ONLINE ORDER ANALYSIS AND PREDICTION E-COMMERCE

[1] G Praveen, [2] A Nithin, [3] B Abhinav, [4] S Hemanth, [5] B Anjali
[1]Assistant Professor, [2345]Students
Department of CSE
Siddhartha Institute of Technology & Sciences, Narapally
badaralaanil_cse@siddhartha.co.in, 23tq1a0562@siddhartha.co.in, 23tq1a0506@siddhartha.co.in, 23tq1a0563@siddhartha.co.in, 24tq5a0501@siddhartha.co.in,

Abstract

The Online Order Analysis and Prediction for E-Commerce system is designed to analyze e-commerce order data and predict future sales, customer orders, and purchasing trends using data analytics and machine-learning techniques. With the rapid growth of online shopping, e-commerce platforms generate large volumes of information related to customers, products, orders, payments, and deliveries. Analyzing this data can help businesses understand customer behavior, optimize inventory, and improve sales performance.

The proposed system collects and processes information such as order ID, customer details, product category, product price, quantity, order date, payment method, shipping location, delivery status, and total order value. The collected data is cleaned and transformed to handle missing values, duplicate records, and inconsistent information before analysis.

The system performs order-wise, product-wise, category-wise, customer-wise, and time-based analysis to identify important sales patterns. Machine-learning and statistical techniques can be applied to historical order data to predict future order volume, sales revenue, product demand, and customer purchasing trends. The system can also identify frequently purchased products and high-demand categories.

Interactive dashboards, charts, graphs, and reports are used to display total orders, revenue, average order value, popular products, monthly sales trends, and prediction results. These insights can help e-commerce businesses improve inventory planning, marketing strategies, sales forecasting, and customer management.

I. Introduction

The rapid growth of e-commerce has significantly changed the way people purchase products and services. Online shopping platforms generate a large amount of data every day through customer orders, product searches, payments, reviews, delivery information, and browsing activities. Analyzing this data helps businesses understand customer behavior, identify popular products, monitor sales performance, and improve their overall decision-making process. However, managing and interpreting large volumes of online order data manually can be difficult, time-consuming, and inefficient.

The **Online Order Analysis and Prediction – E-Commerce** system is designed to analyze historical online order data and generate useful insights from it. The system collects important information such as order details, product categories, quantities,

prices, customer information, order dates, and payment details. By processing this data, the system can identify sales trends, high-performing products, customer purchasing patterns, and variations in demand over different time periods.

A major feature of the proposed system is its ability to perform **prediction** using data analysis and machine learning techniques. Historical order information can be used to predict future sales, product demand, or expected order trends. These predictions can help e-commerce businesses plan their inventory, manage resources, prepare for changes in customer demand, and reduce the possibility of overstocking or stock shortages.

The system also provides a user-friendly interface through which administrators or business users can view analyzed results in the form of dashboards, charts, graphs, and reports. Important performance indicators such as total orders, total sales, frequently purchased products, monthly sales trends, and customer-related statistics can be displayed clearly. This makes complex order data easier to understand and helps users identify important business patterns quickly.

II. Literature Survey

The rapid development of e-commerce has created a large amount of customer, transaction, product, and browsing data. Researchers have increasingly explored machine learning and data analytics techniques to convert this data into useful business insights. A systematic literature review by Policarpo et al. examined machine-learning applications in e-commerce and identified major application areas including recommendation systems, purchase prediction, customer relationship management, and profit-related analysis. The study shows that machine learning can support both customer-oriented services and business decision-making.

Cirqueira et al. studied customer purchase behavior prediction in e-commerce and proposed an analytical framework for understanding predictive approaches. Their review categorized purchase prediction into three major tasks: predicting customer intentions, buying sessions, and purchase decisions. The study highlights the importance of behavioral data and predictive analytics for understanding future customer actions in digital retail environments.

