

E-COMMERCE PURCHASE INTENT PREDICTION USING USER BROWSING PATTERNS

¹K.Suma,²K.Paivasha,³G.Shiva kumar reddy,⁴G.Shravani

¹Assistant Professor, ²³⁴Students

Department of Computer Science and Engineering

Siddhartha institute of technology & sciences,narapally

sumakadari_cse@siddhartha.co.in,23TQ1A0541@siddhartha.co.in,
23TQ1A0554@siddhartha.co.in, 24TQ5A0508@siddhartha.co.in

ABSTRACT

Predicting customer purchase intent is a critical task for e-commerce platforms as it helps businesses understand user behaviour and improve marketing strategies. With the increasing growth of online shopping, large amounts of clickstream and browsing behaviour data are generated by users during their interaction with websites. Analyzing this behavioural data enables companies to identify potential customers who are more likely to make a purchase. This study proposes a machine learning-based approach to predict customer purchase intent using behavioural and session-based features obtained from the Online Shoppers Intention dataset. The proposed methodology includes data preprocessing, feature encoding, model training, and evaluation. Initially, the dataset is cleaned and categorical features are converted into numerical values using label encoding to make them suitable for machine learning algorithms. The processed dataset is then divided into training and testing sets to ensure reliable model performance evaluation.

A Random Forest Classifier is applied to capture complex relationships within the browsing behaviour data and predict whether a user will complete a purchase or leave the website without buying. The performance of the model is evaluated using metrics such as accuracy, confusion matrix, and classification report. The results demonstrate that the proposed model effectively predicts purchase intent and can assist e-commerce businesses in improving personalized marketing strategies, customer targeting, and overall decision-making processes.

1 INTRODUCTION

The rapid growth of e-commerce platforms has significantly transformed the way customers interact with online businesses. Every interaction that a user performs on a website, such as clicking on a product, browsing categories, or spending time on a particular page, generates valuable behavioural data. This data, commonly referred to as clickstream and browsing behaviour data, provides deep insights into customer interests and purchasing patterns. Businesses can utilize this information to understand customer behaviour and predict whether a visitor is likely to complete a purchase. Predicting customer purchase intent has become an important research area in the field of data science and machine learning. Traditional marketing

strategies rely on historical sales data, but modern machine learning approaches allow businesses to analyze real-time behavioural signals to make accurate predictions. By identifying potential buyers during their browsing sessions, companies can personalize recommendations, offer targeted promotions, and improve customer experience. This project focuses on developing a machine learning model to predict whether a customer visiting an e-commerce website will make a purchase. The model analyzes browsing behaviour features such as page visits, session duration, and product interactions. A Random Forest Classifier is used to train the predictive model due to its efficiency and ability to handle complex patterns in the data. The proposed system aims to help e-commerce businesses make data-driven decisions and improve conversion rates. The rapid growth of the internet and digital technologies has significantly transformed the global retail industry. E-commerce platforms have become one of the most popular ways for consumers to purchase products and services. Online shopping offers convenience, a wide range of product choices, and easy price comparisons, which attract millions of users daily. As a result, e-commerce websites generate a large amount of user interaction data through activities such as page visits, product searches, clicks, and browsing sessions. This behavioural data, commonly known as clickstream data, provides valuable insights into customer interests and online behaviour patterns. Understanding customer behaviour has become an essential task for e-commerce businesses in order to improve customer engagement and increase sales. However, one of the major challenges faced by online retailers is that a large number of visitors browse products without actually completing a purchase. Identifying potential customers who are likely to buy products can help businesses improve their marketing strategies, provide personalized recommendations, and increase conversion rates

II LITERATURE SURVEY

Alamuri, O. R. S. G., & Bondalapu, C. K. (2026). Predicting e-commerce purchase intention using machine learning. *ASEAN Journal of Scientific and Technological Reports*, 29(1), e260159-e260159.

Feng, J., & Liu, H. (2026). Next, browse or purchase—A sequential and interpretable deep learning approach for predicting customer purchase intention. *Information Sciences*, 123144. Suo, L., Xia, L.,

Chung, Y., & Kim, E. (2026). Layered Feature Engineering for E-Commerce Purchase Prediction: A Hierarchical Evaluation on Taobao User Behavior Datasets. *Computers, Materials & Continua*, 87(1). Abhichandani, D. (2026, January). Predicting High-Intent Purchases Using Precheckout Basket Signals to Improve Checkout Performance. In *2026 7th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)* (pp. 1006-1012).

