement "Lift" to represent the ratio between the confidence ratio and the expected confidence ratio of each rule. Lift is defined as follows:

$$Lift(R) = \frac{Confidence(R)}{\sum_{X_p \in D} \mu_Y\left(X_p\right)/|D|} \tag{3}$$

In Equation (3), $\mu_Y\left(X_p\right)$. is the matching degree of the X_p with the rule consequent. The measure detects whether the items rule with negative dependence (*Lift* < 1), independence (*Lift* = 1), or positive dependence (*Lift* > 1).

The study used the RKEEL software package of *R* software [39]. The RKEEL software package applies three fuzzy rules algorithm: fuzzy a priori algorithm, genetic fuzzy a priori algorithm, and genetic fuzzy a priori DC algorithm.

3.3.1. Fuzzy A priori Algorithm

The fuzzy a prori method includes the following steps, in brief [40]:

Step 1: Transform the quantitative values *(vj(i))* of each transaction data *(D(i))* for each attribute *(Aj)* into a fuzzy set *(fj(i))* represented as $\sum_{m=1}^{l} \frac{f_{jm}^{(i)}}{R_{jm}}$, where R_{jm} is the n-th fuzzy region of the attribute *Aj*, *fjki* is *vj(i)*'s fuzzy membership value in the region R_{jk}, and *l*(=|*Aj*|) is the number of fuzzy regions of *Aj*.

Step 2: Calculate the number *(Z)* of each attribute region (linguistic term) R_{jk} in the transaction data, $Z_{jk} = \sum_{i=1}^{n} f_{jk}^{(i)}$.

Step 3: Collect each attribute region to form the candidate set Cl.

Step 4: Check whether the Z_{jk} of each R_{jk} is larger than or equal to the predefined minimum support value α. R_{jk} would be included in the set of 1 itemsets (L1).

Step 5: If L_1 is not null, follow the next step. If not, exit the algorithm.

Step 6: Set r = 1; r represents the number stayed into large itemsets. Then, join the large itemsets Lr to generate the candidate set C_{r+1} in a way similar to that in the a priori algorithm except that two regions (linguistic terms) belonging to the same attribute cannot simultaneously exist in an itemset in C_{r+1}.

Step 7: Calculate the fuzzy values of each transaction data in each newly formed (r + 1) itemset s, with itemsets (s1, s2, . . . , sr + 1) in C_{r+1}, and then calculate the count of s in the transactions. If the count of s is larger than or equal to the predetermined minimum support value α, put s into L_{r+1}.

Step 8: If L_{r+1} is null, then take the next step; otherwise, set r = r + 1 and repeat steps 6 and 7.

Step 9: Collect the itemsets.

Step 10: Construct association rules for each large q-itemset s with item (s1, s2, . . . , sq), q ≥ 2.

Step 11: List the output of the association rules with confidence values larger than or equal to the predefined threshold λ.

3.3.2. Genetic Fuzzy A priori Algorithm

The study used the genetic fuzzy a priori algorithm to search for the fuzzy association rules. The algorithm chooses the parent membership function with high fitness values. The evaluation function is then used to qualify the derived membership function sets. The performance of membership functions is then put into the genetic algorithm to promote the quality of the membership functions. This algorithm uses the overlap factor and coverage factor to obtain better derived membership functions.

The overlap factor of the membership functions for an item *Ij* in the chromosome C_q is defined as [41]

$$Overlap\ Factor\left(C_{qj}\right) = \sum_{k \neq i}\left[max\left(\left(\frac{overlap\left(R_{jk},\ R_{ji}\right)}{min\left(w_{jk},\ w_{ji}\right)}\right), 1\right), -1\right]$$

where overlap(R_{jk}, R_{ji}) is the overlap length of R_{jk} and R_{ji}. w_{jk} and w_{ji} represent half the spread of fuzzy regions R_{jk} and R_{ji}. If the overlap length is larger than the minimum span of the two functions, these membership functions are thought of as redundant.

The coverage factor of the membership functions for an item I_j in the chromosome C_q is defined as

$$CoverageFactor\left(C_{qj}\right) = \frac{1}{\frac{range\left(R_{ji},\ldots,R_{jl}\right)}{max\left(I_j\right)}}$$

In the equation above, (R_{j1}, . . . , R_{jl}) is the coverage range of the membership functions, l is the number of membership for I_j, and max(I_j) is the maximum quantity of I_j in the transactions.

In this algorithm, the suitability of the membership functions in a chromosome C_q is defined as

$$\sum_{j=1}^{m}\left[Overlap\ Factor\left(C_{qj}\right) + Coverage\ Factor\left(C_{qj}\right)\right]$$

In the above equation, m is the number of items. The fitness value of a chromosome C_q is defined as

$$f\left(C_q\right) = \frac{|L_1|}{suitability} \tag{4}$$

where $|L_1|$ is the number of large 1-itemsets obtained by using the set of the membership C_q.

The genetic fuzzy a priori algorithm utilizes the max–min arithmetical (MMA) crossover and the one-point mutation.

The brief genetic fuzzy algorithm is introduced as follows:

Step 1: Generate a random population of P individuals. Each individual is a set of membership functions for all m items.

Step 2: Encode each set of membership functions into a string representation.

Step 3: Calculate the fitness value of each chromosome.

Step 4: Execute the crossover operations on the whole population.

Step 5: Execute the mutation operations on the whole population.

Step 6: Using the selection criteria to choose individuals of the next generation.

Step 7: If the termination condition is not satisfied, go to step 3; otherwise, go to the next step.
Step 8: Obtain the set of memberships with the highest fitness value.

3.3.3. Genetic Fuzzy A priori DC Algorithm

The genetic fuzzy a priori DC algorithm also utilizes the max–min arithmetical (MMA) crossover and the one-point mutation and follows the same steps as the genetic fuzzy a priori algorithm.

