IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 24, Issue 2, Ser. IV (Mar. –Apr. 2022), PP 06-10
www.iosrjournals.org

Design of a Novel Customer Satisfaction Model using Segmentation and Machine Learning

Priyanka Meena[1], Chetan Kumar[2]
[1](Computer Science, Kautilya Institute of Technology & Engineering, Jaipur, India)
[2](Computer Science, Kautilya Institute of Technology & Engineering, Jaipur, India)

Abstract: *The proposed system's system design and procedural details are described in this paper. The preliminary and comprehensive design of the system is used to define its workflow. The thorough design explains how the system's components and sub-components interact. The system's procedural details define the whole working procedure, from data collection through analysis and estimation, in step-by-step order.*
Background*: This paper delves into the K-Means and K-Prototype clustering algorithms for consumer behavioural and segmentation models. It comprises the segmentation's theoretical foundation as well as existing E-SAT-based systems. It allows for the statistical analysis of their data properties to be compared. The study went into great detail into the theoretical and algorithmic features of recent customer behavioural and segmentation approaches for E-satisfaction, particularly employing K-Prototype clustering. It depicts the path taken by such systems.*
Methods*: The proposed system's system design and procedural details are described in this paper. The preliminary and comprehensive design of the system is used to define its workflow. The thorough design explains how the system's components and sub-components interact. The system's procedural details define the whole working procedure, from data collection through analysis and estimation, in step-by-step order*
Results*: For automotive related datasets, the proposed system was implemented. Several experiments were carried out. The findings were found using the DBI, and the best values were those that were closest to zero. The datasets yielded four and eight clusters, respectively. It also displayed the results by area to show the different item forms in order.*
Conclusion: *The proposed customer behavior model analyses customers' buying choices in CAR-related datasets to assess their satisfaction and conduct. Through the usage of impacted essential factors, it used K-Prototype to measure user behavior.*
Key Word*: Customer Behavior; K-Prototype; Satisfaction, Customer Segmentation; Machine Learning.*

Date of Submission: 08-04-2022 Date of Acceptance: 25-04-2022

I. Introduction

Responses to opinions, attitudes, and perceptions are used to analyze and segment customer behavior. They provide insight about the sort of individual the customer appears to be. A person is provided with several scenarios; they may not have encountered them previously, but it just evaluates their attitude. Segmentation is a method of dividing customers into different groups based on a variety of characteristics. It is frequently used in marketing to analyze and use data and devise effective activities using K-prototype clustering technique.
The proposed system's system design and procedural details are described in this paper. The preliminary and comprehensive design of the system is used to define its workflow. The thorough design explains how the system's components and sub-components interact. The system's procedural details define the whole working procedure, from data collection through analysis and estimation, in step-by-step order.

II. Literature Review

This paper delves into the K-Means and K-Prototype clustering algorithms for consumer behavioral and segmentation models. It comprises the segmentation's theoretical foundation as well as existing E-SAT-based systems. It allows for the statistical analysis of their data properties to be compared. The study went into great detail into the theoretical and algorithmic features of recent customer behavioral and segmentation approaches for E-satisfaction, particularly employing K-Prototype clustering. It depicts the path taken by such systems.

The analysis fills the void by examining social media and review websites as viable sources of consumer satisfaction data for retail stores [1]. This study's findings reveal the best various means for gathering

customer satisfaction data in retail. Furthermore, utilizing the web-scraping approach, this study devised a way for collecting a significant number of customer happiness data.

The following conclusions are drawn from [2]. The customer's perception of risk is influenced by the quality of the website on the buying platform. Customer happiness is influenced by the quality of an online shopping website. The level of service provided on a shopping website has a negative effect on risk perception. Customer happiness is influenced by the level of service provided on the shopping website platform. Customer repurchase intention is negatively impacted by the perceived risk on the shopping website. Satisfied customer with the shopping site has a positive impact on the likelihood of repurchase.

The causes of customer churn in 4S stores are explored in [3], and numerous solutions are suggested to improve the operation and management of 4S stores. The findings shed insight on customer loyalty and marketing in the car sector. Through data mining methods, [4] examines the effects and reasons of customer churn and provides answers to topics such as how customer churn occurs, the influencing elements of customer churn, and how businesses reclaim churned consumers. The findings of [4] can better serve the practice of relationship management in the telecom industry and serve as a reference for the industry to identify high-risk churned customers in advance, improve customer loyalty and viscosity, maintain "high-value" customers, and keep going to provide consumers with "value" while lowering the cost of preserving clients.

