E-Commerce Production-Profit Enhancement by Customer Behavior Analysis on Social Network Data

R Sudha[1], Dr N Uma Maheswari[2] and Dr R Venkatesh[3]

[1]Assitant Professor,Department Of Information Technology, [2]Professor,Department Of Computer Science and Engineering,
[3]Professor,Department Of Information Technology
 PSNA College of Engineering and Technology, Dindigul,Tamilnadu,India.

Abstract—One of the drastically growing applications under bigdata analytics is consumer behavior analysis for enhancing the e-commerce industry. This industry requires a useful analytics tool for understanding the online customer those who are interacting in the e-commerce site. Most of the marketing people are using the internet to market their company products. Online customers share their opinionson various social websites, such as Facebook, Twitter, and Instagram. By analyzing the tweets and comments, customer behavior can be identified. In this paper, distance-based clustering and Multi-Class Support Vector Machine based classification are applied to the data for predicting customer behavior. From the experiment, it is identified that the proposed method is highly suitable for a small size dataset.

Keyword: Customer Analysis, Behavior Analysis, Social Network Data, E-Commerce Improvement

I. INTRODUCTION

The buying habits with the social trends, background information, and frequency patterns influencing the customer decision to buy anything in the site are called as customer behavior. E-commerce businesses need to understand the behavior of each customer to provide more products and better services which satisfy the customers. Any business people cannot identify the customer one who is buying online, whereas it can be recognized that the way they are shopping. From the online data, it can obtain the sales frequency, preference score of the product, how marketing influences the customer, and the service offers. By analyzing these factors, it is easy to improve the customer relationship with the business to increase production pleasantly. Personal, social, and psychological factors are the three primary factors that influence customer behavior. Analyzing customer behavior is observing qualitative and quantitative information on how the customer communicates with the E-commerce industries.

Figure-1. Bigdata Analytics

Bigdata, behavior analysis, social data, user-experience, privacy, validated learning, etc., are the words that make different meanings for learning customer behavior. Monitor and collect all the above said kinds of data, implement an algorithm, automate the analysis to identify customer behavior. By increasing the self-initiated action, it makes the customer accept the behavior who is using the specific products. The amount of customer data is continuously growing every day and becomes bigdata. Only by overviewing the data, it is not possible to do a decision making. Thus, it needs an efficient analytical tool to makethe better decision making. An overview of the data analytics process is shown in Figure-1.

New system design is created to understand the people and make themselves know how they interact with the e-business. It helps the customer to select their required products over massive competitions. For that, it is assumed that a pool of data is collected, a machine learning algorithm with a prediction engine, a group of people, working, and learning on it. Finally, from the learning,

7743

andsignificant decision-making result is obtained. Customer behavior analysis increases the business benefits. For understanding the need for customer behavior analysis, one of the researchers carried out a case study related to cloud data analysis for e-commerce implementation in Countants (see Figure-2).

From figure-2, it is easy to understand that the retailers of e-commerce can make various benefits obtained from the perceptions, are taken from behavior analysis of the customer. These valuable perceptions lead to a highly personalized method to customer requirements, which can improve their long-time value to the business. Also, the customer behavior analysis in e-commerce helps to reduce the customer attainment costs, improve the product variety endorsements, and increase the lead creation process. Before going to implement the proposed method, it is necessary to understand the issues and challenges faced in the earlier research methods and the real-time e-commerce industry. Hence, this paper discussed a few more first research methods proposed in previous days.
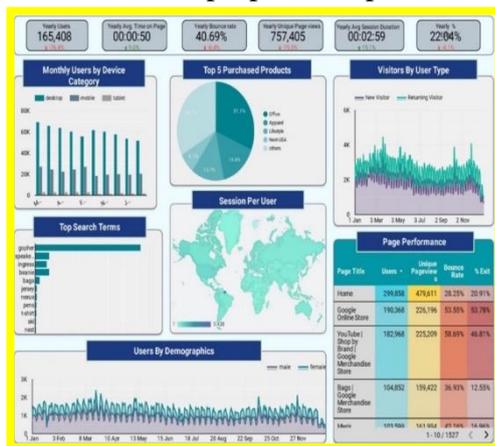

