

# Prediction of Customer Purchase Intention using Social Media Data

Rashmi Kale<sup>1</sup>, Kunal Bidwai<sup>2</sup>, Mandar Maske<sup>3</sup>, Rahul Bansode<sup>4</sup>, Prasad Gurav<sup>5</sup>

Assistant Professor, Department of Computer Engineering<sup>1</sup>

Students, Department of Computer Engineering<sup>2,3,4,5</sup>

Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India

**Abstract:** Predicting customer behavior in the context of e-commerce is becoming increasingly important as people shift from visiting physical stores to shopping online. By facilitating a more personalized shopping process, it can boost consumer happiness and sales, resulting in higher conversion rates and a competitive edge. Models for forecasting consumer behaviors can be constructed using and supplementing customer data. This research looks at how a prominent German apparel shop uses machine learning models to forecast purchases, which is a significant use case. Following that, by doing a descriptive data analysis and individually training the models on the distinct datasets, this study provides insight into the performance differences of the models on sequential and static customer data. Three different algorithms are used.

**Keywords:** Machine Learning, Customer Purchase Intention, SVM, KNN, Random Forest

## I. INTRODUCTION

Web based business, the action of trading items on the web, is one of the many fields upset by information science. One of the fundamental objectives for web-based business organizations is to increment buy change rates, for example the level of site guests who complete the buy at online stores. To accomplish this objective, web-based business organizations as well as specialists in scholarly world have given endeavors in investigating and demonstrating the ways of behaving of website page users. Especially in late year, there has been a pattern in exploration to utilize AI strategies to anticipate the way of behaving of clients is a site that assists the client with imagining and examination the item. The audits of the items are examined in this site by utilizing php and python. It save individuals time by seeing alot of remarks. So, the client can without much of a stretch picture the audit soon. On the off chance that the client can give their determination our site will show the item with their determination and furthermore envision the remarks in the designs format. Now a days the vast majority of the site i.e amazon, flipkart will just presentation their items subtleties and their rating and remarks. Alongside a portion of those detail in amazon flipkart we add a few elements in our website, so that client can without much of a stretch view their item comments. customers, possible clients, conventional clients, or clients with space explicit abilities) and RNPs (i.e., technological creativity, or both innovative and market imaginativeness), as well as the different perspectives on client. To introduce an orchestrated view on factors in the circle of the enhancing organization and the client that impact client mix's prosperity along the extreme advancement improvement process (i.e., disclosure, hatching, and speed increase). Its present roads for future exploration and talk about administrative ramifications of our integrated view.

## II. RELATED WORK

Sushant Kumar, Mikko Murphy the paper Study utilizes the stimuli-organic entity reaction (SOR) hypothesis, which demonstrates that specific natural improvements impact the purchasers inside state or creature and influences their way of behaving, thus. In particular, there study involves selflessness as the boost, supporting neighborhood makers, straightforwardness, fulfillment with naming, and craving for marking as the customers' interior state (organic entity), and buy goals and brand love as the reaction. Cross-sectional information was gathered from 2045 nearby food purchasers related with Facebook-based REKO (fair utilization) bunches in Finland.

Erik Nettet, Ola Bergem, point of the paper was to examine chain devotion impacts of clients' apparent worth of dependability programs in staple retailing. The principal finding was that clients' apparent worth of a reliability program

meaningfully affects chain faithfulness, and the control impacts are chain subordinate. Supervisors ought to subsequently consider fulfillment creation, picture building, and devotion program esteem creation as equal cycles.

Nusrat Parvin; Sayaka Zaman; Samia Amin, Their study expected to decide the relationship between person's web-based communication and e-cigarette buy expectation from Facebook vape gatherings. A cross-sectional review was directed among 214 respondents utilizing the Facebook stage from September to December 2019. Information were gathered by means of a self-controlled survey.

Cost Clement Adoo, He concentrate on the after effects of 1726 datasets from two online business stages recommend that client commitment is essentially connected with followership and buy aim in live-streaming advanced advertising. While cost is a critical mediator, its impacts become immaterial on their buy goals once purchasers become supporters. The outcomes feature the positive effects of social components, including likes, talks, visits, and openness time in friendly business towards conditional (buy) and non-value-based (followership) benefits.