The suggested method used a probabilistic approach to examine changes in customer satisfaction as a function of attribute value variations [5] [6]. They used this data to uncover the link between a feature and customer satisfaction, allowing us to classify the attribute. To assess the efficacy of the suggested method, they looked at artificial and real-world housing datasets. Both dataset's characteristics were appropriately classified using the algorithm. To demonstrate the superiority of the proposed technique, they compared the results to those obtained using the present method. The findings also imply that the suggested method was able of correctly capturing client behavior [5].

[7] explored relevant data sources that allegedly influence consumer satisfaction levels based on existing methodologies to model client satisfaction. A hypothesis-driven approach and a machine learning approach were compared. A retrieved dataset from a real case was used to evaluate the approaches. Next approach examined and evaluated the results of categorization for various methodologies and applications in the production of marketing products for specified segments [8]. Both machine learning algorithms allowed for the identification of persistent interest groupings that were utilized to develop cross-sell operations and customized product recommendations.

[9] learned continuous representations of unstructured text conversation using recurrent neural networks. It discovered a surprising conclusion by studying online chat records from Samsung's customer service: whereas tagged sessions provided by a small fraction of customers received overwhelmingly good reviews, the majority of unlabeled sessions would have received worse ratings from users. The data analytics provided in this research have consequences for detecting disgruntled customers on live chat services, but they also contributed theoretically to determining the extent of bias in online evaluation platforms.

The goals of this study are to first determine the impact of psychographic elements on e-commerce customer happiness and retention, and then to discover segmentation using psychographic and demographic parameters [10]. The four characteristics of the e-service quality model that better predict user behavior are the subject of this research [11]. It examines not just the influence of customer happiness on repurchase desire, word - of - mouth, and website revisit, but also the influence of customer trust. The outcome is expected to broaden understanding of various country cultures in relation to the various importance of e-service quality criteria.

A study of market segmentation of purchase behaviour of clients of health care items in Chandigarh, India, was undertaken for this study [12]. The goal of this research was to uncover the important factors that determine customer happiness as well as a model that can predict consumer satisfaction [13]. The results would provide information on how to select important variables and provided advice to a variety of merchants. Based on three-year historical data from an e-commerce company, four classification machine learning algorithms were assessed.

The XGBoost approach was built on a decision tree and used a linear model for logistic regression [14]. After improving the model, it was shown that the nonlinear model could better utilize these features and produced more accurate predictions. It integrated the single model first, then utilized the model fusion algorithm to fuse the model's prediction outcomes. The goal was to avoid the linear model's easy-to-fit accuracy and the decision tree model's over-fitting. The results showed that the model developed in this article is superior to the single model. Lastly, it was demonstrated by two sets of contrast experiments that the approach used in this study could effectively filter the features, reducing the model's complexity and improving ML classification performance.

The method [15] included Radio Frequency Identification data, which could correctly represent customers' in-store behavior and purchase behavior. It analyzed consumer segmentation using a broad learning algorithm. This approach was one of the most cutting-edge ML algorithms, and it was extremely efficient and

effective in categorization. The client behavior data used in this study came from a real-world Japanese supermarket. Customer segmentation was viewed as a multi-label categorization challenge.

The following method [16] developed a fuzzy data mining methodology for deriving membership functions from navigational data in order to discover fuzzy customer behavior on website features. The features selection technique was also used to track and analyses appropriate e-customer click data.

The goal of this study was to determine which of the aforementioned mediums was most suited for the target client, as well as to determine the type of data that was most acceptable for analysis in reaching this decision [17]. To do this, data was collected via surveys and pre-processed using data from multiple sources.

The strategy proposed a novel approach [18] to precision marketing and the establishment of tiered discount rates for consumers. Furthermore, it provided theoretical support for increasing the customer size and payment ratio, so boosting the expressway core operational decision-making level.

To improve data categorization, the logit leaf model was presented [19]. The concept was that by building models on segments of the data rather than the complete dataset, improved predicted performance could be achieved while keeping the understandability of the models built in the leaves.