Figure-2. A Case Study on Customer Behavior

II. LITERATURE SURVEY

One of the websites which provide a powerful platform for a vast number of people to publish their research methods, opinions, stories about real-time experiences, and experimental methodologies used for data analysis, algorithms belongs to machine learning and deep learning methods [1]. Authors in [2] presented a classification approach for analyzing the behavior

data obtained from the tweeter dataset. The author in [3] explains the importance of the features in data science—the algorithms suitable for particular data, background information, and the application. The author explained about business data, simplification of data, creating a decision tree on the data to predict the purchase behavior of the customers. The author describes the automatic methods used for analysing tweets in [4]. The author explained about Naïve Bayes, SVM, and maximum entropy methods for tweet analysis. From the research, it can be identified that machine learning algorithms are efficient in data analytics. The author in [8-11] explains various algorithms for analyzing social data and the necessity of social data analytics. Different kinds of algorithms are used for several types of data analytics processes like sentiment analysis, text analysis, and machine learning-based classification. The author explained KNN based feature selection for social data with detailed pre-processing steps. From the overall literature survey varied out, it is identified that a sequence of data analytics processes is required for improving the efficiency of the data analytical process. Thus, initially, the various strategy evaluation process is explained.

A. Strategy Evaluation of E-Commerce

Following the modern world economy, each company has to recover the best. For that, it needs to follow constant planning, strategy, and budget analysis. Initialize the protection strategies and cleansing planning models and created a refined and adjusted model by making assumptions from the past[5]. E-commerce companies are working with different scenarios like competitions, rapid response to customer requirements, customer-based, and flexibility and adjustment on the strategies and plans [6]. Before analyzing the e-business strategies, it needs to be diagnosed with purchase strategies. Comparing the business and buying strategies is illustrated in Figure-3.

Figure-3. Comparison of Business Strategies with Purchase Strategies

The companies are focusing on three major phases, such as customer retention, acquisition, and extension. Three contextual strategies evaluating by the business companies are marketing, sales, and customer service [7]. Customer service is the main factor which increases sales and is influenced by marketing. Similarly, the buying strategy focusing on shopping, service evaluation, and the selection of the products.

In another way, it can be said that one of the significant business assets is the behavior of the customer. The action of the customer is analyzed for extracting the internal opinion of the customer's information and expressed the experience of the e-commerce website to the customer. Behavior analysis includes browsing behavior, purchasing behavior, and email behavior. The browsing behavior is used to track the customer's online activity on the e-commerce company. Some of the factors included in browsing behavior are how the customers are attracted to a specific website, products searched, and recording the conversation between the customer and the e-commerce. Purchasing behavior shows the interest of the customer and what kind of products they prefer from the site. The data associated with the purchasing behavior includes some factors are, buying patterns and response of the behavior like getting offers and discounts based on the customer's purchase. The last behavior is email behavior, which says about how the customer reply to the mails sent by the marketing teams. This email behavior is used to analyze customer behavior. It includes three factors, such as the number of business emails read by the customer, the number of users clicked on the emails, and the number of emails promotes the online purchase.