The difference between the genetic fuzzy a priori DC algorithm and the genetic fuzzy algorithm is that the former uses the inequality condition of the central values of membership functions. Let R_{jk} represent the membership function of the k-th linguistic term for an item I_j. c_{jkp} denotes the p-th parameter of the fuzzy region R_{jk}. The inequality condition of the center values of the membership functions is $c_{j12} \leqq c_{j22} \leqq, \ldots, \leqq c_{jl2}$. In each membership function, the inequality condition of the three parameters is $c_{jk1} < c_{jk2} < c_{jk3}$.

In the genetic fuzzy a priori DC algorithm, the overlap ratio of the two membership functions R_{jk} and R_{ji} ($k < i$) is defined as the overlap length divided by the minimum of the right span of R_{jk} and the left span of R_{ji}. The overlap ratio is as follows [42]:

$$overlap\ ratio\left(R_{jk},\ R_{ji}\right) = \frac{overlap\left(R_{jk},\ R_{ji}\right)}{min\left(c_{jk3} - c_{jk2},\ c_{ji2} - c_{ji}\right)}$$

where overlap(R_{jk}, R_{ji}) is the overlap length of R_{jk} and R_{ji}. If the overlap length is larger than the minimum of the above mentioned two spans, the two membership functions are taken as redundant. We can use the overlap ratio to induce the overlap factor of the membership functions for an item I_j in the chromosome C_q as

$$overlap\ factor\left(C_q\right) = \sum_{k<i}\left[max\left(\left(\frac{overlap\left(R_{jk},\ R_{ji}\right)}{min(c_{jk3} - c_{jk2},\ c_{ji2} - c_{ji})}\right), 1\right) - 1\right]$$

The calculation of the coverage factor (C_q) is the same as the genetic fuzzy a priori algorithm. The larger the coverage ratio, the better the derived membership functions are.

The suitability of the set of membership functions in a chromosome is defined as

$$suitability\left(C_q\right) = overlap factor\left(C_q\right) + coverage factor\left(C_q\right)$$

The fitness value of a chromosome C_q is defined as

$$f\left(C_q\right) = \frac{\sum_{X \in L_1} fuzzy\ support(X)}{suitability\left(C_q\right)}$$

In the fitness value function, the fuzzy support (X) means the fuzzy support of the large 1-itemset X from the given database. The fuzzy support is defined as

$$fuzzy\ support(X) = \frac{\sum_{i=1}^{n} f_{jk}^{(i)}}{n}$$

3.4. Data Description

The study assorted the oBike-related web post data via the web crawler method from Taiwan's famous web discussion board "Mobile01" (http://mobile01.com). The discussion board focuses on the discussion of mobile phones, mobile appliance systems, and lifestyle. The research time frame ranged from March 2017 to June 2019. The study then integrated the web post text data into the monthly data. The number of web posts during the research time frame is shown in Figure 2.

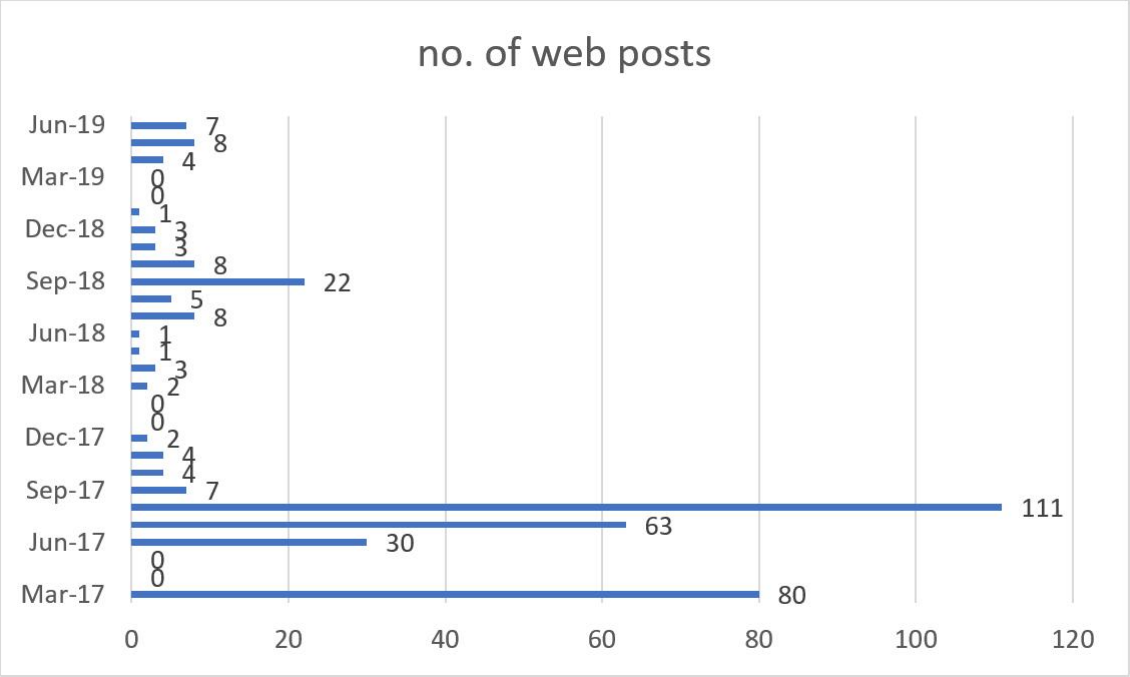

Figure 2. Number of oBike-related web posts.

4. Results

4.1. Text Mining Results

In the beginning, the study analyzed the Chinese text on the social media web discussion boards. After this study obtained the Chinese text, it was necessary to segment the Chinese texts in order to extract the important information. This study utilized the commonly used word segmentation tool "Jieba" on python to obtain the Chinese word TF-IDF data [43]. The study acquired nine keywords from Taiwan's "Mobile01" website posts, including "oBike", "share", "deposits", "government", "deposits", "management", "Ubike", "service provider", "parking", and "user". The nine keywords and the corresponding Chinese text or explanation are listed in Table 1.

Table 1. Major keywords of "Mobile01" web posts.