Aanchal Badgaiyya and Prachi Shankarpale, proposed system is creating a data set from two different sources, for example, utilizing previously existing information and the information removed from Twitter and the information considered was film evaluations. For separating information from Twitter we used the normal language handling ideas utilizing python. When the total information created then the information gave to the BERT model to distinguish the different elements which assume a critical part in parallel grouping according to the evaluations as fortunate or unfortunate.

Jingxing Jiang, Zhubin Wang, Fei Fang, Binqiang Zhao, The proposed client aim forecast model was been generally utilized for the coupon distribution, ad and suggestion on Taobao stage, which incredibly further develop the client experience and shopping effectiveness, and advantage the gross product volume (GMV) advancement too.

Mariya Hendriksen, Pim Nauts addressed the review Based on investigation, they construct two sorts of indicators: (1) an indicator for unknown meetings that can precisely anticipate buy aim in mysterious meetings, beating a creation prepared indicator by more than 17.54% and (2) an indicator for recognized clients that involves meeting information as well as client history and accomplishes a 1 of 96.20% on held-out information gathered from a true retail stage.

Ganesh Dash, Justin Paul, they proposed concentrate on gave an observational trial of this more up to date model by examining every one of the four of its parts with consumer loyalty and buy goal. Utilizing underlying condition displaying to break down 508 forthcoming land first-time homebuyers, this study assesses the job of the parts of Marketing 4.0 in expanding consumer loyalty and affecting buy goals. V kumar and S kesharwani, they gave the investigation of extensive manual for responsiveness examination of model boundaries as to execution in forecast of item evaluations with cost subtleties by tracking down precision estimation.

HS Seippel Their proposed study examines AI models to foresee a buy, which is a pertinent use case as applied by an enormous German apparel retailer. Then, to contrasting models this concentrate further gives knowledge into the presentation distinctions of the models on successive snap stream and the static client information, by leading an expressive information examination and independently preparing the models on the various informational collections. The outcomes demonstrate that a Random Forest calculation is the most ideal for the forecast task, showing the best presentation results, sensible dormancy, offering conceivability and a high vigor.

### **III. SYSTEM ARCHITECTURE AND METHODOLOGY**

The proposed system contains following:

#### **3.1 Data Wrangling**

In this section of the report will load the data, check for cleanliness, and then trim and clean given dataset for analysis. Make sure go with the flow carefully and justify for cleaning and proper decisions.

#### **3.2 Pre-processing**

The data might have some null values to could cause problems. To obtain better outcome data need to be pre-processed so as to quality and efficiency of the algorithm increases. The outliers have to be removed and also variable issues need to be solved.