Evaluation of e-commerce business with the customer can be done separately or by merging together over the evaluation factors. The purchase strategies are evaluated by analyzing the price, quality, trends, design, a model of the product, and experiences with the emotional petition. Decision making is obtained from customer choices. By integrating the strategic business evaluation with the buying strategy evaluation can increase production and sales.
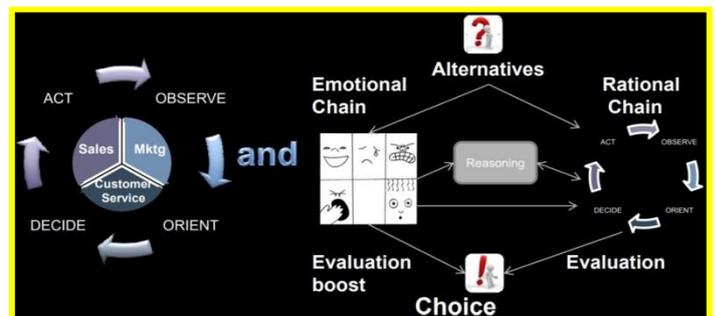

Figure-4. Business and Purchase Evaluation

In terms of business, applying marketing methodologies, provide more customer service through that decide the sales are the evaluation processes. For the buying side, evaluate the old customer emotions, rational chain, choose the best, provide feelings based on the product. These two evaluation steps are reviewed for increasing product sales, which is illustrated in Figure-4. The entire process shown in Figure-4 explains the decision making in the business and customer choice. From the evaluation process, the general information is collected. Based on processing the data, the company or the customer can decide about the sales and purchase. Each customer shows their emotions according to their feelings about the product. From the existing sales or purchase reports, the market effort is obtained, and it helps to create a market dashboard list out the various metrics used in visualizing the results

III.   PROPOSED APPROACH

The overall functionality of the proposed approach is illustrated in Figure-5. It says that the dataset is gathered and converted into a human-computer understandable format. A semi-automatic pre-processing is applied for increasing the quality

7745

of the dataset. Then the distance-based clustering method is used for similar grouping data. Finally, the MSVM method is used for classifying and predicting the purchasing behavior of the customer on a specific product or one category of product.
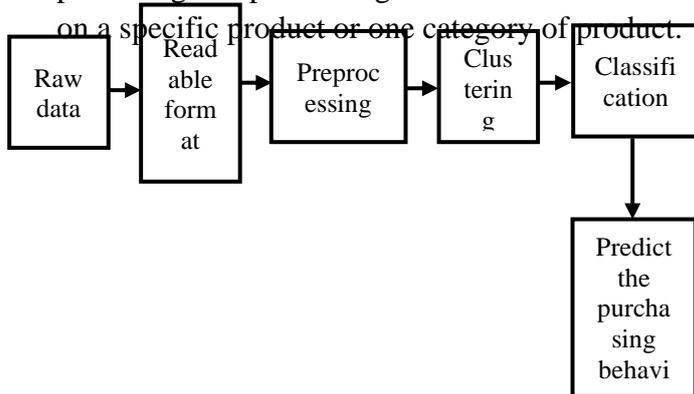

Figure-5. Overall Functionalities of Proposed Method

Initially,the clustering process is applied to the data to create various categories dividing similar customers, called clusters. A cluster represents a collection of points from the given dataset. Those points are very similar to each other among them within the cluster. Also, those points may be associated with the other clusters in the same dataset. One of the significant clustering methods is distance-based clustering, which creates clusters by computing the distance among the points. Points who obtained lesser distance belong to the same cluster. The formula used to compute the distance among the data is measured by mean value is,

$$\min_{y} \sum_{x \in D} \|x - y\|^2$$

Where it gives the geometric median of the data, the distance-based clustering method can provide better accuracy in clustering. The popular K-means clustering method follows the distance-based clustering method based on a center point. Hence, most of the research works used distance-based and k-means clustering methods for simplicity. After clustering, a multi-class SVM (MSVM) is used for classifying the overall data. For extracting various classes available in the data based on the features, the MSVM method is used. However, different classes are obtained from the clusters over the

behavior feature. The efficiency of the proposed method is verified from the experimental implementation.