Keywords	Chinese Word or Explanation
oBike	*a brand name of bike-sharing company*
share	*gongxiang*
deposits	*yajin*
government	*zhengfu*
management	*guanli*
Ubike	*a brand name of public bike-sharing company*
service provider	*yezhe*
parking	*tingche*
user	*shiyongzhe*

The study obtained major keyword frequency data (TF-IDF) according to the text mining process. For each TF-IDF keyword data, we could observe the keyword frequency in each document. The descriptive statistics of major keywords data of oBike related "Mobile01" web posts is shown in Table 2. In Table 2, we can observe the distribution of keywords frequencies in brief.

Table 2. Descriptive statistics of major keywords term frequency inverse document (TF-IDF) data.

Keywords	Maximum	Minimum	Average	Standard Error
oBike	0.834	0	0.337	0.289
share	1.182	0	0.419	0.465
deposits	0.941	0	0.127	0.268
government	0.907	0	0.119	0.230
management	0.89	0	0.154	0.245
Ubike	0.498	0	0.093	0.124
service provider	1.78	0	0.279	0.471
parking	1.09	0	0.349	0.400
user	1.366	0	0.226	0.393

4.2. Fuzzy Association Rules Mining Results

The study used the RKEEL software package and applied the fuzzy a priori algorithm, genetic fuzzy a priori algorithm, and genetic fuzzy a priori DC algorithm, obtaining the average results of different measures, as shown in Table 3.

Table 3. Average results obtained by the three fuzzy algorithms.

Algorithm	#Rules	Support	Lift	Confidence	CF	Imbalance
Fuzzy a priori	6786	0.21	1.16	0.93	0.59	0.70
Genetic fuzzy a priori	165	0.13	2.35	0.92	0.86	0.60
Genetic fuzzy a priori DC	1013	0.14	1.39	0.91	0.73	0.49

According to Table 3, it can be observed that the fuzzy a priori algorithm generated the largest number of rules, and the genetic fuzzy a priori algorithm was the algorithm with the largest average lift, average confidence, and average imbalance value. Although the genetic fuzzy algorithm had the smallest number of rules, it generated better rules.

This study also compared the imbalance value of the three fuzzy algorithm results, as shown in Figure 3. The imbalance ratio calculated the degree of imbalance between two events that the LHS (left hand side, antecedent) and the RHS (right hand side, consequent) contained in a transaction. The imbalance ratio is zero if the conditional probabilities are similar (the antecedents and consequents are common event-sets). The imbalance ratio is close to 1 if the antecedents and consequents are very different [44]. The boxplots of the imbalance ratios showed that the rules generated by the fuzzy a priori algorithm had a larger imbalance ratio but with a larger gap among rules. However, the imbalance ratio of the rules generated by the genetic fuzzy a priori algorithm were evenly distributed in comparison with that of the other two fuzzy algorithms. Although there are three outlier points of the imbalance ratios of the genetic fuzzy a priori algorithm, these outlier points had lower lift ratios and did not affect the research results. The fuzzy association mining results selected for analysis were the RHS and LHS with the highest lift ratio. In order to determine the actual time and the LHS when the RHS (keyword "user") had the highest keyword frequency, this study combined the top five fuzzy association rules to produce the following analysis.

Genetic Fuzzy Apriori Algorithm Results

The study found the genetic fuzzy a priori algorithm generated better results and listed the top 18 association rules in which their lift ratios were larger than 3. The results were shown as the parallel coordinates plot in Figure 4.

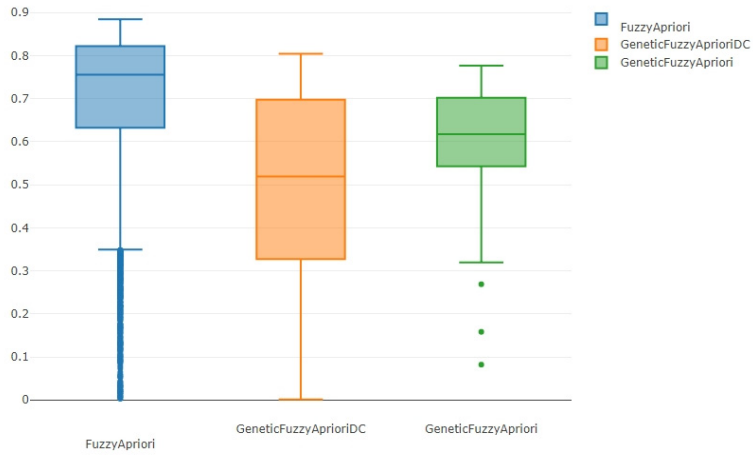

Figure 3. Boxplot of imbalance measure.

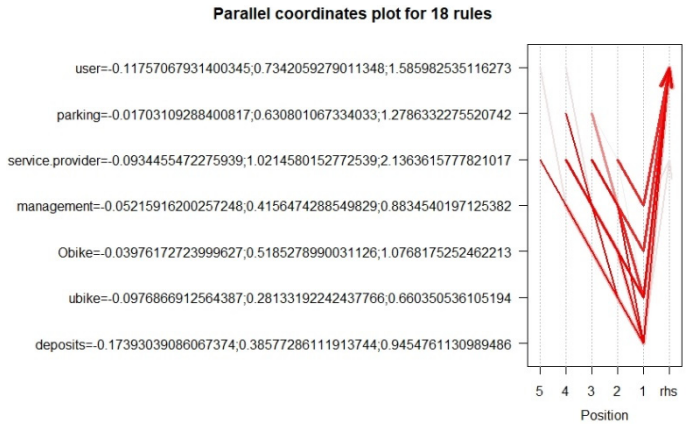

Figure 4. Parallel coordinates plot.

The top five rules are also listed in Table 4. As shown in Figure 4 and Table 4, this study observed that the top five rules had the same RHS (consequent) "user".

Table 4. Top five rules generated by the genetic fuzzy a priori algorithm.

