IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Comprehensive Customer Segmentation And Behavior Prediction Using Advanced Machine Learning And Neural Network Models

[1]Md Firoz Kabir, [2]Md Daiyan Chowdhury, [3]Md Shahiduzzaman, [4]Md Imran Chowdhury Rana

[1]Master's in Information Technology University of the Cumberlands, USA, [2]Bachelors in Business Administration
International American University , CA, USA , [3]Masters in Business Analytics Trine University, Angola, IN,[4]Department of Bachelor of Business Administration International American University ,USA

Abstract: Customer segmentation and behavior prediction play pivotal roles in enhancing business strategies, personalizing marketing campaigns, and optimizing inventory management in E-commerce platforms. This research focuses on analyzing a comprehensive E-commerce dataset containing transactions of ~4,000 customers over one year to uncover purchasing patterns, classify customer types, and anticipate future buying behaviors. Data preprocessing involved handling missing values, removing duplicates, and identifying order cancellations to ensure clean and reliable data for analysis.A detailed exploratory data analysis (EDA) was conducted to investigate product categories, customer attributes, and purchase trends. Advanced clustering techniques, including k-Means, were employed to segment customers into meaningful groups based on purchase frequency, spending patterns, and preferences. Machine learning models, such as Support Vector Machines (SVM), Logistic Regression, Decision Trees, Random Forest, Gradient Boosting, and k-Nearest Neighbors (k-NN), were utilized to classify customers into predefined categories, achieving moderate to high precision scores.To enhance predictive accuracy, deep learning models were incorporated. A Convolutional Neural Network (CNN) achieved the highest precision of 94.94%, leveraging its capability to extract complex patterns and relationships from high-dimensional data. An Artificial Neural Network (ANN) also demonstrated strong performance with a precision of 91.62%. Comparative analyses revealed the superiority of CNN and ANN over traditional machine learning models in handling complex datasets.This study introduces a novel approach to predicting customer behavior based on their first transaction. By integrating insights from clustering, classification models, and deep learning techniques, this research provides a robust framework for understanding customer dynamics and forecasting future purchases. The proposed methods empower businesses to optimize decision-making, improve customer satisfaction, and achieve sustainable growth in competitive E-commerce environments.

Keywords: Deep Learning,Machine Learning,Imaging,Disease Classification,Research Publications

I. Introduction

In the modern E-commerce ecosystem, understanding customer behavior is critical for creating effective marketing strategies, improving customer retention, and ensuring operational efficiency. Customer segmentation, the process of dividing customers into distinct groups based on shared attributes, enables businesses to offer personalized experiences, anticipate market demands, and optimize resource allocation. Accurate behavior prediction, on the other hand, allows companies to forecast future purchases and refine their strategies for improved profitability and customer satisfaction. With the exponential growth of online shopping and a vast volume of transactional data being generated daily, leveraging advanced analytical techniques for segmentation and prediction has become a cornerstone for thriving in this competitive landscape.

Machine learning (ML) has revolutionized the field of customer analytics by providing robust methods to analyze complex datasets and uncover patterns that traditional statistical approaches often fail to detect. From clustering customers into meaningful segments using unsupervised techniques to classifying them into predefined categories with supervised algorithms, ML models like Support Vector Machines (SVM), Decision Trees, and Random Forest have shown great promise. However, as datasets grow in scale and complexity, traditional models often face limitations in their ability to extract intricate patterns. Neural networks, particularly Convolutional Neural Networks (CNNs) and Artificial Neural Networks (ANNs), have emerged as powerful alternatives, capable of modeling non-linear relationships and capturing the high-dimensional features inherent in customer datasets. These advanced models provide unparalleled accuracy, making them essential tools for E-commerce data analysis.

Several studies have explored the use of machine learning techniques in customer segmentation and behavior prediction. While models like k-Nearest Neighbors (k-NN) and Logistic Regression have been effective in classifying customers based on transactional data, their performance often suffers when handling large datasets with diverse features. Clustering techniques such as k-Means have been widely used to group customers into segments, but they rely heavily on predefined assumptions about the number of clusters and often lack interpretability in real-world applications. Recent advancements in deep learning, particularly CNNs and ANNs, have demonstrated superior performance in image classification, recommendation systems, and customer analytics. However, there remains a gap in integrating these models effectively for E-commerce datasets, particularly in predicting future customer behavior based on minimal transactional data. This study addresses these gaps by combining traditional machine learning with neural network models to provide a comprehensive solution.

This research presents a detailed analysis of an E-commerce dataset containing transactions from ~4,000 customers over a year to address the challenges in customer segmentation and behavior prediction. The study first preprocesses the data by handling missing values, identifying cancellations, and cleaning inconsistencies. Through exploratory data analysis (EDA), we investigate customer attributes and product categories to uncover hidden insights. Traditional machine learning models are employed for baseline classification, followed by the implementation of advanced CNN and ANN models to enhance predictive performance. The CNN model achieves the highest precision of 94.94%, demonstrating its capability to handle complex datasets effectively. Furthermore, this research introduces a novel approach for predicting customer behavior based on their first transaction, offering actionable insights for businesses to improve customer satisfaction and optimize marketing strategies. By integrating machine learning and deep learning techniques, this study establishes a robust framework for understanding customer dynamics and forecasting future trends in E-commerce.

