

An Efficient Machine Learning Technique for Prediction of Consumer Behaviour with High Accuracy

¹Poonam Sonwane, ²Prof. Monika Raghuwanshi

¹Research Scholar, ²Assistant Professor

^{1&2}Department of Computer Science Engineering,

^{1&2}RKDF College of Engineering, Bhopal, India

Abstract : Consumers choose to purchase a given item by looking at these evaluations and surveys. Such substance can be positive or negative audits made by consumers who have recently utilized the item. The Machine Learning Calculation can help us to visual portrayal of the information and vectorize the information. This paper presents the Naïve Bayes and Logistic Regression method to investigate the consumer behavior. The logistic regression strategies improved preferred performance over others. The current issues are examined, and afterward, current answers for these issues are introduced and talked about. The trial results show that the proposed strategy has higher precision, recall and F1 score. The strategy is end up being effective with high accuracy on remarks. The reproduction and analysis is finished utilizing the python spyder 3.7 software.

IndexTerms - Naïve Bayes, Logistic Regression, Precision, Recall, F1 Score, Spyder.

I. INTRODUCTION

A Machine learning (ML) is the scientific investigation of calculations and factual models that PC frameworks use to play out a specific task without utilizing unequivocal guidelines, contingent upon examples and derivation all things considered. It is viewed as a subset of mechanized thinking. Machine learning calculations develop a mathematical model dependent on example information, known as "planning information", to make forecasts or decisions without being explicitly customized to play out the task.[1] Machine learning calculations are utilized in a wide arrangement of utilizations, for example, email sifting and PC vision, where it is problematic or infeasible to develop a conventional count for enough playing out the task.

Machine learning is firmly identified with computational insights, which focuses on making forecasts utilizing PCs. The investigation of scientific smoothing out conveys strategies, theory and application spaces to the field of machine learning. Information mining is a field of study inside machine learning, and focuses on exploratory information analysis through unaided learning.[3][4] In its application across business issues, machine learning is likewise implied as perceptive investigation.

A few learning calculations target finding better portrayals of the data sources gave during training.[11] Exemplary models join head segments analysis and group analysis. Feature learning calculations, likewise called portrayal learning calculations, often try to save the information in their information yet also transform it to such an extent that makes it useful, often as a pre-handling venture before performing classification or expectations. This methodology permits recreation of the data sources starting from the unknown information delivering dissemination, while not being fundamentally given to configurations that are unlikely under that conveyance. This replaces manual component planning, and permits a machine to both get comfortable with the features and use them to play out a specific task.

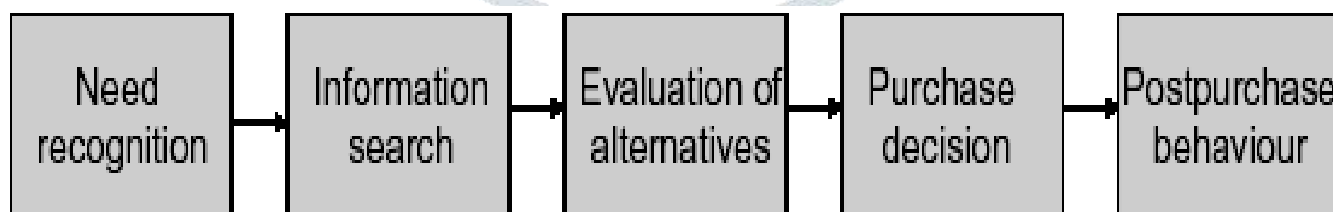


Figure 1: Five-Stage Model of the buying process

This model suggests that consumers go through each of the five phases in purchasing an item. This might be the situation in high-including buys. In low-inclusion buys, consumers may skip or converse a portion of these stages. This model shows the full scope of contemplations that emerge when a consumer face a profoundly including new buy.

Consumer satisfaction provided by three general components. It can be identified in extant definitions:

- 1) Consumer satisfaction is a response (Emotional or Cognitive)
- 2) The response pertains to a particular focus (expectations, product, consumption experience, etc.)
- 3) The response occurs at a particular time (after consumption, after choice, based on accumulated experience, etc.)

The pre-deals stage where there are the assumptions for the item, the profits, the cost and the accessibility of item. The business stage when client plate the climate, the item, the kind of administration, the conveyance, the quality and the change from the

market. The after-deals stage when client anticipates the support or the advices, the substitution of item or the arrival of aggregate, fixes and cycles of charges.