LHS [1]	RHS [1]	Support	Confidence	Lift	CF	Imbalance
deposits = −0.17;0.38;0.94, management = −0.05;0.41;0.88, service provider = −0.09;1.02;2.13	user = −0.11;0.73;1.58	0.114	1.00	3.424	1.00	0.607
deposits = −0.17;0.38;0.94, Obike = −0.03;0.51;0.76, management = −0.05;0.41;0.88, service provider = −0.09;1.02;2.13	user = −0.11;0.73;1.58	0.111	1.00	3.424	1.00	0.617
Ubike = −0.09;0.28;0.66, Obike = −0.03;0.51;0.76, management = −0.05;0.41;0.88, parking = −0.01;0.63;1.27	user = −0.11;0.73;1.58	0.115	1.00	3.424	1.00	0.603
deposits = −0.17;0.38;0.94, Ubike = −0.09;0.28;0.66, management = −0.05;0.41;0.88, service provider = −0.09;1.02;2.13	user = −0.11;0.73;1.58	0.114	1.00	3.424	1.00	0.607
deposits = −0.17;0.38;0.94, Ubike = −0.09;0.28;0.66, Obike = −0.03;0.51;0.76, management = −0.05;0.41;0.88, service provider = −0.09;1.02;2.13	user = −0.11;0.73;1.58	0.111	1.00	3.424	1.00	0.617

[1] Values after two decimal places were omitted.

Figure 5 represents the keyword frequency data of the keyword "user" with the two cut-off values lines in Table 4. The gray line represents the upper cut-off value, and the orange line represents the lower cut-off value. Since the lowest cut-off values were negative in Table 4 and all keyword frequencies of the major keywords were positive, this study only analyzed the other two positive cut-off values of all LHS and RHS keywords. Because the fuzzy association mining results did not indicate the actual time of each rule, this study explored the keyword frequency and compared the two positive cut-off values of each keyword. This study observed that the keyword frequency of "user" belonged to the upper range (0.73–1.58) in September 2017, October 2017, and December 2017. It indicated that the consequent keyword "user" is the top issue on these months.

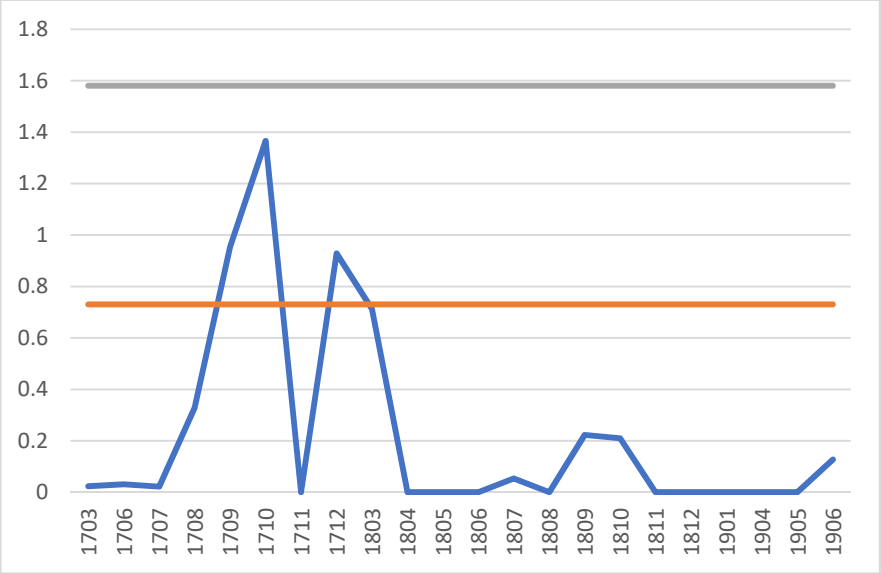

Figure 5. Keyword frequency of the consequent "users".

Figure 6 represents the keywords frequency of the antecedent "deposits". The gray line represents the upper cut-off value, and the orange line represents the lower cut-off value. It indicates that the keyword "deposits" belonged to the upper range (0.38–0.94) in March 2018, December 2018, and June 2019.

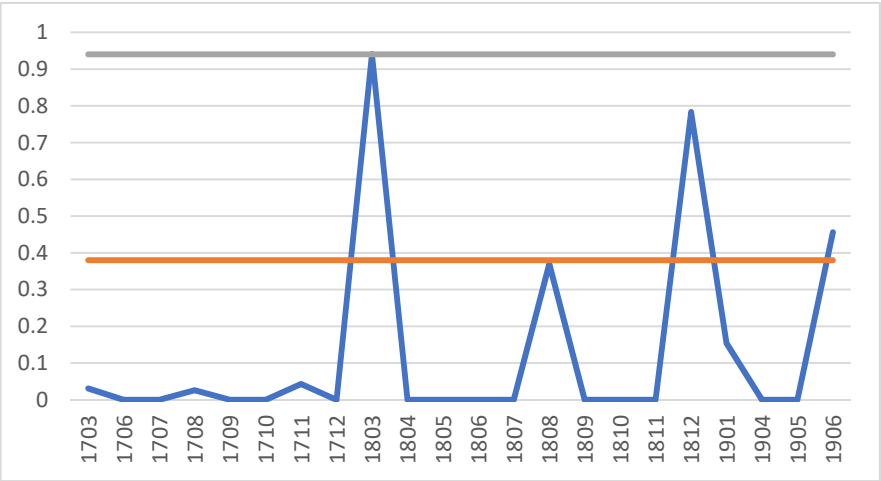

Figure 6. Keyword frequency of the antecedent "deposits".

Figure 7 represents the keyword frequency of the antecedent "management". The gray line represents the upper cut-off value, and the orange line represents the lower cut-off value. It indicates

that the keyword "management" belonged to the upper range (0.41–0.88) in September 2017, May 2018, and December 2018.

