

The investigation of accuracy level prediction of Fintech customers loyalty by using data mining algorithm decision tree

Mohammadreza Babakhanian^a, Seyyed Abdollah Amin Mousavi^b, Roya Soltani^c, Hamidreza Vakilifar^d

^aDepartment of Information Technology Management, Qeshm Branch, Islamic Azad University, Qeshm, Iran

^bDepartment of Management, Central Tehran Branch, Islamic Azad University, Tehran, Iran

^cDepartment of Industrial Engineering, Faculty of Engineering, Khatam University, Tehran, Iran

^dDepartment of Accounting, Tehran Science and Research Branch, Islamic Azad University, Tehran, Iran

(Communicated by Seyyed Mohammad Reza Hashemi)

Abstract

The emergence of mobile applications is forcing ambitious companies hoping to build loyalty for customers' brands to rush towards marketing their brand applications. The present research was conducted with the aim of classifying loyal customers and measuring their loyalty level using data mining algorithms. The present research method is based on applied-descriptive and the statistical population included the customers of Asan Pardakht Company which were considered number ten thousand people and with the number of 700,000 transactions. These customers were separated by clustering operation and classified for performing different tests. By using the data of Fintech customers of Asan Pardakht Company, it was attempted by using the decision tree algorithm, in addition, to identifying active customers, to implement this algorithm, a way is made in order to increase customer loyalty and ultimately increase their profitability and create satisfaction among managers. In the present research, by implementing the different stages of Crisp methodology, clustering and testing different artificial intelligence algorithms, the most useful algorithm in order to identify the best customers and also to make them loyal and policies and implementable programs to be formulated in order to increase the satisfaction percentage and finally customers' loyalty was explained and mentioned.

Keywords: Fintech, Customer loyalty, Decision tree, Data mining algorithms

2020 MSC: 65Y04, 68W40

1 Introduction

In the present era, where consumers spend more time in digital environments, companies prioritize being online to keep in touch with their customers anytime and anywhere [12]. The emergence of the Internet has changed the prioritization of consumers and the decision-making process by customers. The online opinions which is shared by others about their experience, is profoundly influence on consumer decisions [22]. Reliable internet connectivity and the massive increase of smartphone use are also involved in the change way of the consumers acquire goods and services. This issue has also led to a sharp increase in the availability of various products and services through the

Email addresses: m.babakhanian@yahoo.com (Mohammadreza Babakhanian), iranianman2010@yahoo.com (Seyyed Abdollah Amin Mousavi), r.soltani@khatam.ac.ir (Roya Soltani), vakilifard.phd@yahoo.com (Hamidreza Vakilifar)

online medium [34]. Therefore, relying on online data sources for service information has become more common during these days. Mobile application programs (“apps”) are software programs that are installed on a mobile device and often represent a brand’s identity [36, 3]. They have become important marketing drivers in shaping customers’ brand experience in e-commerce [9]. These apps have integrated the physical and digital world together [7]. Retailers’ mobile applications play an important role in this integration due to their market penetration and importance in mobile shopping. These apps allow users to communicate with retailers in an easy, fast and convenient way, both to access information and to perform transactions [32]. In addition, these apps can make users have more frequent purchases and spend more money for their purchases compared to non-app users [11]. Even though an increasing percentage of customers purchase through retailer apps, most users download and use only a few of them which mostly are belong to large companies [17]. Furthermore, users usually uninstall the app after accessing information or advertising or completing a purchase [15]. Therefore, it is valuable and important to classify and examine the effective factors in customer loyalty in order to maintain them. Until now, the researches performed on the use of mobile application programs has mainly focused on why consumers download them, intention to use the application, frequency number of use, and purchase intention. While few studies have analyzed customer post-adoption behaviors [23, 14]. Therefore, creating and applying a personalized approach in contact with digital customer is critical in maintaining customer relationships [13]. Previous studies of online consumer behavior have identified a range of factors that influence customer loyalty. These factors include customization, interaction, interface design, community, information, incentives, wide range of products, overall image [27], satisfaction and reputation [25], application quality [26], usefulness, ease of use [21] and perceived value [16] etc. In general, research on customer loyalty in the business-to-consumer field has a long tradition. In the battle between attracting new customers and retaining old customers, this idea that retaining customers is more cost-effective has been agreed. In this article, it is tried that using various data of customers, to exploit fintechs and the use of data and past experiences of customers, automatic performance and also the use of machine learning techniques were investigated and by clustering customers and examining different recruitment and retention policies (using machine learning technique) in relation to customer loyalty are acted on different clusters. Theoretically, it is believed that customers who meet their needs in using a Fintech supplier are loyal to it and recommend its use to others. In recent years, mobile payment has become one of the new services offered by the Bank Iran mobile system for customers. The multiple banking system, including electronic banking, which is one of its components of meter payment, has become highly competitive in Iran, so customer loyalty to such programs has very importance because loyal customers have always considered as assets of any business. The objective of this research is to classify loyal customers and measure their loyalty amount using data mining algorithm in mobile payment systems.