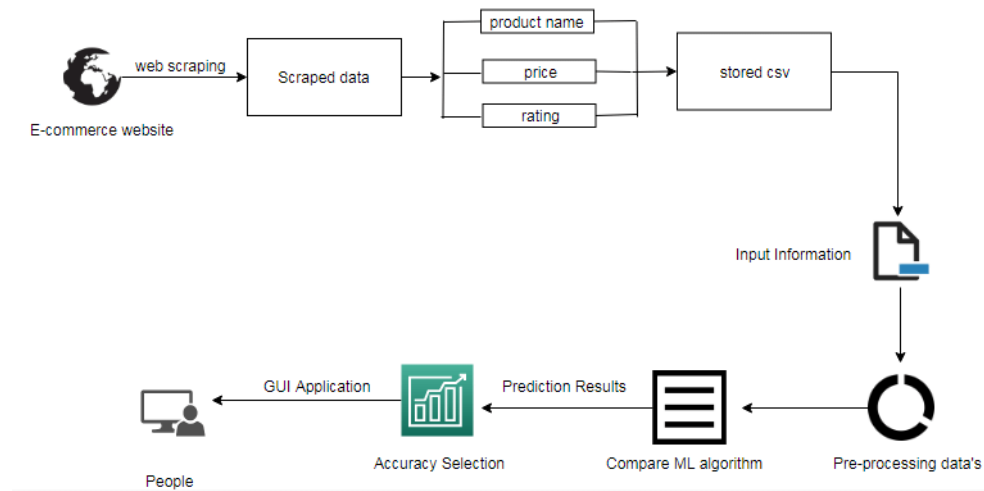


Figure 1: System Architecture

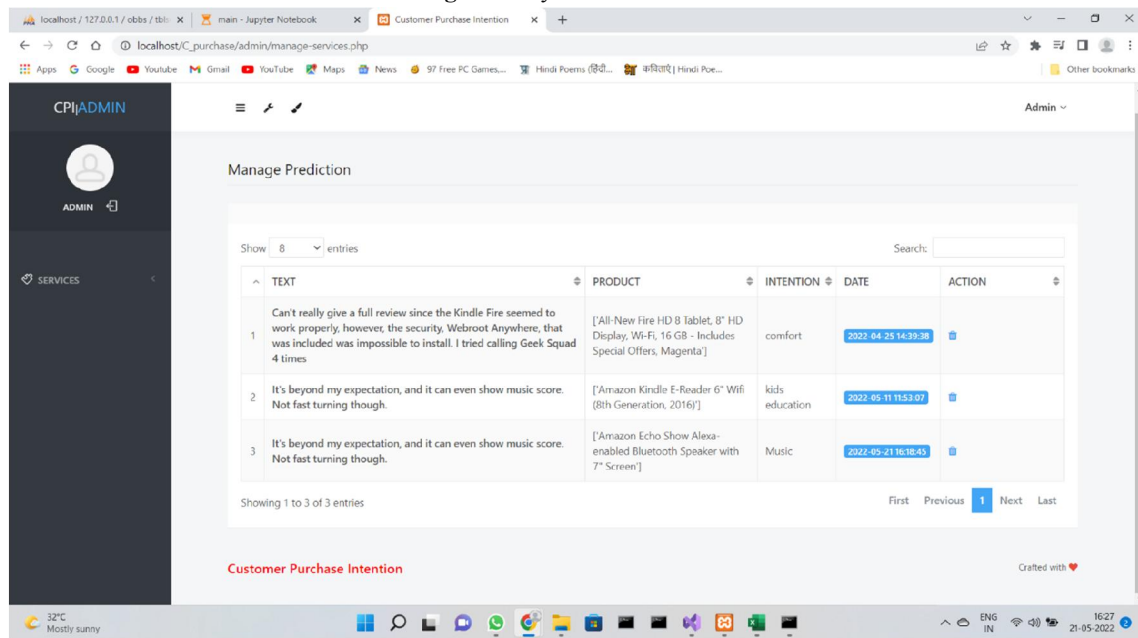


Figure 2: Output Expectations

#### IV. CONCLUSION

With the proposed system, we will be able to implement an online system that will help in the selling and buying of agricultural products with good cost approximation and quality in mind, as well as good quality processed food for the needful persons, all while using the required software flawlessly for farmer consumers, NGO, and hotels/farmer selling products, see through that the food or selling product doesn't go to dump and go to the needful persons.

#### REFERENCES

- [1]. Jingxing Jiang, Zhubin Wang, Fei Fang, Binqiang Zhao, "TPG-DNN: A Method for User Intent Prediction Based on Total Probability Formula and GRU Loss with Multi-task Learning", IEEE Conference 2020
- [2]. Mariya Hendriksen, Pim Nauts, Analyzing and Predicting Purchase Intent in E-commerce: Anonymous vs. Identified Customers IEEE Conference 2020

- [3]. Ganesh Dash, Justin Paul, "Marketing-to- Millennials: Marketing 4.0, customer satisfaction and purchase intention", Journal of Business Research, Elsevier 2020
- [4]. V kumar and S kesharwani "Online Customer Satisfaction and Loyalty in Apparel Purchase: A Review of Extant Literature", 2020
- [5]. HS Seippel, "Customer purchase prediction through machine learning", Electrical Engineering, Mathematics & Computer Science, 2018
- [6]. Sushant Kumar, Mikko Murphy "What drives brand love and purchase intentions toward the local food distribution system? A study of social media- based REKO (fair consumption) groups", 2021
- [7]. Erik Nettet, Ola Bergem, "Building chain loyalty in grocery retailing by means of loyalty programs – A study of the Norwegian case Journal of Retailing and Consumer Services 2021
- [8]. Nusrat Parvin; Sayaka Zaman; Samia Amin," E- cigarette Purchase Intention through Facebook social media: A Cross-Sectional Study", Science Publication 2021
- [9]. Clement Adoo," Customer engagement and purchase intention in live-streaming digital marketing platforms", Research Gate 2021
- [10]. Aanchal Badgaiyya and Prachi Shankarpale. Application of Sentiment Analysis Based on Hybrid Database of Movie Ratings" , International Research Journal of Engineering and Technology 2021