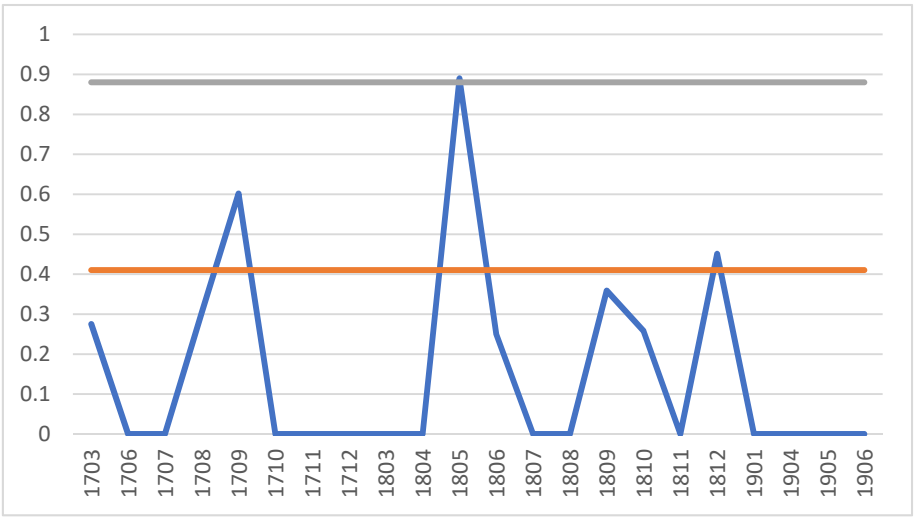

Figure 7. Keyword frequency of the antecedent "management".

Figure 8 represents the keyword frequency of the antecedent "service provider". The gray line represents the upper cut-off value, and the orange line represents the lower cut-off value. It indicates that the keyword "service provider" belonged to the upper range (1.02–2.13) in April and May 2018.

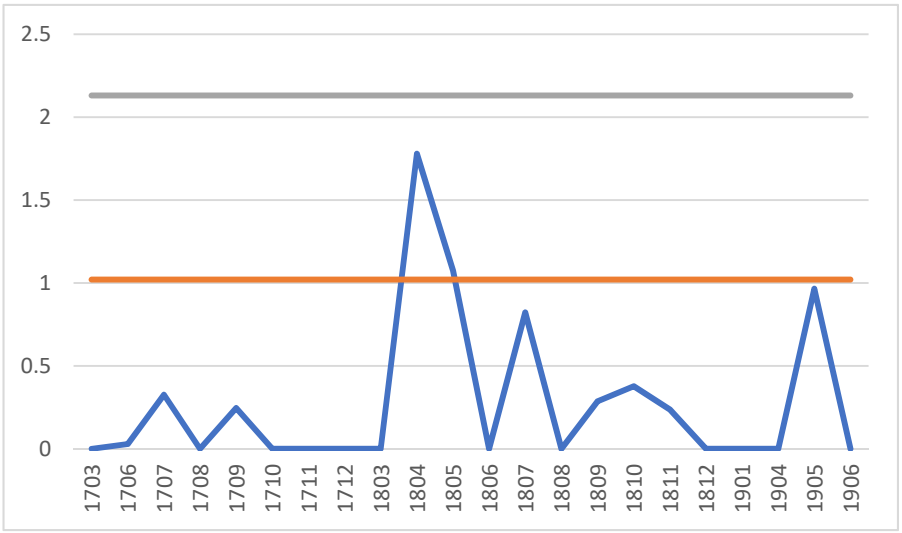

Figure 8. Keyword frequency of the antecedent "service provider".

Figure 9 represents the keyword frequency of the antecedent "Obike". The gray line represents the upper cut-off value, and the orange line represents the lower cut-off value. It indicates that the keyword "Obike" belongs to the upper range (0.51–0.76) in December 2017 and January 2019. However, the keyword frequency in June 2018, April 2019, and May 2019 were above the upper cutoff point (0.76).

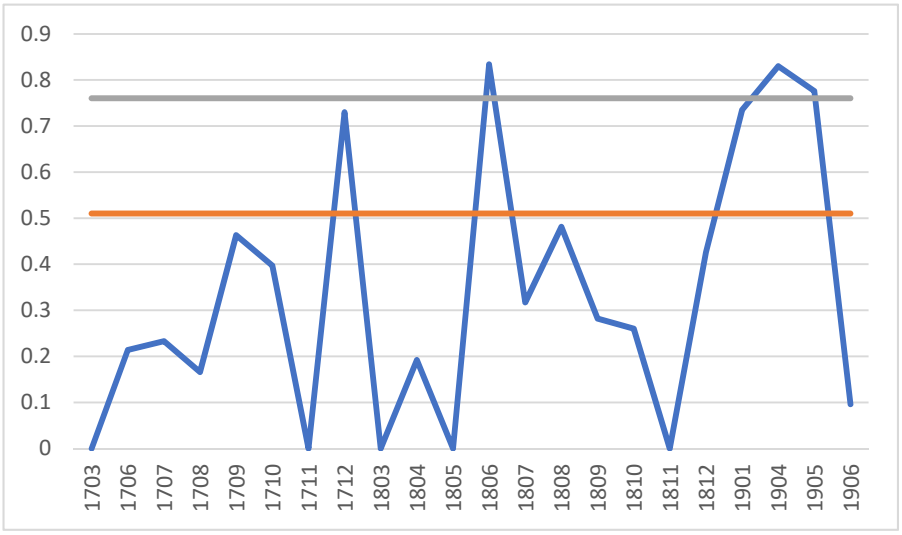

Figure 9. Keyword frequency of the antecedent "Obike".

Figure 10 represents the keywords frequency of the antecedent "Ubike". The gray line represents the upper cut-off value, and the orange line represents the lower cut-off value. It indicates that the keyword "Ubike" belonged to the upper range (0.28–0.66) in April 2019.

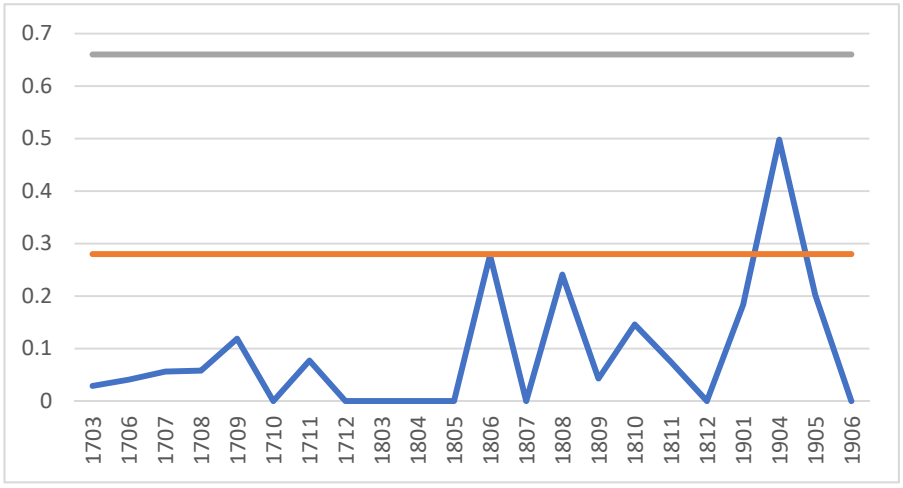

Figure 10. Keyword frequency of the antecedent "Ubike".