A. Dataset

The main objective of this paper is to provide necessary information about customer behavior analysis for improving the business profit. For that, Mall_Customers data is taken from the Kaggle data store available at https://www.kaggle.com/akram24/mall-customers. It has a lot of feature values like name, age, phone number, address, city, state, mother tongue, occupation, annual income, frequency of visits, etc. Also, twitter dataset related to purchasing cameras, cars, iPad, tablet, and smartphones online is collected and used in the experiment. From the manual analysis, it is found that some of the words or verbs, such as bought, purchased, received, good, like, most like, cheaper, and quite common, are used to determine the buying behavior of the user. The volume of the data used in the experiment is 5,48,573.

B. Experimental Results and Discussion

The overall experiment is carried out in MATLAB software, and the results are verified. The proposed method includes pre-processing, clustering, and classification. From the classification label, it is easy to identify customer purchasing behavior for a specific product or several products. In the experiment, after pre-processing, 11% of the dataset is eliminated due to error, missing values, duplicate and irrelevant data from the whole dataset.
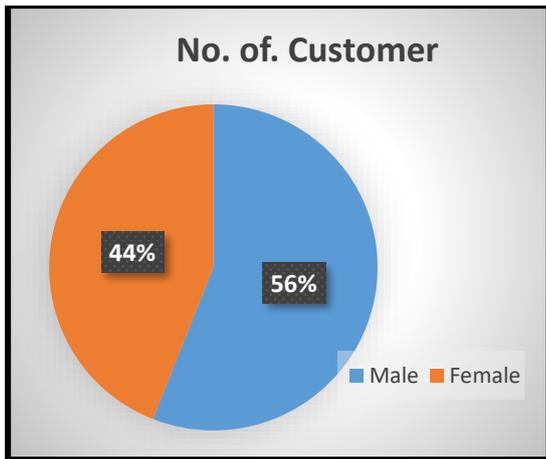
Figure-6. Data Clustering Regarding Gender

After clustering, the entire dataset is clustered according to various clusters. Visualizing all the clusters in a single screen is a crucial task. Hence, two cluster data are given in Figure-6 and Figure-7. The customers are also clustered according to gender. It helps to identify what are which gender prefers all the products. For example, TV, smartphones, iPad, and other electronic products are common for both male and female customers. The clustering process based on the gender feature is obtained from the experiment, and its relevant result is given in Figure-6. The e-commerce business needs to understand the product types and the age of the customer. From the result, it is identified that online customers belong to the age from 18yrs to 78 yrs. The highest percentage of customers belongs to 25 to 35 prefer online purchases.
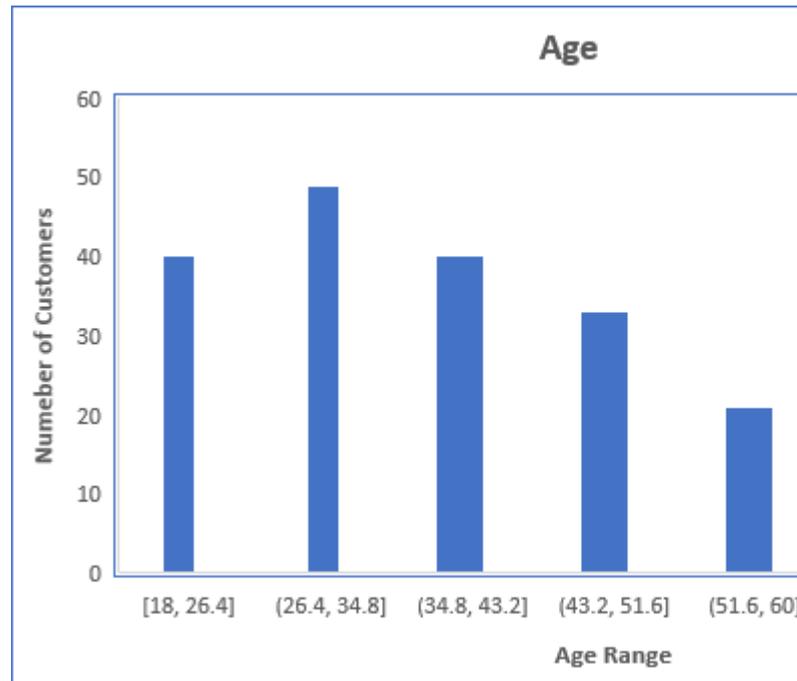