II. PROPOSED METHODOLOGY

Focusing on scientific approach to assess how assistance is acknowledged in the public arena, we created consumer behavior displaying framework. An exact analysis of this client created substance can be helpful to online business associations to acquire bits of knowledge and comprehend their consumers' goals and prerequisites. Machine Learning Calculations can help us plot exact visual portrayals of such consumer behavior. Machine learning classifiers incorporate Naïve Bayes, Logistic Regression are utilized in the planning of the framework.

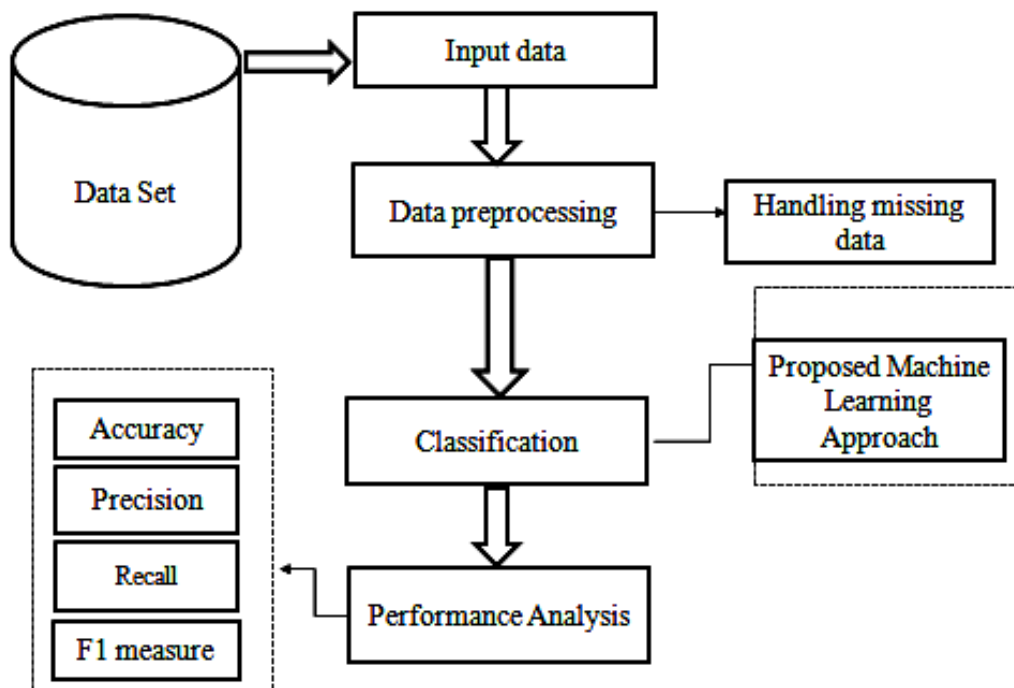


Figure 2: Flow Chart

Figure 2 is showing the proposed flow chart. The flow of work starts with to define input configuration and taken consumer behavior data set from the kaggle machine learning repository. Now before apply machine learning techniques, firstly apply the steps for data pre-processing. The sample of data is taken in this step, it is also known as training data.

Now apply proposed approach based on the logistic regression. At last all training data is process and give predication of diseases. Now, Results graph generation and calculation of necessary parameters is done.

III. RESULT AND ANALYSIS

The implementation of the proposed algorithm is done over python spyder 3.7. The sklearn, numpy, pandas, matplotlib, pyplot, seaborn, os library helps us to use the functions available in spyder environment for various methods like decision tree, random forest, naive bayes etc.

Index	id	name	asins	brand	category
0	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
1	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
2	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
3	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
4	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
5	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
6	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
7	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
8	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
9	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
10	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
11	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
12	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...
13	AVqkIhwDvBe3...	All-New Fire HD 8 Tablet...	B01AHB9CN2	Amazon	Electro...

Figure 3: Dataset

Figure 3 is showing the amazon data set. Total 69000 person dataset given in this file.