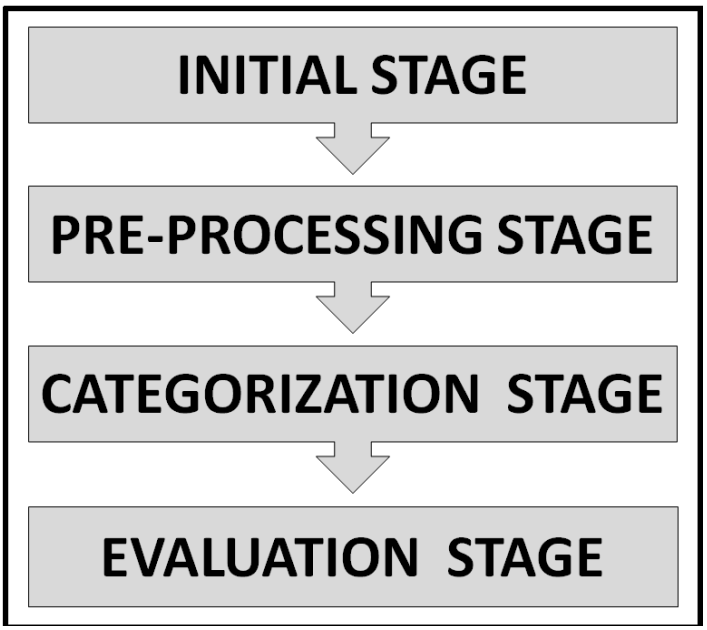
The new characteristics of client behavior were first examined [20]. It suggested an embedded consumer value model with thirteen indicators in three dimensions: buy value, interaction value, and market diffusion value, based on complex network theory and the RFM model, and taking into account the value created by customer connection and interaction.

III. Proposed Behavioral Model

The suggested Customer Segmentation and Behavioral Analytic System uses a machine learning technique called K-Prototype to evaluate customer behavior. It has the fundamental characteristics of all five effective elements of behavioral analysis. It focuses mostly on client retention, loyalty, and requirement clustering.

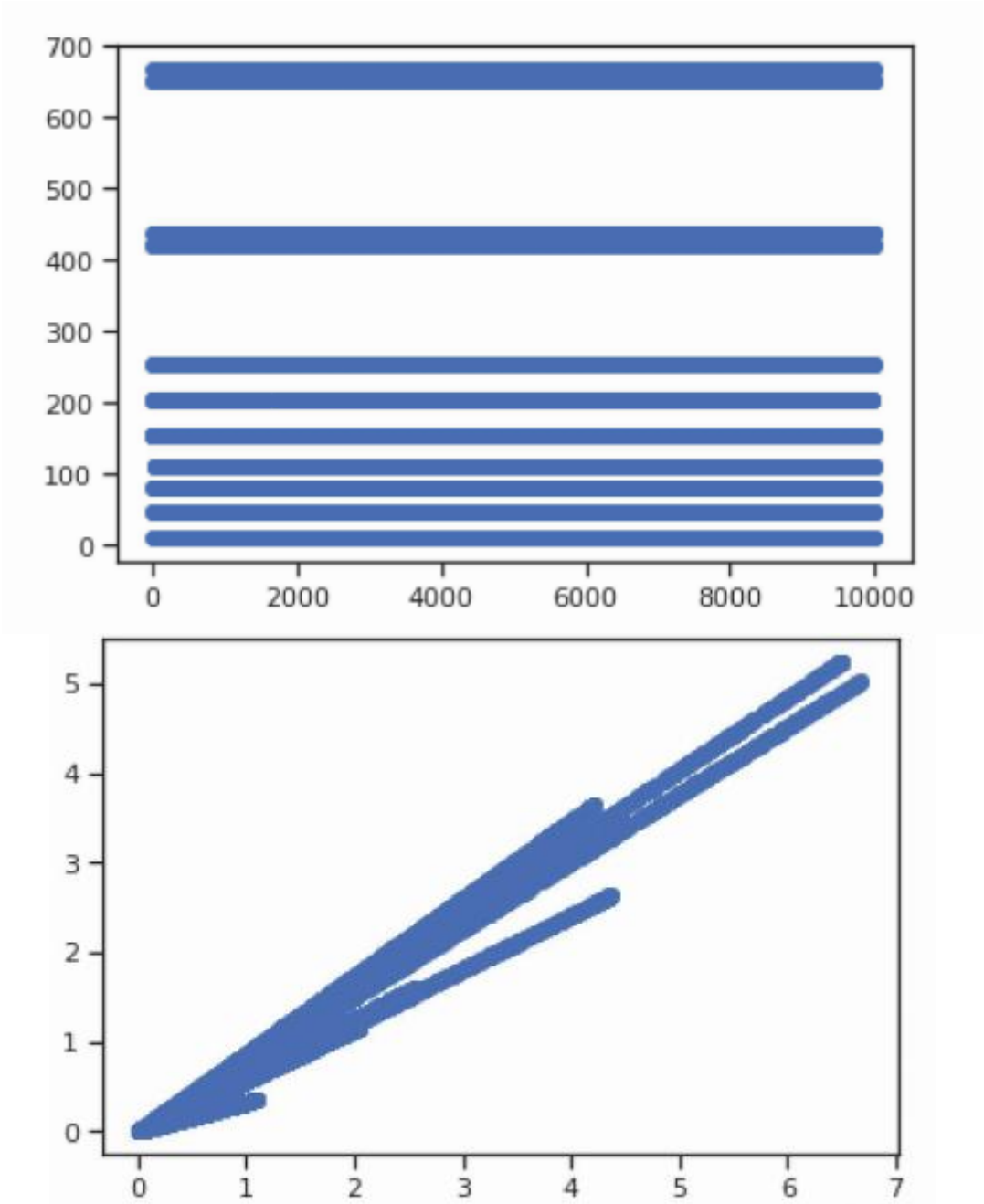
From a customer segmentation standpoint, Fig. 1 displays the system's early stage design. The starting stage, pre-processing stage, categorization stage, and evaluation stage are the four steps of its stage-wise architecture. It shows how the K-Prototype clustering technique accepts, processes, identifies, and groups input patterns.