IEEE. El Youbi, R., Loukili, M., Loukili, R., & El Aalouche, O. (2026). From Clicks to Conversions: Leveraging Apriori and Behavioural Segmentation in E-Commerce. *Statistics, Optimization & Information Computing*, 15(1), 848-858.

Li, N., & Gu, L. (2026). IFP-DeepFM: Integrating Improved FP-Growth and Attention-Based Deep Factorization Machines for User Purchase Behavior Modeling and E-commerce Recommendation. *Informatica*, 50(7).

Song, Y., Liu, L., Li, X., Zhou, W., & Fan, X. (2026, January). Attention prediction model for e-commerce pages based on deep learning and eye tracking data. In *International Conference on Advanced Electronics, Intelligent Technology, and Computing (AEITC 2025)* (Vol. 14010, pp. 23-27).

SPIE. Behara, N. Data Privacy Concerns, Platform Trust and Purchase Intentions in AI-Driven E-Commerce. Naim, A., Mohamad, M., Alnfiai, M. M., & Ayedee, N. (2026). Digital Shopping Behavior Model (DSBM): Exploring the Role of Personalization, Reviews, and Purchase Intent on Amazon. *International Journal of Consumer Studies*, 50(1), e70162.

Guda, V. R. (2026). Context-Aware Onboarding Flow Adaptation Using First-Run Prediction Models in E Commerce Applications. *International Journal of AI, BigData, Computational and Management Studies*, 7(1), 45 48.

Hu, S., Zhang, Y., Yu, X., & Yao, B. (2026, February). Research on User Purchase Behavior Prediction Based on Machine Learning Methods. In *2025 4th International Conference on Mathematical Statistics and Economic Analysis (MSEA 2025)* (pp. 361-368). Atlantis Press.

III SYSTEM ANALYSIS

E-commerce purchase intent prediction is a data-driven system that analyzes user browsing behavior to determine whether a user is likely to make a purchase. The system collects data such as page views, click patterns, time spent on product pages, cart additions, and navigation paths. By applying machine learning algorithms, the system identifies hidden patterns and correlations in user behavior. This helps businesses understand customer intent in real-time and enables personalized recommendations, targeted advertisements, and improved conversion rates. The system plays a crucial role in enhancing user experience and optimizing business strategies in competitive online markets.

Existing system

The existing system in e-commerce platforms mainly relies on basic analytics tools and rule-based approaches. These systems track user activity but lack the intelligence to accurately predict purchase intent. They often depend on historical data and simple metrics like page visits or previous purchases without analyzing deeper behavioral patterns. As a result, decision-making is reactive rather than proactive, leading to missed opportunities for customer engagement and lower conversion rates

DisAdvantages of Existing system

- Low prediction accuracy
- No real-time decision-making
- Limited understanding of user behavior

- Cannot handle large-scale dynamic data efficiently
- Lack of personalization

Proposed system

The proposed system uses advanced machine learning and deep learning techniques to predict purchase intent based on real-time user browsing patterns. It integrates data collection, preprocessing, feature extraction, and predictive modeling into a unified architecture. Algorithms such as Logistic Regression, Decision Trees, Random Forest, and Neural Networks are used to analyze user interactions and generate predictions. The system continuously learns from new data, improving its accuracy over time. It enables real-time recommendations, targeted marketing, and better decision-making, ultimately increasing sales and customer satisfaction.

Advantages of Proposed System

- High prediction accuracy
- Real-time analytics and prediction
- Better personalization of products
- Efficient handling of big data
- Improved customer experience

IV METHODOLOGY

Step 1: Data Collection

User interaction data is gathered from the e-commerce platform, including:

- Page views
- Clickstream data
- Time spent on each page
- Product views
- Add-to-cart actions
- Search queries

This data forms the foundation for prediction.

Step 2: Data Preprocessing

Collected data is cleaned and prepared:

- Remove missing or duplicate values
- Handle noisy/incomplete data
- Normalize and scale data