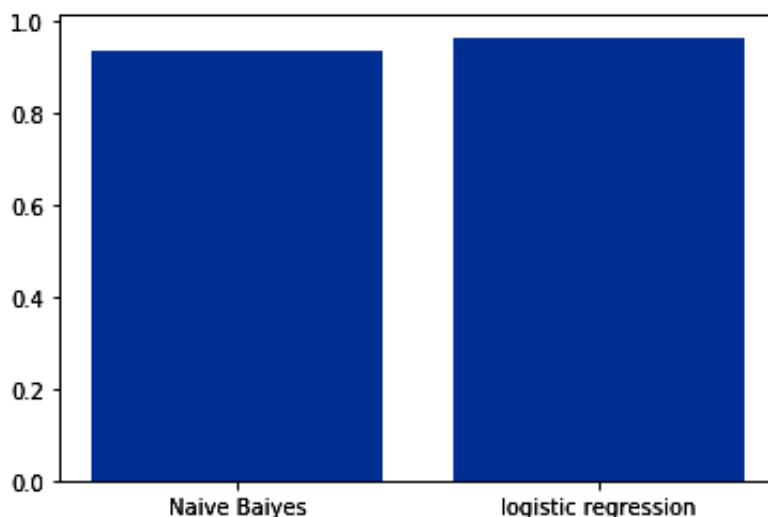


Figure 4: Accuracy comparison

Table 1: Comparison of Proposed work with previous work

Sr. No.	Parameters	Previous work [1]	Proposed Work
1	Method	Naive Bayes	logistic regression
2	Accuracy %	93.41 (94)	96.62 (97)
3	Classification error %	6	3
4	Precision %	92	98
5	Recall %	93	97
6	F-measure %	92	97

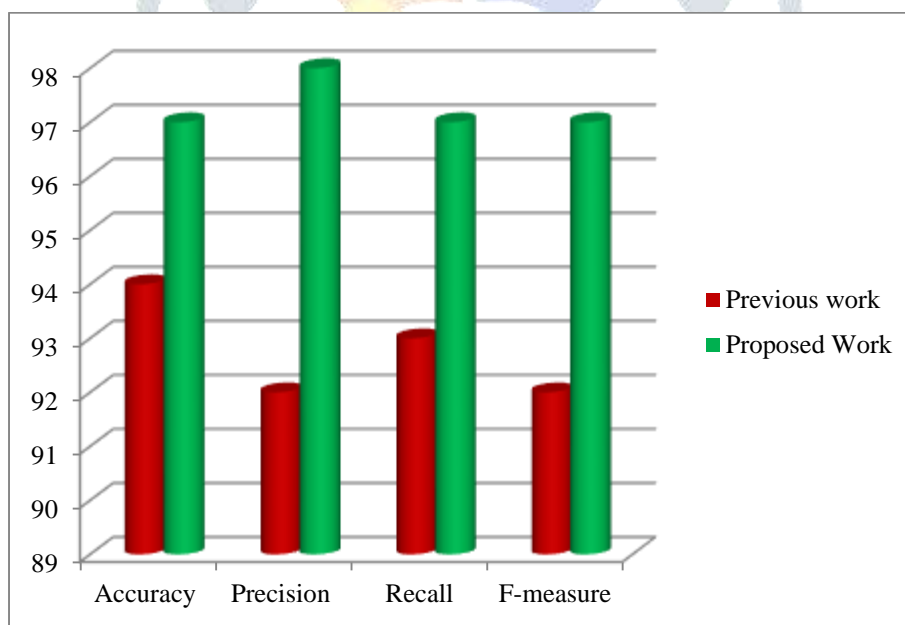


Figure 5: Parameters comparison

Table 1 and figure 5 is showing that the comparison of proposed and previous work results. It is clear that proposed method gives 97% accuracy while in previous there is 94% accuracy.

IV. CONCLUSION

This paper presents an approach to help the organizations in knowing their customers and incorporating targeted marketing techniques to increase their customer base and profits. Sentiment analysis helped us evaluate consumers' sentiments related to various products which in turn helped us analyze the product's performance in the market. It is clear from simulated results that proposed approach gives 97% accuracy while in previous there is 94% accuracy. The classification error is 3% in proposed while 6% in previous approach. The precision value is 98% and F-measure is 97% in proposed while previously it 92% and 92% respectively.