Figure-7. Customer Age Range

Next, the spending score of the customer is analyzed from the data, and the obtained result is given in Figure-8. Each customer has clustered according to the appending score. It means, from the customer's income, how much amount of spending on online purchase is calculated, and according to their score, the customers are clustered. From the payment, the overall spending score of the customers is from 12% to 54%, and it is shown in Figure-8. From the overall spending score, it is identified that a greater number of customers spending the highest income in the online purchase is 40% to 60%.
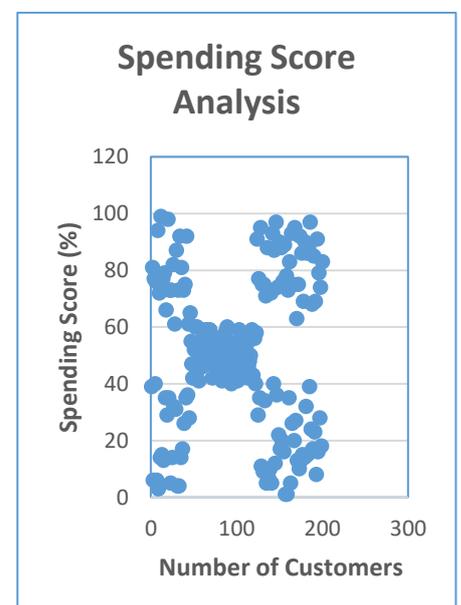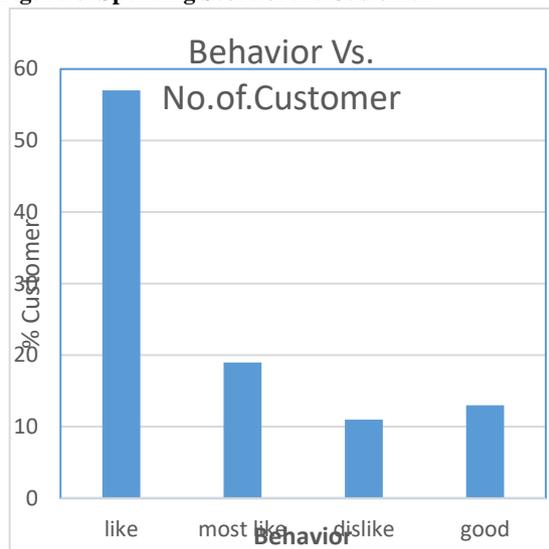

7747

Figure-8. Spending Score of the Customer

Figure-9. Customer Behavior

Finally, the customers are grouped according to the purchasing behavior of the customers. The entire dataset is clustered over the behavior words, such as like, most like, dislike, and good is given in Figure-9. From the result, it is identified that the highest amount (57%) of customers like online purchasing, and 11% of the customers dislike online purchasing.

From the above discussion and experimental results, it is clear that for a small size dataset, the proposed approach is suitable for analyzing customer behavior.

IV. CONCLUSION

All e-commerce companies are focusing on tracking customer behavior for increasing business profit and production. Though the company is small or big, the customer's opinion, idea, and behavior determine the products projected in the market. Customer behavior says their experience over product, website usage, and other impacts satisfy them. This paper is motivated to provide an overview of the introduction of e-commerce, business strategies, purchasing strategies, customer behavior, how it influences the business. It also provides a simple method for customer data analysis, including pre-processing, clustering, and classification. From the results, it is identified that the proposed method is essential and suitable for small size dataset analysis.