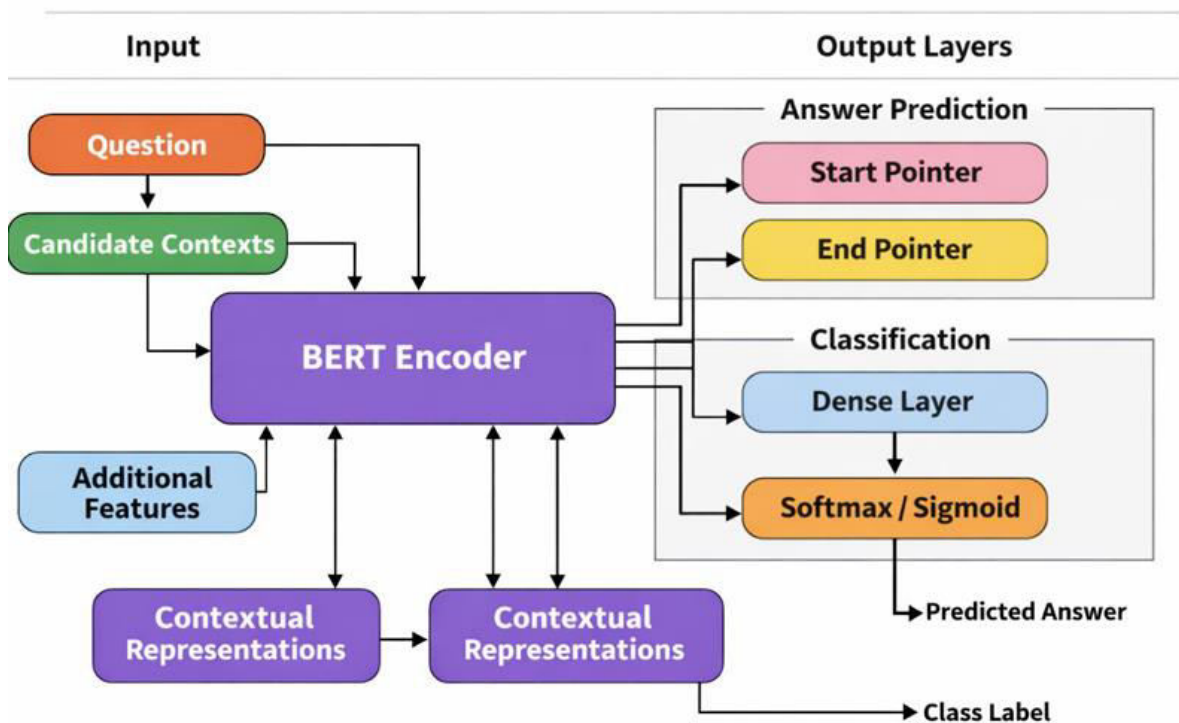
Step 3: Feature Engineering

- Important features are extracted from raw data:
- Number of clicks per session
- Session duration
- Bounce rate
- Product categories viewed
- Frequency of visits
- Cart abandonment rate
- These features represent user behavior patterns.

Step 4: Model Selection & Training

- Machine learning models are applied:
- Logistic Regression
- Decision Tree
- Random Forest
- Support Vector Machine (SVM)

System Architecture



The system architecture for e-commerce purchase intent prediction is designed as a multi-layered pipeline that processes user browsing data and converts it into actionable predictions. It integrates components such as data collection, preprocessing, feature engineering, machine learning models, and output delivery. The architecture ensures smooth data flow from user interaction to prediction results, enabling real-time decision-making and personalized user experiences.

V RESULTS & OUTPUT

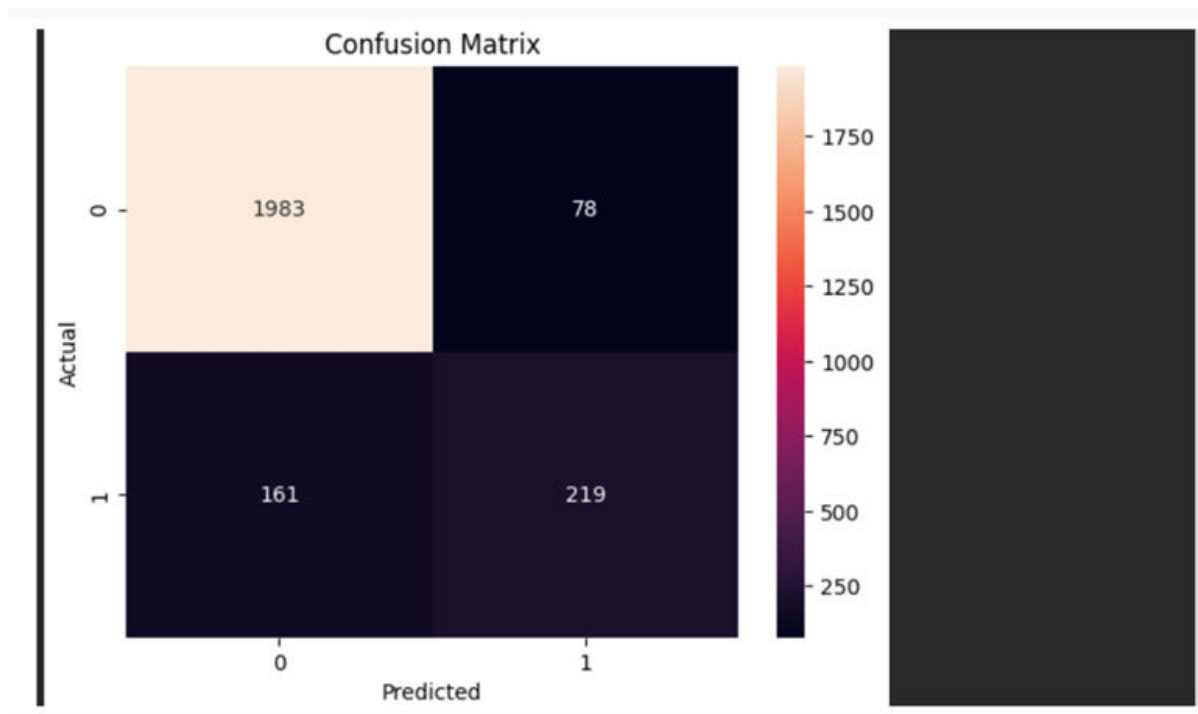
1	0.003774	0.021304	10.240430	0.0	Mar	4
2	0.012500	0.029167	0.000000	0.0	Dec	2
3	0.012121	0.023485	0.000000	0.0	Nov	3
4	0.000000	0.033333	0.000000	0.0	Mar	1