At the outset, this system accepts the data and specifies the dataset properties and effective elements. It cleans and normalizes the data at the pre-processing step. The next step is to use the K-prototype and Elbow technique to reduce the number of segments. All of the metrics and influential factors are analyzed and evaluated in the final stage.



IV. Results

Many experimental results were obtained on the proposed system. It obtained the results for the "Units Sold" and "Unit Price". It is shown in the Fig. 2. Fig. 3 depict the results between “Total Revenue” and “Total Cost”.



The proposed system provided very promising relsuts and the system accuracy was 92%.

V. Discussion

For automotive related datasets, the proposed system was implemented. Several experiments were carried out. The findings were found using the DBI, and the best values were those that were closest to zero. The datasets yielded four and eight clusters, respectively. It also displayed the results by area to show the different item forms in order.

VI. Conclusion

The proposed customer behavior model analyses customers' buying choices in CAR-related datasets to assess their satisfaction and conduct. Through the usage of impacted essential factors, it used K-Prototype to measure user behavior. It evaluated a large wide variety of data and came up with valuable results. Its scope is wide and it can be extended for many brands, companies and applications.

References

[1]. C. J. Udokwu, F. Darbanian, N. F. M. Falatouri, and P. Brandtner. Evaluating technique for capturing customer satisfaction data in retail supply chain. The 4th International Conference on E-commerce, E-Business and E-Government, ACM, NY, USA, pp. 89–95, 2020.

[2]. H-F. Chen, and S-H Chen. "How Website Quality, Service Quality, Perceived Risk and Customer Satisfaction Affects Repurchase Intension? A Case of Taobao Online Shopping", In: Proceedings of the 10th International Conference on E-Education, E-Business, E-Management and E-Learning, ACM, NY, USA, 326–328, 2019.

[3]. S. Zhang, X. Tan, J. Wang, J. Chen, and X. Lai. "Modeling Customers' Loyalty using Ten Years' Automobile Repair and Maintenance Data: Machine Learning Approaches", In: Proceedings of the 2019 2nd International Conference on Data Science and Information Technology, ACM, NY, USA, pp. 242–248, 2019.

[4]. M. Zhao, Q. Zeng, M. Chang, Q. Tong, and J. Su, "A Prediction Model of Customer Churn Considering Customer Value: An Empirical Research of Telecom Industry in China", Discrete Dynamics in Nature and Society, Vol. 2021, pp. 1-12, 2021.

[5]. A. Ahmad, and O. Barukab. "A Novel Approach to Identify the Categories of Attributes for the Three-Factor Structure in Customer Satisfaction", Complexity, Vol. 2020, pp. 1-12, 2020.