2 A review on research literature

Customer loyalty can be understood as favorable behavior towards a company, which is evidenced by a customer’s willingness to repeat business with a given retailer, preference for a particular brand, and word-of-mouth support [35]. Customer loyalty is not a new concept. It has been explored since the 1930s, and since then, many definitions of this variable have been proposed. According to the previous research in the field of customer loyalty, the authors have proposed to define this concept as a positive attitude of the customer towards the service (and its characteristics), that is, an attitude that increases the customer’s preference in the form of purchasing certain types of goods and services [33]. The development of customer loyalty, from Oliver’s (1999) conceptualization, has been considered as an evolutionary and incremental process that begins with logical and rational reasons (cognitive loyalty), then it changes to emotional attachment (affective loyalty), and finally behavioral attachment (Mutual loyalty and action) [6, 19]. In summary, previous research on customer loyalty has considered it as a hierarchical, incremental, and cognitive phenomenon. This approach has not accepted customer loyalty as a meaningful phenomenon in which consumers actively participate in assigning meaning to loyalty in different situations [19]. How to maintain profitable customers and increase customer loyalty, is one of the constant concerns of companies [19]. As it was stated in the introduction, many factors affect customer loyalty. In mobile commerce literature, several factors influence loyalty, such as convenience [30], ease of use, usefulness, enjoyment [9], personal relationships [31] and customer satisfaction [20] are prominent. In the present era, data mining is the most important technology for the effective use of large and huge data, and its importance has increased. In order to obtain results from the data provided by Asan Pardakht Company, a data mining methodology should be used. Data mining is the analysis and discovery of huge amounts of data that can be understood in order to discover valid, new and potentially useful patterns.

In an informal definition, data mining is a process, automatically, for extracting patterns that represent knowledge, which has been implicitly stored in huge databases, data warehouses, and other large information repositories. Data mining simultaneously benefits from several scientific fields such as: database technology, artificial intelligence,

machine learning, neural networks, statistics, pattern recognition, Knowledge-based systems, knowledge-acquisition, information retrieval, high-performance computing, and data visualization. Artificial intelligence and data mining is one of the elements of customer relationship management and can help companies' movement towards customer-centricity. Raw and primary data can be collected from various resources and then are entered into the management data warehouse through extraction, translation and calling processes. In the data preparation section, the data are removed from the warehouse and becomes a suitable format for data mining.