	Browser	Region	TrafficType	VisitorType	Weekend	Revenue
0	2	1	20	Returning_Visitor	False	False
1	2	3	8	Returning_Visitor	False	True
2	2	1	2	Returning_Visitor	False	False
3	2	3	2	Returning_Visitor	False	False
4	1	1	3	Returning_Visitor	False	False

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 12330 entries, 0 to 12329
Data columns (total 18 columns):
#   Column                                Non-Null Count  Dtype
---  ---                                -
0   Administrative                        12330 non-null  int64
1   Administrative_Duration               12330 non-null  float64
2   Informational                         12330 non-null  int64
3   Informational_Duration                12330 non-null  float64
4   ProductRelated                       12330 non-null  int64
5   ProductRelated_Duration               12330 non-null  float64
6   BounceRates                           12330 non-null  float64
7   ExitRates                             12330 non-null  float64
8   PageValues                            12330 non-null  float64
9   SpecialDay                            12330 non-null  float64
```

FIG.1 ONLINE SHOPPERS INTENTION

```
RandomForestClassifier
RandomForestClassifier(random_state=42)
```



VI CONCLUSION

In conclusion, predicting customer purchase intent plays an important role in improving the effectiveness of e-commerce platforms. With the increasing growth of online shopping, large amounts of clickstream and browsing behaviour data are generated by users during their interaction with websites. Analyzing this data using machine learning techniques allows businesses to better understand customer behaviour and identify potential buyers. In this project, a machine learning-based approach was implemented to predict whether a customer will complete a purchase based on their browsing behaviour. The Online Shoppers Intention dataset was used to train and test the predictive model.

Data preprocessing techniques such as data cleaning and label encoding were applied to convert categorical variables into numerical form suitable for machine learning algorithms. A Random Forest Classifier was used to build the predictive model due to its ability to handle complex patterns in large datasets and provide accurate predictions. The performance of the model was evaluated using metrics such as accuracy, confusion matrix, and classification report. The results demonstrate that the model can effectively identify patterns in user browsing behaviour and predict purchase intent. This system can help e-commerce businesses improve personalized marketing strategies, optimize product recommendations, and enhance overall decision making through data-driven insights.

REFERENCE

- [1] Kumar, R. D., Prudhviraaj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145-158). CRC Press.
- [2] Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on

Artificial Intelligence and Smart Environment (pp. 557-564). Cham: Springer Nature Switzerland.

[3] Sv satyakrishna, shirisha rangu ,bhargavi nalacheruve.(2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm

[4] Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.

[5] V. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, “Real-Time Object Detection in Drone Surveillance Using YOLOv5,” in Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.

[6] B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, “Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks,” in Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart Environment, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.

[7] R. D. Kumar, V. N. S. Manaswini, “Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology,” in Blockchain for Smart Cities, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.

[8] Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, “An advanced movie recommender using collaborative filtering and sentiment analysis,” International Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.

[9] Ravi Kumar Banoth, Ramana Murthy B V, “Automatic crop recommendation system using LightGBM and decision tree machine learning models,” Journal of Machine and Computing, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.

[10] Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, “Smart agriculture through IoT and machine learning for analyzing carbon footprints,” in Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE), Apr. 2025.

[11] Ravi Kumar Banoth, B. V. Ramana Murthy, “Soil image classification using transfer learning approach: MobileNetV2 with CNN,” SN Computer Science, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.