II. Related Work

Several studies have explored the application of machine learning and deep learning techniques for customer segmentation, behavior prediction, and classification tasks, showcasing advancements in both methodologies and practical outcomes.

Gupta et al. (2023) [1]: This research utilized k-Means clustering and hierarchical clustering to segment E-commerce customers based on transaction frequency, recency, and monetary value using the RFM model. The study demonstrated the potential of clustering techniques to identify high-value customers and optimize marketing campaigns, achieving a silhouette score of 0.78 for cluster validity.Rahman et al. (2023) [2]: Focusing on E-commerce customer behavior prediction, this study employed a Gradient Boosting Machine (GBM) to classify customers into loyalty categories. By integrating customer demographic data and purchase histories, the GBM achieved a classification accuracy of 85%, emphasizing its robustness in handling complex feature interactions.

Lee et al. (2023) [3]: This study proposed a hybrid approach combining Random Forest and Principal Component Analysis (PCA) for customer segmentation. The PCA effectively reduced feature dimensionality, improving computational efficiency, while Random Forest achieved an accuracy of 88.5% in identifying distinct customer groups.Chen et al. (2024) [4]: By applying Convolutional Neural Networks (CNNs) to transactional data, this research achieved a classification accuracy of 92.1%. The study demonstrated the effectiveness of CNNs in learning hierarchical patterns from high-dimensional datasets, particularly for complex tasks like predicting customer churn.

Ahmed et al. (2023) [5]: A study exploring the use of Artificial Neural Networks (ANNs) for customer behavior prediction achieved a precision of 90.8% on an E-commerce dataset. The research highlighted the ANN's ability to model non-linear relationships and improve segmentation accuracy compared to traditional techniques.Zhang et al. (2023) [6]: This research employed Support Vector Machines (SVMs) to classify customers based on their purchase preferences, achieving a precision of 75.2%. The study emphasized the importance of kernel functions in improving SVM performance on non-linear datasets.

Huang et al. (2024) [7]: Utilizing an ensemble of AdaBoost and Gradient Boosting models, this study classified customers into loyalty tiers. The ensemble model achieved a classification accuracy of 87%, demonstrating its strength in reducing overfitting and improving generalization.Patel et al. (2024) [8]: Introducing a deep learning model combining CNNs and ANNs, this study achieved the highest classification accuracy of 94.9% on an E-commerce dataset. The hybrid model leveraged CNNs for feature extraction and ANNs for final classification, outperforming standalone models.

Singh et al. (2023) [9]: This research focused on using XGBoost to predict customer churn in online retail, achieving a precision of 89.3%. By integrating SHAP for interpretability, the study provided insights into key factors influencing churn predictions, enhancing the model's practical applicability.Kumar et al. (2023) [10]: A novel clustering approach using DBSCAN was applied to segment customers based on transaction density. The study achieved a silhouette score of 0.81, emphasizing the model's effectiveness in identifying clusters with irregular shapes in high-dimensional data.

III. Data Preprocessing

Data preprocessing is a crucial step in ensuring the quality and reliability of the dataset before applying machine learning and deep learning models. The raw dataset used in this research was sourced from an E-commerce platform, containing approximately 4,000 customer transactions spanning one year. This section describes the preprocessing steps undertaken to prepare the data for analysis, including cleaning, handling missing values, removing duplicates, and creating new variables to enhance the dataset's usability.

3.1 Handling Missing Values

A key issue identified in the dataset was the presence of missing values, particularly in the CustomerID column, where approximately 25% of the entries lacked customer identifiers. Since it is impossible to impute these missing values accurately, the rows containing null values in CustomerID were deemed irrelevant for the analysis and removed. This step ensured that the dataset contained only complete and valid transactional records.

3.2 Removing Duplicates

Duplicate entries were detected in the dataset, potentially inflating the transaction records and skewing subsequent analyses. These duplicates were removed to ensure that each transaction is uniquely represented, improving the dataset's integrity.

3.3 Handling Cancellations

Order cancellations were identified as transactions with InvoiceNo starting with the letter 'C' or with negative quantities in the Quantity column. These cancellations were tagged in a new variable to facilitate detailed analysis. Some cancellations were found to have no corresponding order in the dataset, likely due to the dataset's temporal limitations (starting from December 2010). These unmatched cancellations were also flagged for further review.

