

PREDICTIVE DATA ANALYSIS OF E-COMMERCE SALES AND CUSTOMER BEHAVIOUR

**Prof. Dr. Kalpana Malpe¹, Mrs. Rittika R. Sardar², Mrs. Khushi S. Lunge³,
Mr. Shahzad S. Sheikh⁴, Mr. Mithilesh M. Gondane⁵**

1Assistant Professor , Department of Computer Science & Engineering, Govindrao Wanjari College of Engineering & Technology, (Dr.Babasaheb Ambedkar Technological University)

kmalpe@gmail.com

2Student , Department of Computer Science & Engineering, Govindrao Wanjari College of Engineering & Technology, (Dr.Babasaheb Ambedkar Technological University)

ritikasardar2005@gmail.com

3Student , Department of Computer Science & Engineering, Govindrao Wanjari College of Engineering & Technology, (Dr.Babasaheb Ambedkar Technological University)

Khushilunge22@gmail.com

4Student , Department of Computer Science & Engineering, Govindrao Wanjari College of Engineering & Technology, (Dr.Babasaheb Ambedkar Technological University)

romiosheikh08@gmail.com

5Student ,Department of Computer Science & Engineering, Govindrao Wanjari College of Engineering & Technology, (Dr.Babasaheb Ambedkar Technological University)

gondanemithilesh05@gmail.com

Abstract

Predictive data analysis in e-commerce is used to study past sales records, customer shopping activity, and market trends to understand future business possibilities. In the fast-growing online shopping industry, companies generate a large amount of data every day from customer searches, product views, cart activity, purchases, and feedback. If this data is properly analyzed, it can help businesses make better decisions related to sales planning, customer service improvement, and product management. The main purpose of predictive data analysis is to identify useful patterns from historical data and use them to estimate future outcomes. For example, businesses can predict which products will have high demand in upcoming seasons, which customers are likely to buy again, and which products customers may be interested in based on their previous activity. This helps businesses prepare stock in advance, reduce product wastage, and improve delivery planning. In this system, data is collected from e-commerce platforms including customer purchase history, browsing behavior, product ratings, seasonal sales trends, and customer preferences. This collected data is cleaned and processed to remove errors and unwanted information. After processing, the data is analyzed to identify customer buying habits and product demand patterns. The results help businesses provide personalized product suggestions, improve marketing strategies, and design better discount offers.

Predictive data analysis also helps in customer behavior understanding. Businesses can group customers based on their buying patterns, spending habits, and product interests. This helps companies target the right customers with the right products at the right time. It also improves customer satisfaction because customers receive product suggestions that match their interests.

Keywords: Predictive Analysis, E-Commerce Sales, Customer Behaviour, Data Analysis, Sales Forecasting, Customer Segmentation, Market Trends, Business Intelligence, Demand Prediction, Online Shopping Data.

I. Introduction

E-commerce has changed the way people shop. Customers now prefer online shopping because it is fast, convenient, and offers more choices. Because of this rapid growth, online businesses generate a large amount of data every day. This data includes customer searches, purchase history, product ratings, and browsing behaviour.

However, collecting data is not enough. Businesses need to understand this data to make better decisions. Predictive data analysis helps businesses study past data and identify patterns that can help predict future outcomes. For example, businesses can predict which products customers may buy next or which season will have higher sales.

Predictive analysis helps companies understand customer needs, improve product recommendations, and manage stock more efficiently. It also helps businesses design better marketing strategies. This improves customer satisfaction and increases company profit.

The goal of this project is to study how predictive data analysis can be used in e-commerce platforms to improve sales performance and understand customer behaviour. The system focuses on analysing customer activity, purchase patterns, and product demand trends.