3 A review on research empirical background

Rehman and Ha [24] in a research stated that for success, focusing on service quality, customer trust and product quality leads to customer loyalty for Fintech Company compared to other traditional options for cross-border payments through the mediating effect of customer satisfaction. Tien and his colleagues [29], were investigated the role of advertising compatible with customer behavior and their loyalty by bank customers, as well as the role of new Fintech technologies as an intermediary between them. In this study, they used adaptive behavior scales, relationship marketing, and customer loyalty scales. The findings of the analysis of this model show that the mediating effect of Fintech marketing plays a role between customer behavior and their transaction loyalty. Mulyono and Pasaribu [18] in a research showed that mobile phone service quality and brand image have a positive effect with the intervening variable of customer satisfaction on customer loyalty. Lele and Shaw [10] have considered pleasure motivation, efficiency, perceived value, perceived privacy risks, system availability and fulfillment as important factors in customer loyalty to apps. Akhmedova et al. [1] investigated winning strategies for customer loyalty with the combined methods. They introduced organizing information stored in apps, responsiveness and platform reliability and customer interaction with the peer provider as important factors in customer loyalty. Tien et al. [29] investigated the role of advertisements compatible with customers' behavior and their loyalty by bank customers, as well as the role of new Fintech technologies as a mediator between behavior and loyalty. This study has used appropriate and adaptive behavior scales, relationship marketing, and customer loyalty scale. The findings of the analysis of this model show that the mediating effect of Fintech marketing plays an effective role between customer behavior and their transactions loyalty.

4 Research method

The present research is based on applied and descriptive method which will be studied by examining the relationships between variables and customers' loyalty and using data mining techniques to attract, retain, and customers' loyalty to Fintechs which the processes of research execution were studied in the form of diagram (1):

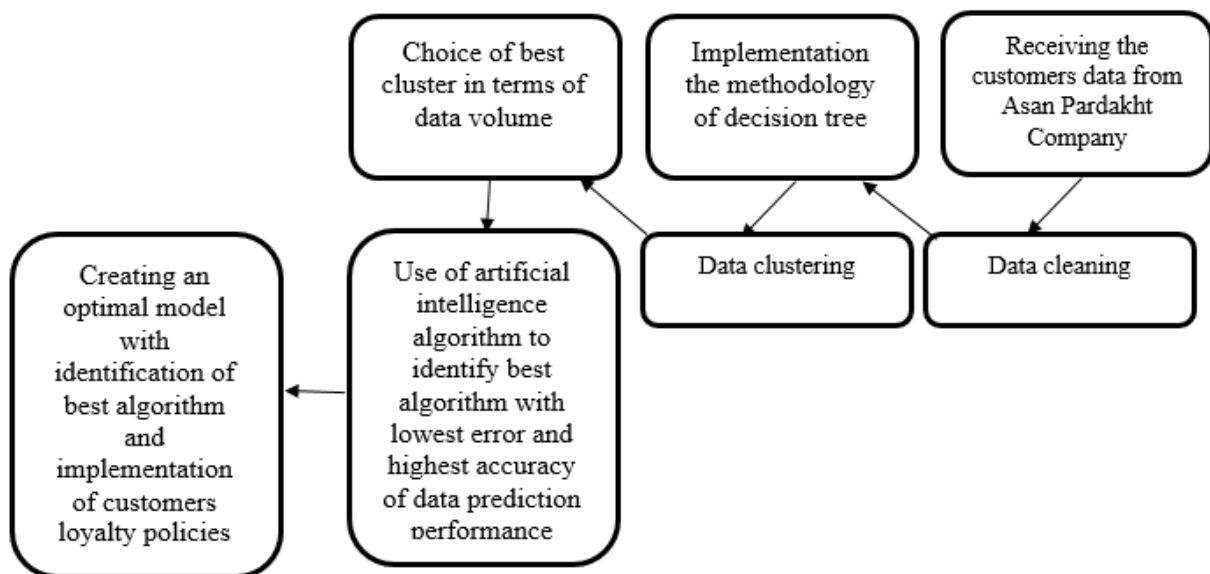


Figure 1: Research execution processes

Sample, Explore, Modify, Model, and Assess (SEMMA), CRISP, Knowledge Discovery in Database (KDD) are among the types of methodologies, which are used for data mining [2, 4].