Figure 11 represents the keyword frequency of the antecedent "parking". The gray line represents the upper cut-off value, and the orange line represents the lower cut-off value. It indicates that the keyword "parking" belongs to the upper range (0.63–1.27) in June to September 2017, July 2018, and April 2019.

Table 5 indicates the time when the consequent "user" was in the higher range (0.73–1.58) of keyword frequencies. The consequent "user" was in the higher range in September 2017, October 2017, and December 2017. In order to explore the relationship among the antecedents and the consequents in the higher range of the top five fuzzy association rules, this study combined the antecedents of the association rules as part of the analysis. As far as the antecedent in Table 5 is concerned, the keywords "management" and "parking" were in the higher range of the keyword frequencies in September 2017, while the keyword "Obike" was in the higher range of keywords frequencies in December 2017. However, the keywords "deposits", "management", "service provider", "Obike", "Ubike", and "parking" were in the lower range of the keyword frequencies.

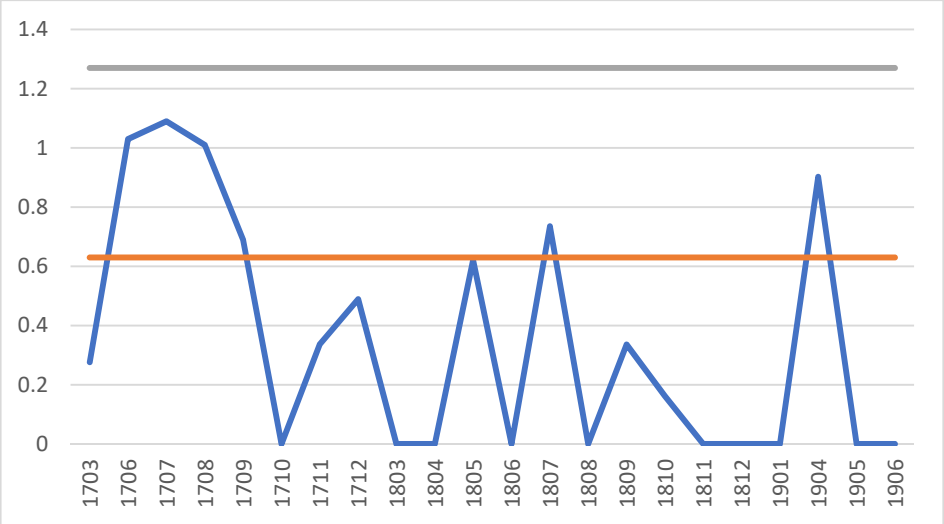

Figure 11. Keyword frequency of the antecedent "parking".

Table 5. The time when the consequent "user" was in the higher range of frequency.

Time	Antecedent [1]	Consequent [1]
Sep 2017	deposits(−), management(+), service provider(−), Obike(−), Ubike(−), parking(+)	user (+)
Oct 2017	deposits(−), management(−), service provider(−), Obike(−), Ubike(−), parking(−)	user (+)
Dec 2017	deposits(−), management(−), service provider(−), Obike(+), Ubike(−), parking(−)	user (+)

[1] (+) means the keywords were in the higher range of frequency; (−) means the keywords were in the lower range of frequency.

Table 5 shows the causal relationship between LHS and RHS in terms of higher keyword frequency. It indicates that the keywords "deposits", "management", "service provider", "Obike", "Ubike", and "parking" had a relationship with the larger frequency of "user". It implies that these keywords relate to public opinion about Obike users found on social websites.

5. Discussion

This study intended to explore the possible reason for oBike's failure in Taiwan. By means of the fuzzy association rules mining, this study concluded the major findings of this study as follows:

1. According to the fuzzy association rules mining results, the study found the genetic fuzzy a priori algorithm to have a better performance in terms of obtaining the fuzzy association rules with the higher average lift ratio. The study concluded the first 18 fuzzy association rules with the same consequent "user". This meant the oBike-related web posts involved the "user" concept.
2. The study also focused on the top five fuzzy association rules with the highest lift ratio (3.424), as shown in Table 4, as well as the antecedents of the top 5 rules. The antecedents of the top 5 fuzzy association rules included "deposits", "management", "service provider", "oBike", "Ubike", and "parking". It can be concluded that oBike-related web posts concerned the deposits, management, service provider, oBike, Ubike (a brand name of the public bike-sharing service provider) and parking, and that these keywords affected the frequency of the keyword "user". This indicated that public opinion concerned the users' viewpoints on the key elements of the oBike service: deposits, management, service provider, oBike's brand image, oBike's competitor "Ubike", and parking. Because the related articles on the "mobile 01" social website discussion boards were filled with negative evaluation of oBike, the text mining and fuzzy association

rules mining results reflect the users' experience and the defects of the oBike system, such as its management problem.