II. Literature Survey

This book explains important data mining methods such as classification, clustering, regression, and association rule mining. It describes how very large datasets can be studied to discover useful patterns and hidden knowledge. The authors also explain predictive modelling methods that are commonly used for forecasting sales and studying customer purchase behaviour. The ideas discussed in this book provide a strong base for understanding predictive data analysis used in e-commerce applications.[1]

This book explains how data science techniques can be used in real business situations. It covers predictive analytics, customer grouping, and decision-making models in an easy and structured way. The authors explain how using data-based strategies can improve marketing performance and help businesses retain customers. These concepts are very useful for studying customer buying behaviour in e-commerce platforms.[2]

This book explains machine learning concepts such as probability models, regression methods, and classification techniques. It focuses on the mathematical background required for prediction systems. The book supports building accurate prediction models using both supervised and unsupervised learning approaches. These techniques are helpful for sales prediction and customer behaviour trend analysis in e-commerce systems.[3]

This book discusses statistical learning techniques including linear regression, decision trees, clustering, and boosting methods. It also explains the importance of testing and validating models before real use. The methods explained in this book help in creating reliable prediction systems for e-commerce sales forecasting and customer classification.[4]

This research paper explains the challenges and benefits of big data analysis. It focuses on scalable algorithms and distributed data processing methods for managing large datasets. Since e-commerce platforms produce huge amounts of customer data, the methods discussed in this paper are useful for predictive sales analysis and customer behaviour study.[5]

This paper explains the importance of analytics in modern marketing systems. It focuses on customer insights, personalized marketing methods, and predictive modelling techniques. The study shows how data-based strategies help improve customer engagement and increase sales performance. These ideas support predictive data analysis systems used in e-commerce platforms.[6]

The reviewed studies show that predictive data analysis is very important for understanding e-commerce sales trends and customer buying behaviour. Most research focuses on data mining methods, statistical models, prediction techniques, and big data processing for finding useful insights from customer data. The literature also discusses challenges such as prediction accuracy, data quality issues, and system scalability. These studies provide both theoretical knowledge and practical guidance for building a Predictive Data Analysis System for E-Commerce Sales and Customer Behaviour.

III. Proposed Work

The proposed work focuses on the design and development of a Predictive Data Analysis System for E-Commerce Sales and Customer Behavior that helps online businesses analyze past sales data and customer activity to estimate future trends. The system aims to provide a reliable, scalable, and user-friendly solution that supports business decision making, sales planning, and customer experience improvement in online shopping platforms.

In the proposed system, data is collected from e-commerce platforms including customer purchase history, product search records, browsing behaviors, product ratings, and seasonal sales data. This data is stored in a centralized database where it is processed and organized for analysis. Data cleaning techniques are applied to remove duplicate, missing, or incorrect values to improve analysis accuracy.

The system analyzes historical sales data to identify product demand trends and customer buying patterns. Based on these patterns, the system estimates future product demand and identifies high selling products. The system also studies customer purchase frequency, preferred product categories, and spending patterns to understand customer behavior.

The system supports personalized product suggestion features by matching customer interests with available products. It also helps businesses plan inventory levels by predicting which products may have higher demand in future periods. This helps reduce stock shortage and overstock problems.

The system is connected to a data storage and processing platform where sales and customer data can be accessed and analyzed in real time. A dashboard interface is proposed to allow business users to monitor sales performance, customer activity, and demand trends. The interface also helps generate reports that support marketing planning and sales strategy development.

Data security and privacy are considered important in the proposed system. Secure data storage and authorized access methods are used to protect customer information and business data. The system is designed to be flexible and expandable so that additional data sources and analysis features can be added in the future.

Overall, the proposed work presents a cost-effective and practical solution for e-commerce businesses to understand customer behavior, improve sales planning, and increase customer satisfaction. The system demonstrates how data analysis techniques can be used effectively to support business growth and market competition.

IV. Methodologies

The methodology of the Predictive Data Analysis System for E-Commerce Sales and Customer Behavior explains the step-by-step process followed for designing, developing, and implementing the system. The methodology ensures proper data collection, accurate analysis, and reliable prediction results for business decision making.