Koehn, Lessmann, and Schaal investigated the use of deep learning for predicting online shopping behavior from clickstream data. Their research used recurrent neural networks such as LSTM and GRU to analyze customer interactions and predict shopping-related outcomes. The study demonstrates that clickstream information can provide valuable signals for predicting customer behavior and supporting targeted e-commerce interventions.

Chaudhuri et al. focused on predicting whether customers on an e-commerce platform would make a purchase. Their study applied deep neural networks to customer purchase data and reported that deep learning models could provide more accurate predictions than several conventional machine-learning techniques. The research also identified customer and platform characteristics that can contribute significantly to purchase prediction.

Lin et al. proposed a machine-learning-based e-commerce purchase prediction model using historical customer information. Their approach used a Random Forest decision model to extract features from historical purchase data and predict customer purchasing behavior. The research indicates that machine learning can be useful for identifying potential purchasing patterns and supporting e-commerce sales improvement.

Sharma, Chakraborti, and Jha investigated sales prediction in the Amazon marketplace in India using machine-learning techniques. Their work emphasized that accurate demand prediction is important for supply-chain management because it can help businesses avoid overproduction and underproduction. The study also considered factors such as product type, price, discounts, delivery options, and review sentiment when predicting customer demand.

Recent research has also examined the use of machine learning and deep learning specifically for e-commerce sales forecasting. A 2025 review discussed the application of deep-learning approaches to sales forecasting and highlighted their ability to model nonlinear relationships and long-term dependencies in complex e-commerce data. Such techniques can support more accurate forecasting of future sales and demand.

A systematic review of machine-learning and deep-learning techniques in e-commerce identified several commonly used methods, including Support Vector Machines, Decision Trees, Random Forests, neural networks, and recurrent neural networks. It also identified sales prediction and customer purchase behavior prediction as important research areas, while noting challenges such as imbalanced data, overfitting, generalization, interpretability, and personalization.

Research Gap

From the reviewed literature, it can be observed that many existing studies concentrate on a specific prediction task, such as customer purchase intention, clickstream behavior, or sales forecasting. There is an opportunity to develop an integrated system that combines online order analysis, visualization of historical sales patterns, customer/product analysis, and future order prediction in a single platform. The proposed Online Order Analysis and Prediction – E-Commerce system addresses this gap by using historical order data to generate meaningful analytical insights and predictive results that can support better inventory planning, sales forecasting, and business decision-making.

III. System Analysis

The **Online Order Analysis and Prediction – E-Commerce** system is designed to analyze online transaction data and predict future sales or order trends. E-commerce platforms generate large amounts of data through customer orders, products, prices, quantities, payment methods, and order dates. Analyzing this information manually is difficult when the volume of data increases. The proposed system studies historical order data to identify important patterns and trends. It can calculate total orders, total sales, product performance, customer purchasing behavior, and category-wise sales. Data preprocessing is performed to remove missing, duplicate, and inconsistent

records before analysis. Statistical analysis and visualization techniques are used to represent important information through charts and graphs. Machine learning techniques can then be applied to historical data for predicting future sales or order quantities. The prediction results can help businesses understand expected demand and prepare their operations accordingly. The system also provides meaningful information to support inventory and sales planning. Overall, system analysis helps transform raw e-commerce transaction data into useful information for business decision-making. Research shows that machine learning is increasingly used in e-commerce for purchase prediction, customer behavior analysis, and business optimization.

Existing System

In the existing system, e-commerce order information is generally stored in databases or spreadsheets and is mainly used for recording completed transactions. Business users may depend on manual reports or basic database queries to understand sales performance. Historical order data can be analyzed using simple calculations such as total sales, number of orders, and product-wise revenue. However, traditional systems generally focus more on recording and reporting past transactions than predicting future demand. When the amount of order data becomes large, manually analyzing the information becomes time-consuming and difficult. Identifying customer purchasing patterns may also require considerable effort. Many traditional systems do not provide advanced visualization or automated prediction capabilities. Decision-making may therefore depend heavily on previous experience and manually prepared reports. Changes in customer demand can be difficult to identify quickly using static reports. Traditional forecasting methods may also have limitations when e-commerce data contains complex relationships and changing customer behavior. Modern research indicates that machine learning can provide additional capabilities for analyzing e-commerce data and predicting sales and purchasing behavior.