3. Since the fuzzy association mining rules results did not clearly indicate the actual time of the association rules, this study attempted to compare the keyword frequency and positive cut-off values, as shown in Table 4. This study chose the time points of the consequent "user" in the higher range of the keyword frequency. This study also combined the top five fuzzy association rules in order to obtain the results showcasing that the consequent "user" was in a higher range of frequency in September, October, and December 2017. As for the antecedent in these months, in September 2017, "management" and "parking" were also in the higher range of keyword frequency. This implies that the users' higher level of concern about management and parking was the main focus at that time. In October 2017, "deposits", "management", "service provider", "oBike", "Ubike", and "parking" were in the lower range of the keyword's frequencies. This means that none of these factors had a larger impact on the keyword frequency of "user". In December 2017, "oBike" was in the higher range of keyword frequency. This meant that web post users' focus on "oBike" had a positive impact on the keyword frequency of "user", and vice versa.
4. This study got the fuzzy association rules results of the oBike-related "Mobile01" web post keywords and obtained the influencing factors of oBike's failure as follows:

    - Management and parking problems: According to the combined top five fuzzy association rules analysis, the users' higher-level concern about management and parking was the main focus in September 2017. According to oBike's business history in Taiwan, as mentioned in the introduction section, oBike began its business in Taiwan in 2017. However, "Mobile01" web posts related to oBike pointed out the management and parking problems. As was mentioned in the introduction part, oBike's business expansion was fast in Taiwan, but its business soon terminated in 2018. It can be induced that the "Mobile01" web posts related to the oBike reflected the oBike's problem before it ceased operation.
    - Competitor's impact: The results indicate that web post comments related to Ubike positively affected the frequency of the keyword "user". This implies that the successful business model of Ubike influences users' evaluation and made it difficult for oBike to gain the market share in the bike-sharing market. The results also confirm the competitor's success.

5. In terms of previous research about the bike-sharing system problem in Mainland China, Wu and Lei (2019) concluded that the bike-sharing problem in Mainland China included the deposits, management, and sustainability problems [45]. However, the results of this study indicated the main problems of oBike's failure in Taiwan included management, parking, deposits, and the competitor's impact. This implies the business operation of oBike in Taiwan encountered the same problems: management, parking, and deposits. However, the fuzzy association rules mining results in this study point out the consequences of the top five rules are "users", and the competitor's impact is also the factor for oBike's failure in Taiwan. This means that oBike's failure in Taiwan related to the management problem and its negative impact on users.

6. Conclusions

This study used the text mining method and fuzzy association rules mining method to analyze the keyword relationships of web posts, finding that the main reasons of oBike's failure included management, parking, deposits, and the impact of competitors. These factors affect users' experience and the bike-sharing company's sustainable business operation. This indicates that the bike-sharing companies should be user-oriented in order to grasp the market share and stay in the market. The contribution of this study is that this research used social website discussion boards as the research target, concluded the reasons for the failure of "oBike", and filled the literature gap in terms of failed cases of commercial bike-sharing systems. This study offers us more insights into bike-sharing system

research. Further studies could focus on the means of improvement in terms of the current commercial and public bike-sharing system.

Author Contributions: The author is responsible for the whole contribution of this article. The author has read and agreed to the published version of the manuscript.

Funding: This research was funded by the Macau Institutions of Higher Learning Research Fund of Humanities and Social Science and Faculty Research Grant of Macau University of Science and Technology.

Conflicts of Interest: The author declares no conflict of interest. The funders had no role in the design of the study and in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Lu, C.C. Robust Multi-period Fleet Allocation Models for Bike-Sharing Systems. *Netw. Spat. Econ.* **2016**, *16*, 61–82. [CrossRef]
2. DeMaio, P. Bike-sharing: History, impacts, models of provision, and future. *J. Public Transp.* **2009**, *12*, 3. [CrossRef]
3. Mateo-Babiano, I.; Kumar, S.; Mejia, A. Bicycle sharing in Asia: A stakeholder perception and possible futures. *Transp. Res. Procedia* **2017**, *25*, 4966–4978. [CrossRef]
4. Pucher, J.; Buehler, R.; Seinen, M. Bicycling renaissance in North America? An update and re-appraisal of cycling trends and policies. *Transp. Res. Part A Policy Pract.* **2011**, *45*, 451–475. [CrossRef]
5. Fishman, E. Bikeshare: A Review of Recent Literature. *Transp. Rev.* **2016**, *36*, 92–113. [CrossRef]
6. Shaheen, S.; Guzman, S. *Worldwide Bikesharing*; UC Berkeley ACCESS Magazine: Berkeley, CA, USA, 2011; Volume 1.
7. Lu, M.; Hsu, S.C.; Chen, P.C.; Lee, W.Y. Improving the sustainability of integrated transportation system with bike-sharing: A spatial agent-based approach. *Sustain. Cities Soc.* **2018**, *41*, 44–51. [CrossRef]
8. Dept. of Household Registration, M. of the I.R. of C.Dept. of Household Registration Survey Data. Ministry of the Interior. Republic of China (Taiwan). *Dept. Househ. Regist. Minist. Inter. Repub. China (Taiwan)*, 2018.
9. Lin, Y.; Liang, F. Simulation for Balancing Bike-Sharing Systems. *Int. J. Modeling Optim.* **2017**, *7*, 7763.
10. Say, A.L.; Guo, R.; Chen, C. Disruption or new order?: The emergence of the unicorn bike-sharing entrepreneurship in China. In Proceedings of the 2018 Portland International Conference on Management of Engineering and Technology (PICMET), Honolulu, HI, USA, 19–23 August 2018; pp. 1–9.
11. Soltani, A.; Mátrai, T.; Camporeale, R.; Allan, A. Exploring shared-bike travel patterns using big data: Evidence in chicago and budapest. *Lect. Notes Geoinf. Cartogr.* **2019**, 53–68. [CrossRef]
12. Bike-Sharing May Change Our Cities. Available online: https://www.nst.com.my/cbt/2017/09/278219/bike-sharing-may-change-our-cities (accessed on 25 April 2020).
13. oBike Expands into Ten New Countries|BikeBiz. Available online: https://www.bikebiz.com/obike-expands-into-ten-new-countries/ (accessed on 26 April 2020).
14. oBike Apocalypse Ends as Taipei Scraps Last o . . . |Taiwan News. Available online: https://www.taiwannews.com.tw/en/news/3743822 (accessed on 25 April 2020).
15. Sun, F.; Chen, P.; Jiao, J. Promoting public bike-sharing: A lesson from the unsuccessful Pronto system. *Transp. Res. Part D Transp. Environ.* **2018**, *63*, 533–547. [CrossRef]
16. Laa, B.; Emberger, G. Bike sharing: Regulatory options for conflicting interests–Case study Vienna. *Transp. Policy* **2020**. [CrossRef]
17. Bill Tabled to License Bike-Sharing Firms to Curb Illegal Parking, Transport News & Top Stories-The Straits Times. Available online: https://www.straitstimes.com/singapore/transport/bill-tabled-to-license-bike-sharing-firms-to-curb-illegal-parking (accessed on 27 April 2020).
18. Heymes, C. Stationless in Sydney: The Rise and Decline of Bikesharing in Australia. *Transp. Find.* **2019**, *2018*, 1–8. [CrossRef]
19. Hsu, C.C.; Liou, J.J.H.; Lo, H.W.; Wang, Y.C. *Using a Hybrid Method for Evaluating and Improving the Service Quality of Public Bike-Sharing Systems*; Elsevier B.V.: Amsterdam, The Netherlands, 2018; Volume 202, ISBN 8862277121.