Data mining: It is an essential stage of KDD process in which various statistical methods and machine learning are used to extract patterns [5]. Data mining is often considered as part of the process of “Knowledge Discovery in Database (KDD)”. Knowledge Discovery in Database (KDD) is a process that turns raw data into useful knowledge, which in addition to data mining, includes two stages of pre-processing and post-processing.

A) Pre-processing: The purpose of pre-processing is to convert raw data into a format which is appropriate for further analysis, this stage includes; 1) Data cleaning 2) Data integration 3) Data transformation 4) Data reduction 5) Imaging; In order to reduce the dimension, a reduced or compressed representation of the original data is obtained. 6) Data mining [8].

B) Post-processing: Post-processing includes all the operations that must be done to make it easier for analysts to understand the results of data mining [28].

The statistical population includes the customers of Asan Pardakht Company. The business environment of the app application is also Persian Asan Pardakht Company software. The research data obtained from the transactions of the customers of Asan Pardakht Company and interviews with the experts who are the managers of the company were collected and used. The tests used in this research have used the Python programming language in order to cluster and perform various data mining tests on the data. After the clustering operation, the data were analyzed using quantitative and qualitative methods. In the present research, the software used in the Python programming language was used to cluster and analyze data and measure customer loyalty to Fintech services. In addition, R software was applied to create the final model and the effects of artificial intelligence on increasing the efficiency of Fintechs.

Research findings

The application of artificial intelligence algorithms to predict and identify loyal customers In this section, by using of different algorithms of artificial intelligence, we are investigated different operations of predicting the amount of customer loyalty and their profitability in the system by applying two items of error rate and accuracy, which respectively the error rate with 8.09 and accuracy rate with 91.76% is transferred to the output. In the examination of this item, the data available in the largest cluster are divided into 80%, 20% of the test data and the result which using the decision tree algorithm, this operation was implemented on the data as above, which can be predict the level of loyalty and also the future purchases of the customer and based on the level of loyalty of each customer, define a plan for them to maintain and provide future offers as well as to increase influence in the market. It seems that the amount of prediction made in the purchase and the amount of customers' loyalty has a high error and its accuracy amount of that number has been shown in the output. A decision tree is a map of the possible outcomes of a series of related choices or options that allow an individual or organization to evaluate possible measures in terms of costs, probabilities, and benefits.

Calculating the error of our model with MAE (Mean Absolute Error)

```
In [92]: from sklearn.metrics import mean_absolute_error

        predicted_user_purchases = charge_purchase_model.predict(X)
        mean_absolute_error(y, predicted_user_purchases)

Out[92]:0.0
```

Figure 2: Calculating the error of our model with mean absolute error(MAE)

In the above part and sections of the data aggregated and evaluated by the decision tree function, it can be seen that the purchase forecast by primary customers has been transferred to the output in the order of 140, 9, 168, 23, 1, In order to investigate the accuracy and correctness of the output of the future prediction of the purchase by the decision tree algorithm, the evaluated output data by the decision tree is compared with the real data to determine the output error. The cross-validation operation in the following code fragment is performed using the Sclearn library in Python and entering the train- test-split.

As it can be observed, this time and by changing the test data, a very large error was created which shows the improper performance of this algorithm in predicting the level of customers' loyalty. Using two items, error rate and accuracy were investigated, respectively, the error rate with 8.09 and accuracy rate with 91.76% was transferred to the output.