1. System Requirement Analysis :-The first step involves identifying system requirements. This includes selecting data sources such as e-commerce transaction data, customer browsing data, and product information. The system requirements focus on data accuracy, fast processing, scalability, and data security. Business needs such as sales forecasting, customer behavior understanding, and demand prediction are also considered.
2. Data Collection Method :-In this step, data is collected from e-commerce platforms including customer purchase history, product views, search history, ratings, and seasonal sales records. The collected data is stored in a centralized database for further processing and analysis.
3. Data Cleaning and Preparation:-The collected data may contain missing values, duplicate entries, or incorrect records. Data cleaning techniques are applied to remove unwanted data and improve data quality. After cleaning, data is formatted into a structured format suitable for analysis.
4. Data Processing and Storage :-After cleaning, the data is processed and stored properly in database tables. Data indexing and data organization methods are used to improve data retrieval speed and system performance. This step ensures efficient handling of large amounts of customer and sales data.
5. Pattern Analysis and Trend Identification :-The processed data is analyzed to identify useful patterns such as customer buying trends, product demand trends, seasonal sales patterns, and customer category preferences. These patterns help businesses understand customer behavior and market demand.
6. Sales Prediction Process :-Based on historical sales data and identified patterns, the system estimates future product demand and sales trends. The prediction results help businesses plan inventory levels and marketing strategies.
7. Customer Behavior Analysis :-Customer activity data is analyzed to understand customer interests, purchase frequency, and spending behavior. Customers can be grouped based on similar buying patterns. This helps businesses provide personalized product recommendations and targeted marketing.
8. Testing and Validation :-The system is tested using real and sample data to verify prediction accuracy and system reliability. Testing ensures correct data processing, accurate prediction results, and stable system performance.
9. Deployment and Performance Evaluation :-After successful testing, the system is deployed in real business environments. Performance metrics such as prediction accuracy, system response time, and data processing speed are evaluated. Continuous monitoring helps improve system performance.

V. Conclusion

The Predictive Data Analysis System for E-Commerce Sales and Customer Behavior successfully demonstrates how sales data and customer activity data can be used to support business decision making in online shopping platforms. By analyzing past sales records, customer purchase history, browsing behavior, and product demand patterns, the system helps businesses understand market trends and customer preferences more clearly.

The system provides several advantages such as improved sales planning, better inventory management, and enhanced customer experience. By estimating future product demand, businesses can maintain proper stock levels and reduce problems like stock shortage and overstock. Customer behavior analysis helps businesses design better marketing strategies and provide personalized product suggestions, which increases customer satisfaction and improves sales performance.

The system also supports data security and reliable data storage to protect customer and business information. The structured data analysis approach helps businesses respond quickly to market changes and

customer demand variations. The system is designed to be flexible and scalable, allowing future improvements and additional data sources to be integrated easily.

References

- [1] Data Mining: Concepts and Techniques, J. Han, M. Kamber, J. Pei, Morgan Kaufmann Publication, 2012.
- [2] Data Science for Business, F. Provost, T. Fawcett, O'Reilly Media Publication, 2013.
- [3] Pattern Recognition and Machine Learning, C. M. Bishop, Springer Publication, 2006.
- [4] The Elements of Statistical Learning, T. Hastie, R. Tibshirani, J. Friedman, Springer Publication, 2009.
- [5] Data Mining with Big Data, X. Wu et al., IEEE Transactions on Knowledge and Data Engineering, 2014.
- [6] Marketing Analytics for Data-Rich Environments, M. Wedel, P. K. Kannan, Journal of Marketing, 2016.
- [7] Introduction to Data Mining, P. Tan, M. Steinbach, V. Kumar, Pearson Publication, 2005.
- [8] Business Intelligence and Analytics, R. Sharda, D. Delen, E. Turban, Pearson Publication, 2014.
- [9] Predictive Analytics, E. Siegel, Wiley Publication, 2013.
- [10] Customer Relationship Management Concepts and Technologies, F. Buttle, Elsevier Publication, 2009.