20. YouBike: Public Bikes. Available online: https://taipei.youbike.com.tw/news/list?5cb582c1060db454916c642a (accessed on 30 August 2020). (In Chinese).
21. Farrell, J.; Saloner, G. Installed Base, and Compatibility, and Innovation. *Am. Econ. Rev.* **1986**, *76*, 910–955.
22. Katz, M.L.; Shapiro, C. Network externalities, competition, and compatibility. *Am. Econ. Rev.* **1985**, *75*, 424–440.
23. Kathan, W.; Matzler, K.; Veider, V. The sharing economy: Your business model's friend or foe? *Bus. Horiz.* **2016**, *59*, 663–672. [CrossRef]
24. Chen, C.-F.; Chen, Y.-X.; Mu, P.-J.; Yu, Y.-H. Beyond adoption of shared bike: A case study on antecedents and consequences of brand attachment in Taiwan. *Res. Transp. Bus. Manag.* **2020**, 100574. [CrossRef]
25. Moro, S.; Rita, P.; Vala, B. Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach. *J. Bus. Res.* **2016**, *69*, 3341–3351. [CrossRef]
26. Chen, C.L.P.; Zhang, C.Y. Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Inf. Sci.* **2014**, *275*, 314–347. [CrossRef]
27. Lu, H.-M.; Zeng, D.; Trujillo, L.; Komatsu, K.; Chen, H. Ontology-enhanced automatic chief complaint classification for syndromic surveillance. *J. Biomed. Inform.* **2008**, *41*, 340–356. [CrossRef]
28. Yang, M.; Kiang, M.; Shang, W. Filtering big data from social media-Building an early warning system for adverse drug reactions. *J. Biomed. Inform.* **2015**, *54*, 230–240. [CrossRef]
29. Morinaga, S.; Yamanishi, K.; Tateishi, K.; Fukushima, T. Mining product reputations on the web. In Proceedings of the KDD'02: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton, AB, Canada, 23–25 July 2002; pp. 341–349.
30. Kloptchenko, A.; Eklund, T.; Karlsson, J.; Back, B.; Vanharanta, H.; Visa, A. Combining data and text mining techniques for analysing financial reports. *Intell. Syst. Account. Financ. Manag. Int. J.* **2004**, *12*, 29–41. [CrossRef]
31. Kumar, B.S.; Ravi, V. A survey of the applications of text mining in financial domain. *Knowl.-Based Syst.* **2016**, *114*, 128–147. [CrossRef]
32. Salton, G.; Buckley, C. Term-weighting approaches in automatic text retrieval. *Inf. Process. Manag.* **1988**, *24*, 513–523. [CrossRef]
33. Allahyari, M.; Pouriyeh, S.; Assefi, M.; Safaei, S.; Trippe, E.D.; Gutierrez, J.B.; Kochut, K. A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. *arXiv* **2017**, arXiv:1707.02919.
34. Zimmermann, H. Fuzzy set theory. *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 317–332. [CrossRef]
35. Srikant, R.; Agrawal, R. Mining generalized association rules. *Futur. Gener. Comput. Syst.* **1997**, *13*, 161–180. [CrossRef]
36. Pancho, D.P.; Alonso, J.M.; Alcalá-Fdez, J.; Magdalena, L. Analyzing fuzzy association rules with Fingrams in KEEL. In Proceedings of the 2014 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Beijing, China, 6–11 July 2014; pp. 2352–2359.
37. Kianmehr, K.; Kaya, M.; ElSheikh, A.M.; Jida, J.; Alhajj, R. Fuzzy association rule mining framework and its application to effective fuzzy associative classification. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2011**, *1*, 477–495. [CrossRef]
38. Segura, A.; Pérez-Pérez, R.; Alcalá-Fdez, J. New open source modules in KEEL to analyze and export fuzzy association rules. In Proceedings of the 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Vancouver, BC, Canada, 24–29 July 2016; pp. 225–232.
39. Moyano, J.; Sanchez, L. RKEEL: Using KEEL in R code. In Proceedings of the 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), Vancouver, BC, Canada, 24–29 July 2016; pp. 257–264.
40. Hong, T.P.; Kuo, C.S.; Chi, S.C. Trade-off between computation time and number of rules for fuzzy mining from quantitative data. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **2001**, *9*, 587–604. [CrossRef]
41. Hong, T.P.; Chen, C.H.; Wu, Y.L.; Lee, Y.C. A GA-based fuzzy mining approach to achieve a trade-off between number of rules and suitability of membership functions. *Soft Comput.* **2006**, *10*, 1091–1101. [CrossRef]
42. Hong, T.; Chen, C.; Lee, Y.; Wu, Y. Genetic-Fuzzy Data Mining With Divide-and-Conquer Strategy. *IEEE Trans. Evol. Comput.* **2008**, *12*, 252–265. [CrossRef]
43. Jieba · PyPI. Available online: https://pypi.org/project/jieba/ (accessed on 28 May 2020).

44. Wu, T.; Chen, Y.; Han, J. Re-examination of interestingness measures in pattern mining: A unified framework. *Data Min. Knowl. Discov.* **2010**, *21*, 371–397. [CrossRef]
45. Wu, S.; Lei, X. The Analysis of the Influencing Factors on the Problems of Bike-Sharing System in China. *IEEE Access* **2019**, *7*, 104000–104010. [CrossRef]

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).