Disadvantages of Existing System

- Manual analysis of large amounts of order data is time-consuming.
- Limited ability to predict future sales and customer demand.
- Difficult to identify hidden patterns in large datasets.
- Heavy dependence on manual reports and calculations.
- Static reports may not provide real-time business insights.
- Limited visualization of sales and order trends.
- Difficulty in identifying changing customer purchasing behavior.
- Higher possibility of human errors during manual analysis.
- Inventory planning may become less accurate without demand prediction.
- Traditional methods may not handle complex relationships in data effectively.
- Difficult to compare multiple products and categories efficiently.
- Decision-making may depend mainly on historical observations.

Proposed System

The proposed Online Order Analysis and Prediction – E-Commerce system provides an integrated platform for analyzing historical order data and predicting future trends. The system accepts order information containing details such as order date, product,

category, quantity, price, customer information, and payment details. The collected data is first cleaned and transformed into a suitable format for analysis. The system performs exploratory data analysis to identify sales patterns, product performance, customer behavior, and order trends. Interactive charts and graphs can be used to present important information in an understandable form. Machine learning algorithms can be trained using historical order data to predict future sales or demand. Different algorithms can be evaluated using suitable performance measures to select an effective prediction model. Research on e-commerce sales prediction has explored methods such as Random Forest, Decision Tree, XGBoost, neural networks, and other machine-learning approaches. The prediction module can help businesses estimate upcoming demand and improve inventory planning. The system reduces dependence on manual calculations and makes analytical information easier to access. Overall, the proposed system combines data preprocessing, analysis, visualization, and prediction to support efficient and data-driven e-commerce decision-making.

Advantages of Proposed System

- Automates the analysis of online order data.
- Reduces manual calculations and human errors.
- Provides clear sales and order trends.
- Helps identify high-performing products.
- Helps understand customer purchasing patterns.
- Provides graphical representation of important information.
- Predicts future sales or order demand.
- Supports better inventory management.
- Helps businesses make data-driven decisions.
- Saves time when analyzing large datasets.
- Allows comparison of products, categories, and time periods.
- Can improve sales planning and operational efficiency.
- Can be extended with advanced machine-learning algorithms.