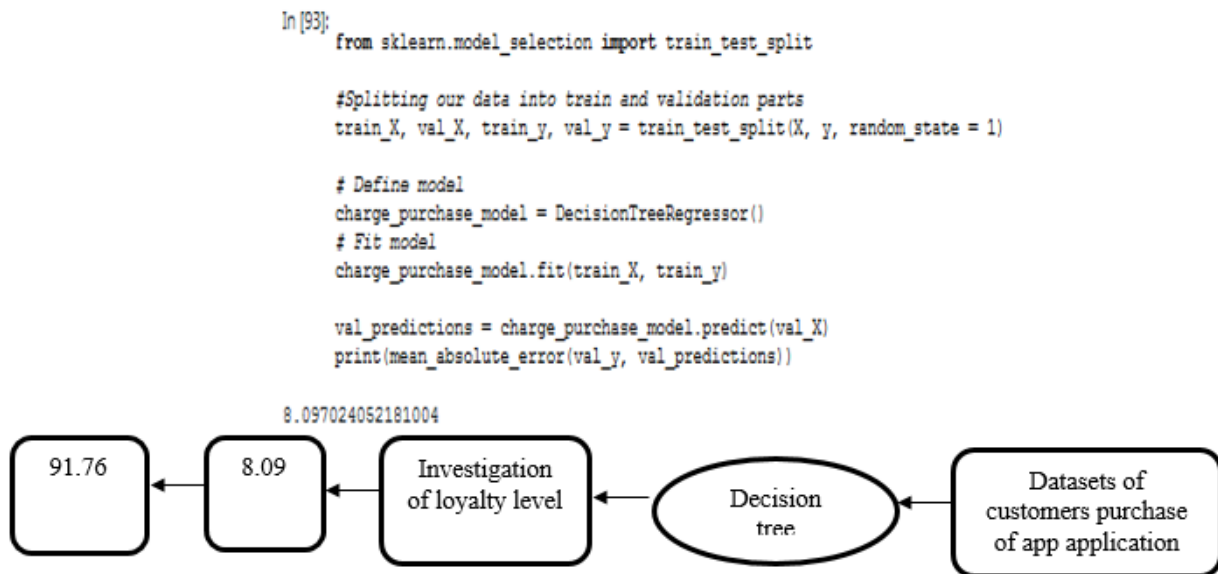


Figure 3: Prediction of customer loyalty and their profitability in the system using error rate and accuracy

5 Discussion and conclusion

The objective of this research was to classify loyal customers and measure their loyalty level by using data mining algorithm and decision tree. After analyzing the data as well as cleaning them in different stages and preparing the data, the data clustering operation was conducted by performing various tests on them and the largest cluster which was cluster number 4, was identified. By identifying the largest cluster, in fact, the most profitable cluster was identified. This category of customers, are among the customers who perform the highest purchases, and considering that the fee amount is based on each transaction, then the number of more customers is equal to the higher number of transactions and ultimately more profitability. In addition, by identifying some characteristics related to this category of customers, it is possible to predict providing the best possible service to customers, and in fact, and then to identify the future needs of the customer and his interests and were resolved in the best and fastest way. The use of different algorithms of artificial intelligence also has a significant effect in improving the process of identifying and meeting the needs of customers. In this research, the best algorithm due to the large number of transactions was the decision tree algorithm which by using this algorithm, it is even possible to identify the cluster of new customers and according to the amount of their future purchases, identify the type of customer in terms of loyalty. This algorithm has the lowest error in customer review and the highest accuracy in predicting customer performance. With this description, we were able to identify the best cluster of customers by using this algorithm and examine their loyalty by using different marketing policies. Also, by using this algorithm, it is possible to identify unfaithful or less interested customers to use and purchase through Fintech and by placing incentives, these customers are able to use of loyalty application. It seems that the existing findings, which using artificial intelligence algorithms were researched, have an effective and bold role in Fintechs smartization more than before. In case of identification of different categories of customers, in addition to update the database (for example, eliminating the deceased customers), it is possible to take a more effective action to revive and build customer loyalty, such as communication with abandoned customers, recognition of high transaction customers in many years, etc. for making customers loyal. Establishing a larger and more general model by using some specific characteristics of customers such as education level, occupation, age, etc., in order to create new clusters and provide unique services to different categories of customers by reviewing the level of accuracy and error rate in line with loyalty of customers.