[6]. J. Han, M. Kamber, and J. Pei. Data Mining: Concepts and Techniques. 3rd Edition, Waltham, Massachusetts: Morgan Kaufmann Publishers.

[7]. G. Anderst-Kotsis, and J. Ratzenböck. "Evaluating Customer Satisfaction using Behavioral Analytics of Data-Intensive Software Systems", In: Proceedings of the 20th International Conference on Information Integration and Web-based Applications Services, ACM, NY, USA, pp. 129–133, 2018.

[8]. I. Derevitskii, O. Severiukhina, and K. Bochenina. "Clustering Interest Graphs for Customer Segmentation Problems", In: Sixth International Conference on Social Networks Analysis, Management and Security, IEEE Press, pp. 321-327, 2019.

[9]. K. Park, M. Cha, and E. Rhim. "Positivity Bias in Customer Satisfaction Ratings", In: Companion Proceedings of the Web Conference: International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, ACM, pp. 631–638, 2018.

[10]. B. G. Muchardie, A. Gunawan, and B. Aditya. "E-Commerce Market Segmentation based on the Antecedents of Customer Satisfaction and Customer Retention", In: International Conference on Information Management and Technology, IEEE Press, pp. 103-108, 2019.

[11]. P. Rita, T. Oliveira, and A. Farisa. "The Impact of E-service Quality and Customer Satisfaction on Customer Behavior in Online Shopping", Heliyon, Elsevier, Vol. 5, Issue 10, pp. 1-14, 2019.

[12]. G. Kumari, Varsha, and K. M. Pandey. "Market Segmentation of Buying Behavior of Consumers of Health Care Products in Chandigarh India", International Journal of Recent Technology and Engineering, Vol. 8, Issue 4, pp. 610-614, 2019.

[13]. A-N. Wong, and B. P. Marikannan. "Optimising E-Commerce Customer Satisfaction with Machine Learning. Journal of Physics: Conference Series: International Conference On Computational Physics in Emerging Technologies, Vol. 1712, pp. 1-9, 2020.

[14]. C-J. Liu, T-S. Huang, P-T. Ho, J-C. Huang, and C-T. Hsieh. "Machine Learning-Based E-Commerce Platform Repurchase Customer Prediction Model, PLOS ONE, pp. 1-15, 2020.

[15]. Z. Wang, Y. Zuo, T. Li, C. L. P. Chen, and K. Yada. "Analysis of Customer Segmentation based on Broad Learning System", In: International Conference on Security, Pattern Analysis, and Cybernetics, IEEE Press, pp. 75-80, 2019.

[16]. H. Zaim, M. Ramdani, and A. Haddi. "Fuzzy-Based Mining Framework of Browsing Behavior to Enhance E-Commerce Website Performance: Case Study from Kelkoo.com", In: Proceedings of the 12th International Conference on Intelligent Systems: Theories and Applications, ACM, NY, USA, pp. 1–6, 2018.

[17]. I. T. Afolabi, R. E. Worlu, O. P. Adebayo, and O. Jonathan. "Predicting Customer Behavior with Combination of Structured and Unstructured Data", Journal of Physics: IOP Conference Series: 3rd International Conference on Science and Sustainable Development, Vol. 1299, pp. 1-14, 2019.

[18]. C. Qian, M. Yang, P. Li, and S. Li. "Application of Customer Segmentation for Electronic Toll Collection: A Case Study", Journal of Advanced Transportation, Vol. 2018, pp. 1-9, 2018.

[19]. A. D. Caigny, K. Coussement, and K. W. D. Bock. "A New Hybrid Classification Algorithm for Customer Churn Prediction based on Logistic Regression and Decision Trees", European Journal of Operational Research, Elsevier, Vol. 269, Issue 2, pp. 760-772, 2018.

[20]. J. Li, A. D. Nigatu, B. Yu, Q. Gu, and Y. Yang. "Understanding Customer Value in the Mobile Internet Era", Discrete Dynamics in Nature and Society, Vol. 2021, pp. 1-14, 2021.

Priyanka Meena, et. al. "Design of a Novel Customer Satisfaction Model using Segmentation and Machine Learning." *IOSR Journal of Computer Engineering (IOSR-JCE)*, 24(2), 2022, pp. 06-10.