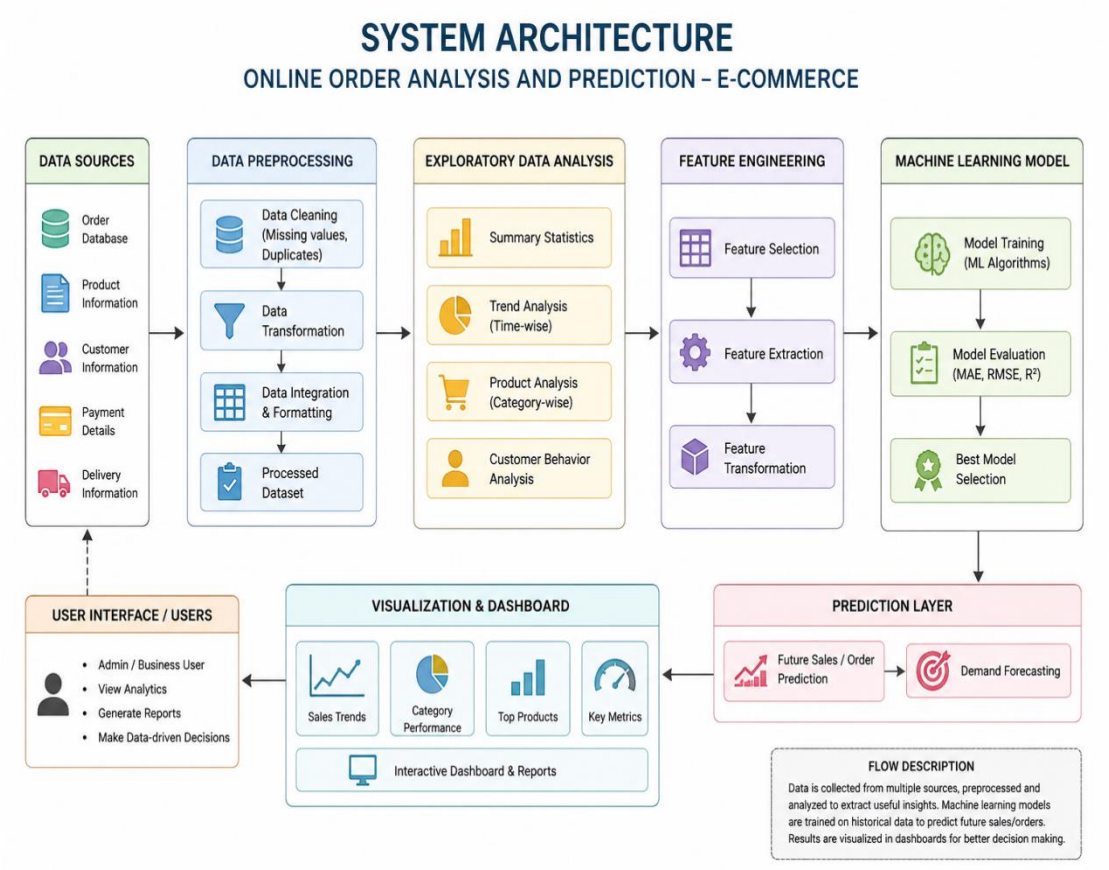
IV. Methodology

The methodology of the Online Order Analysis and Prediction – E-Commerce system consists of several sequential stages. First, the required e-commerce order dataset is collected from available transaction records. The collected data may contain attributes such as order ID, product, category, quantity, price, customer details, payment method, and order date. In the next stage, data preprocessing is performed to handle missing values, duplicate records, incorrect entries, and inconsistent formats. After preprocessing, exploratory data analysis is conducted to understand the characteristics of the dataset and identify important sales patterns. Statistical measures and visualization techniques are used to analyze order frequency, revenue, product performance, and time-based trends. Relevant features are selected and transformed into a suitable format for machine-learning algorithms. The dataset is then divided into training and testing portions for model development and evaluation. Suitable prediction algorithms such as Linear Regression, Decision Tree, Random Forest, XGBoost, or other forecasting models can be considered. Model performance can be evaluated using metrics such as MAE, RMSE, and R^2, which are commonly used in sales prediction research. Finally, the selected model is used to generate predictions,