REFERENCES

- [1]. V. Shirame, J. Sabade, H. Soneta and M. Vijayalakshmi, "Consumer Behavior Analytics using Machine Learning Algorithms," 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), Bangalore, India, 2020, pp. 1-6, doi: 10.1109/CONECCT50063.2020.9198562.
- [2]. E. Kafeza, A. Kanavos, C. Makris, G. Pispirigos and P. Vikatos, "T-PCCE: Twitter Personality based Communicative Communities Extraction System for Big Data," in IEEE Transactions on Knowledge and Data Engineering, vol. 32, no. 8, pp. 1625-1638, 1 Aug. 2020, doi: 10.1109/TKDE.2019.2906197.
- [3]. M. A. Sharkh, Y. Xu and E. Leyder, "CloudMach: Cloud Computing Application Performance Improvement through Machine Learning," 2020 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), London, ON, Canada, 2020, pp. 1-6, doi: 10.1109/CCECE47787.2020.9255686.
- [4]. S. Shahriar, A. R. Al-Ali, A. H. Osman, S. Dhou and M. Nijim, "Machine Learning Approaches for EV Charging Behavior: A Review," in IEEE Access, vol. 8, pp. 168980-168993, 2020, doi: 10.1109/ACCESS.2020.3023388.
- [5]. J. Edmond Meku Fotso, B. Batchakui, R. Nkambou and G. Okereke, "Algorithms for the Development of Deep Learning Models for Classification and Prediction of Behaviour in MOOCS.," 2020 IEEE Learning With MOOCS (LWMOOCS), Antigua Guatemala, Guatemala, 2020, pp. 180-184, doi: 10.1109/LWMOOCS50143.2020.9234363.
- [6]. P. A. Savenkov and A. N. Ivutin, "Organizations Data Integrity Providing through Employee Behavioral Analysis Algorithms," 2020 9th Mediterranean Conference on Embedded Computing (MECO), Budva, Montenegro, 2020, pp. 1-3, doi: 10.1109/MECO49872.2020.9134271.
- [7]. J. R. Goodall et al., "Situ: Identifying and Explaining Suspicious Behavior in Networks," in IEEE Transactions on Visualization and Computer Graphics, vol. 25, no. 1, pp. 204-214, Jan. 2019, doi: 10.1109/TVCG.2018.2865029.
- [8]. D. Damkevala, R. Lunavara, M. Kosamkar and S. Jayachandran, "Behavior Analysis using Serverless Machine Learning," 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2019, pp. 1068-1072.
- [9]. Asniar and K. Surendro, "Predictive Analytics for Predicting Customer Behavior," 2019 International Conference of Artificial Intelligence and Information Technology (ICAIT), Yogyakarta, Indonesia, 2019, pp. 230-233, doi: 10.1109/ICAIT.2019.8834571.
- [10]. F. D. Pereira, E. H. T. Oliveira, D. Fernandes and A. Cristea, "Early Performance Prediction for CS1 Course Students using a Combination of Machine Learning and an Evolutionary Algorithm," 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT), Macei??, Brazil, 2019, pp. 183-184, doi: 10.1109/ICALT.2019.00066.
- [11]. M. A. Salitin and A. H. Zolait, "The role of User Entity Behavior Analytics to detect network attacks in real time," 2018 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), Sakhier, Bahrain, 2018, pp. 1-5, doi: 10.1109/3ICT.2018.8855782.
- [12]. A. Bouhoute, R. Oucheikh, K. Boubouh and I. Berrada, "Advanced Driving Behavior Analytics for an Improved Safety Assessment and Driver Fingerprinting," in IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 6, pp. 2171-2184, June 2019, doi: 10.1109/TITS.2018.2864637.
- [13]. M. Callara and P. Wira, "User Behavior Analysis with Machine Learning Techniques in Cloud Computing Architectures," 2018 International Conference on Applied Smart Systems (ICASS), Medea, Algeria, 2018, pp. 1-6, doi: 10.1109/ICASS.2018.8651961.
- [14]. R. G. Rajasekaran, S. Manikandaraj and R. Kamaleshwar, "Implementation of Machine Learning Algorithm for predicting user behavior and smart energy management," 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI), Pune, 2017, pp. 24-30, doi: 10.1109/ICDMAI.2017.8073480.
- [15]. Y. Zhang, R. Yang, K. Zhang, H. Jiang and J. J. Zhang, "Consumption Behavior Analytics-Aided Energy Forecasting and Dispatch," in IEEE Intelligent Systems, vol. 32, no. 4, pp. 59-63, 2017, doi: 10.1109/MIS.2017.3121551.