References

- [1] A. Akhmedova, F. Marimon and M. Mas-Machuca, *Winning strategies for customer loyalty in the sharing economy: A mixed-methods study*, J. Bus. Res. **112** (2020), 33–44.
- [2] D. Anand and M. Mantrala, *Responding to disruptive business model innovations: The case of traditional banks facing fintech entrants*, J. Bank. Financ. Technol. **3** (2019), no. 1, 19–31.

- [3] S. Bellman, R.F. Potter, S. Treleaven-Hassard, J.A. Robinson and D. Varan, *The effectiveness of branded mobile phone apps*, J. Interact. Market. **25** (2011), no. 4, 191–200.
- [4] P. Cabena, P. Hadjinian, R. Stadler, J. Verhees and A. Zanasi, *Discovering data mining from concepts to implementation*, Prentice Hall, 1998.
- [5] U. Fayyad, G. Piatetsky-Shapiro, P. Smith and R. Uthurusamy, *Advances in knowledge discovery and data mining*, AAAI/MIT Press, 1996.
- [6] M. Fraering and M.S. Minor, *Beyond loyalty: customer satisfaction, loyalty, and fortitude*, J. Serv. Market. **27** (2013), no. 4, 334–344.
- [7] W. Gao, H. Fan, W. Li and H. Wang, *Crafting the customer experience in omnichannel contexts: The role of channel integration*, J. Bus. Res. **126** (2021), 12–22.
- [8] M. Ghazanfari, S. Alizadeh and B. Teymorpour, *Data mining and knowledge discovery*, Iran University of Science and Technology Publication, First edition, Tehran, 2012.
- [9] D.S. Kumar, K. Purani and S.A. Viswanathan, *Influences of ‘appscape’ on mobile app adoption and m-loyalty*, J. Retail. Consum. Serv. **45** (2018), 132–141.
- [10] A. Lele and N. Shaw, *Augmented reality: does it encourage customer loyalty?*, Int. Conf. Human-Comput. Interact., Springer, Cham., 2021, pp. 105–119.
- [11] H. Liu, L. Lobschat, P.C. Verhoef and H. Zhao, *App adoption: The effect on purchasing of customers who have used a mobile website previously*, J. Interact. Market. **47** (2019), 16–34.
- [12] W. Maroengsit, T. Piyakulpinyo, K. Phonyiam, S. Pongnumkul, P. Chaovalit and T. Theeramunkong, *A survey on evaluation methods for chatbots*, Proc. 7th Int. Conf. Inf. Educ. Technol., 2019, pp. 111–119.
- [13] G. McLean and A. Wilson, *Evolving the online customer experience. . . is there a role for online customer support?*, Comput. Human Behav. **60** (2016), 602–610.
- [14] G. McLean, K. Osei-Frimpong, K. Al-Nabhani and H. Marriott, *Examining consumer attitudes towards retailers’ m-commerce mobile applications—An initial adoption vs. continuous use perspective*, J. Bus. Res. **106** (2020), 139–157.
- [15] S. Molinillo, R. Aguilar-Illescas, R. Anaya-Sánchez and E. Carvajal-Trujillo, *The customer retail app experience: Implications for customer loyalty*, J. Retail. Consum. Serv. **65** (2022).
- [16] S. Molinillo, R. Aguilar-Illescas, R. Anaya-Sánchez and F. Liébana-Cabanillas, *Social commerce website design, perceived value and loyalty behavior intentions: the moderating roles of gender, age and frequency of use*, J. Retail. Consum. Serv. **63** (2021).
- [17] S. Molinillo, A. Navarro-García, R. Anaya-Sánchez and A. Japutra, *The impact of affective and cognitive app experiences on loyalty towards retailers*, J. Retail. Consum. Serv. **54** (2020).
- [18] R.A. Mulyono and L.H. Pasaribu, *The impact of mobile service quality and brand image on customer loyalty*, Enrichment: J. Manag. **12** (2021), no. 1, 1–7.
- [19] E. Närvänen, H. Kuusela, H. Paavola and N. Sirola, *A meaning-based framework for customer loyalty*, Int. J. Retail Distrib. Manag. **48** (2020), no. 8, 825–843.
- [20] S. Omar, K. Mohsen, G. Tsimonis, A. Oozeerally and J.H. Hsu, *M-commerce: The nexus between mobile shopping service quality and loyalty*, J. Retail. Consum. Serv. **60** (2021), 102468.
- [21] K. Purani, D.S. Kumar and S. Sahadev, *E-Loyalty among millennials: Personal characteristics and social influences*, J. Retail. Consum. Serv. **48** (2019), 215–223.
- [22] I. Ranggadara, G. Wang and E.R. Kaburuan, *Applying Customer Loyalty Classification with RFM and Naïve Bayes for Better Decision Making*, Int. Seminar Appl. Technol. Inf. Commun. (iSemantic), 2019, pp. 564–568.
- [23] B.T. Ratchford, *The history of academic research in marketing and its implications for the future*, Spanish J. Market.-ESIC. **24** (2020), no. 1, 3–36.
- [24] U. Rehman and K.S. Ha, *Factors influencing Fintech’s customer loyalty for cross border payments: Mediating*