V. REFERENCES

[1]. https://towardsdatascience.com / analyzing – feature – importance – user – behaviour – and – ux – performance -cbf32d55eff8.

[2]. Chithra R G, Harshitha G M, AnuPrakash M P, and Rakshitha H B, (2019), "Behavioural Analysis of Tweeter data: A Classification Approach", International Journal of Engineering Research & Technology, Vol. 7, No. 8, PP. 1-4.

[3]. Matias Eiletz, "Analysing User behaviour and UX performance with Feature importance", https://towardsdatascience.com/analyzing-feature-importance-user-behaviour-and-ux-performance-cbf32d55eff8.

[4]. Sahayak, Vijaya Shete, Apashabi Pathan, (2015), "Sentiment Analysis on Twitter Data: Varsha ", International Journal of Innovative Research in Advanced Engineering, Vol. 2, No. 1, pp. 178-183.

[5]. https://www.ecapitaladvisors.com/blog/planning-budgeting-and-forecasting/

[6]. Bhandari, Gokul, Bliemel, Michael, Harold, Allan, Hassanein, Khaled, (2004), "Flexibility in e-Business Strategy: A requirement for Success", Global Journal of Flexible Systems Management, Vol. 5, No. 22, pp. 11-22.

[7]. Ang, Lawrence, Buttle, Francis, (2010), "Managing for Successful Customer Acquisition: An Exploration", Journal of Marketing Management, Vol. 22, pp. 295-317. Doi: 10.1362/026725706776861217.

[8].Prabu, A. V., & Kumar, G. S. (2019). Hybrid MAC based adaptive preamble technique to improve the lifetime in wireless sensor networks. J. Adv. Research in Dynamical & Control Systems, 11(1), 240-249.

[9].Srinivas, K., Prabu, A. V., & Sambasivarao, K. (2019). A Real Time Prototype Model for Enhancing the Security Features in the ATM Units International. Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(7), 1936-1939.

[10].B. Vinuthna, P. Ravi kiran., & A. V.Prabhu(2019).Smart Electricity Bill Generation using Mobile App .International Journal of Innovative Technology and Exploring Engineering (IJITEE),8(6), 1698-1702

[11].K Vijaya Manasa , A V Prabu , M Sai Prathyusha , S Varakumari (2018) .Performance monitoring of UPS battery using IoT" International Journal of Engineering & Technology, 7 (2.7).352-355

[12].Manikandan, R and Dr.R.Latha (2017). "A literature survey of existing map matching algorithm for navigation technology. International journal of engineering sciences & research technology", 6(9), 326-331.Retrieved September 15, 2017.

[13].A.M. Barani, R.Latha, R.Manikandan, "Implementation of Artificial Fish Swarm Optimization for Cardiovascular Heart Disease" International Journal of Recent Technology and Engineering (IJRTE), Vol. 08, No. 4S5, 134-136, 2019.

[14].Manikandan, R., Latha, R., & Ambethraj, C. (1). An Analysis of Map Matching Algorithm for Recent Intelligent Transport System. Asian Journal of Applied Sciences, 5(1). Retrieved from https://www.ajouronline.com/index.php/AJAS/article/view/4642

[15].R. Sathish, R. Manikandan, S. Silvia Priscila, B. V. Sara and R. Mahaveerakannan, "A Report on the Impact of Information Technology and Social Media on Covid–19," 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Thoothukudi, India, 2020, pp. 224-230, doi: 10.1109/ICISS49785.2020.9316046.

[16].Manikandan, R and Dr.R.Latha (2018). "Map Matching Algorithm Based on a Hidden Markov Model for Vehicle Navigation" International Journal of Advanced Technology in Engineering and Science, 6(6), 36-42.

[17].Manikandan, R and Dr.R.Latha (2018). "GLOBAL POSITIONING SYSTEM FOR VEHICLE NAVIGATION" International Journal of Advances in Arts, Sciences and Engineering (IJOAASE), 6(13), 1-9.