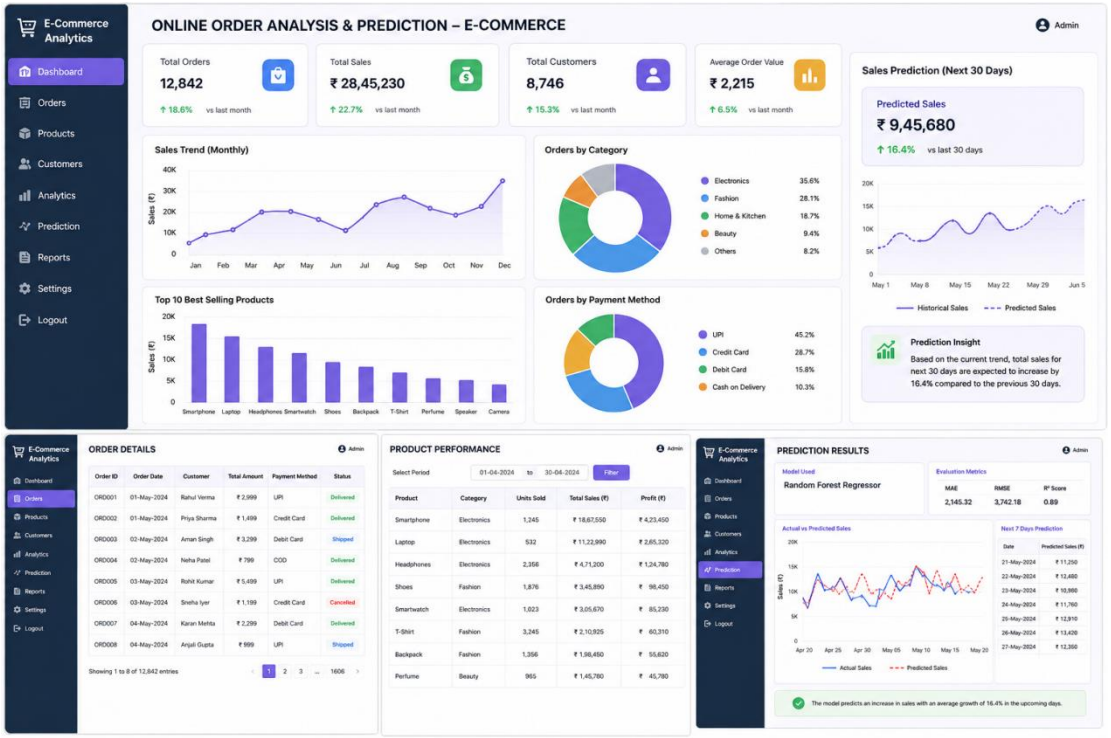
while the analytical results are presented through a user-friendly dashboard or reporting interface.

System Architecture

The system architecture represents the complete flow of the Online Order Analysis and Prediction – E-Commerce system from data collection to final results. The first layer is the Data Source Layer, which contains historical online order information such as customers, products, quantities, prices, and order dates. The data is transferred to the Data Preprocessing Layer, where missing values, duplicate records, and inconsistent data are handled. The cleaned information is then passed to the Data Analysis Layer, where statistical calculations and exploratory analysis are performed. The Feature Engineering Layer selects and transforms relevant attributes required for prediction. The processed data is provided to the Machine Learning Layer, where prediction models are trained using historical order information. The trained models are evaluated using appropriate performance metrics, and the best-performing model is selected for prediction. The Prediction Layer generates future sales or order-demand estimates. The Visualization Layer presents sales trends, product performance, order statistics, and prediction results using charts, graphs, and dashboards. Finally, the User Interface Layer allows administrators or authorized users to view the analytical results and predictions. This layered architecture provides a systematic flow from raw e-commerce data to meaningful insights and predictive results. Modern e-commerce research similarly emphasizes combining data analysis and machine learning to support business decisions and forecasting.



V. Result and Output



VI. Conclusion

The Online Order Analysis and Prediction – E-Commerce system provides an effective solution for analyzing large volumes of online order data and generating meaningful business insights. The system helps identify important patterns related to sales, products, customers, categories, and order trends. Through data preprocessing and analysis, raw transaction data is converted into useful and understandable information. The visualization module presents analytical results using charts, graphs, tables, and dashboards, making it easier for users to understand business performance. The prediction module uses historical order data and machine-learning techniques to estimate future sales or order demand. These predictions can support better inventory planning, sales forecasting, and resource management. The system also reduces the effort and time required for manual data analysis and reporting. By combining data analysis with predictive capabilities, businesses can make more informed and data-driven decisions. The proposed system can help identify high-performing products and changing customer demand more efficiently. It can also be extended in the future with real-time data processing, advanced forecasting algorithms, recommendation systems, and automated alerts. Overall, the project demonstrates how data analytics and machine learning can be effectively applied to e-commerce order management and prediction, improving operational efficiency and supporting better business decision-making.

References

[1] Kumar, R. D., Prudhviraj, 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.