- customer satisfaction*, Asia-Pacific J. Bus. Ventur. Entrepreneur. **16** (2021), no. 6, 287–297.
- [25] P.G. Rodríguez, R. Villarreal, P.C. Valiño and S. Blozis, *A PLS-SEM approach to understanding E-SQ, e-satisfaction and e-loyalty for fashion e-retailers in Spain*, J. Retail. Consumer Serv. **57** (2020), 102201.
- [26] S.K. Roy, W.M. Lassar and G.T. Butaney, *The mediating impact of stickiness and loyalty on word-of-mouth promotion of retail websites: A consumer perspective*, Eur. J. Market. **48** (2014), no. 9/10, 1828–1849.
- [27] S.S. Srinivasan, R. Anderson and K. Ponnawolu, *Customer loyalty in e-commerce: an exploration of its antecedents and consequences*, J. Retail. **78** (2002), no. 1, 41–50.
- [28] P.N. Tan, M. Steinbach and V. Kumar, *Introduction to data mining*, Pearson Education India, 2016.
- [29] C.T. Tien, H.K. Cheng and S. Pei-Ling, *The mediated effect of relationship marketing on the influences of irritation advertising in fintech times*, Proc. 2nd Int. Conf. E-Educ. E-Bus. E-Technol., 2018, pp. 99–101.
- [30] R. Thakur, *Understanding customer engagement and loyalty: a case of mobile devices for shopping*, J. Retail. Consumer Serv. **32** (2016), 151–163.
- [31] T.H. Tseng and C.T. Lee, *Facilitation of consumer loyalty toward branded applications: The dual-route perspective*, Telemat. Inf. **35** (2018), no. 5, 1297–1309.
- [32] G. Van Noort and E.A. Van Reijmersdal, *Branded apps: Explaining effects of brands' mobile phone applications on brand responses*, J. Interact. Market. **45** (2019), 16–26.
- [33] N. Vilkaite-Vaitone and I. Skackauskiene, *Service customer loyalty: An evaluation based on loyalty factors*, Sustainability **12** (2020), no. 6.
- [34] J. Xu, I. Benbasat and R.T. Cenfetelli, *Integrating service quality with system and information quality: an empirical test in the e-service context*, MIS Quart. **37** (2013), no. 3, 777–794.
- [35] V.A. Zeithaml, L.L. Berry and A. Parasuraman, *The behavioral consequences of service quality*, J. Market. **60** (1996), no. 2, 31–46.
- [36] Z. Zhao and C. Balagué, *Designing branded mobile apps: Fundamentals and recommendations*, Bus. Horizons **58** (2015), no. 3, 305–315.