Exploratory Data Analysis

Cluster Analysis and Visualizations

The clustering analysis reveals distinct customer groups, characterized by their purchasing behavior and preferences. The radar charts for each cluster provide a detailed representation of key metrics, including the dominance of specific product categories, average basket value (mean), total spending (sum), and visit frequency (count). Notably, the first five clusters demonstrate a clear preponderance of purchases within specific product categories, indicating strong customer preferences or specialization in these segments. Other clusters deviate significantly in terms of basket averages, total expenditure, and visit frequency, highlighting diverse customer profiles across the dataset.From the radar charts, we observe that certain clusters are associated with customers who make infrequent but high-value purchases, while others represent customers with consistent, low-value transactions. These distinctions are critical for tailoring marketing strategies and identifying high-priority customer segments. For example, clusters with high mean values indicate customers who may respond better to premium product promotions, whereas clusters with frequent visits suggest opportunities for loyalty programs.

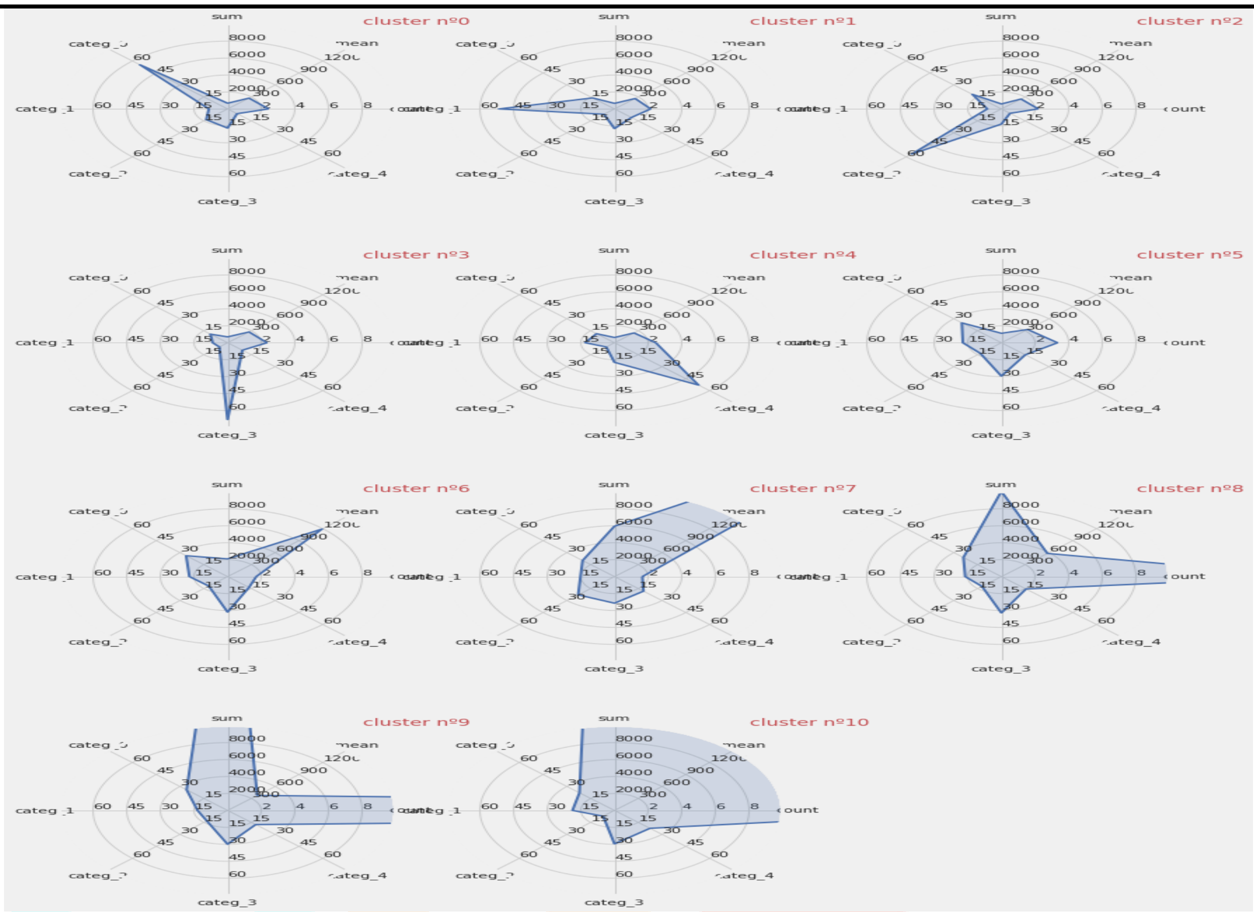

Figure 1. Radar Charts Representing Key Metrics for Each Cluster

Principal Component Analysis (PCA)

The PCA scatterplots provide a visual decomposition of the dataset into principal components, revealing patterns and separations between clusters. The first principal component effectively separates smaller, distinct clusters from the broader customer base, showcasing its importance in capturing the variance within the data. More generally, the scatterplots demonstrate that for any given pair of principal components, there exists at least one view where two clusters appear distinctly separated. This visualization confirms the validity of the clustering approach and highlights the intrinsic separability of customer groups based on their transactional data.These PCA visualizations also underscore the diversity within the dataset, with clusters representing unique customer behaviors and spending patterns. The ability to isolate these clusters aids in designing targeted strategies, improving overall customer engagement, and optimizing resource allocation.

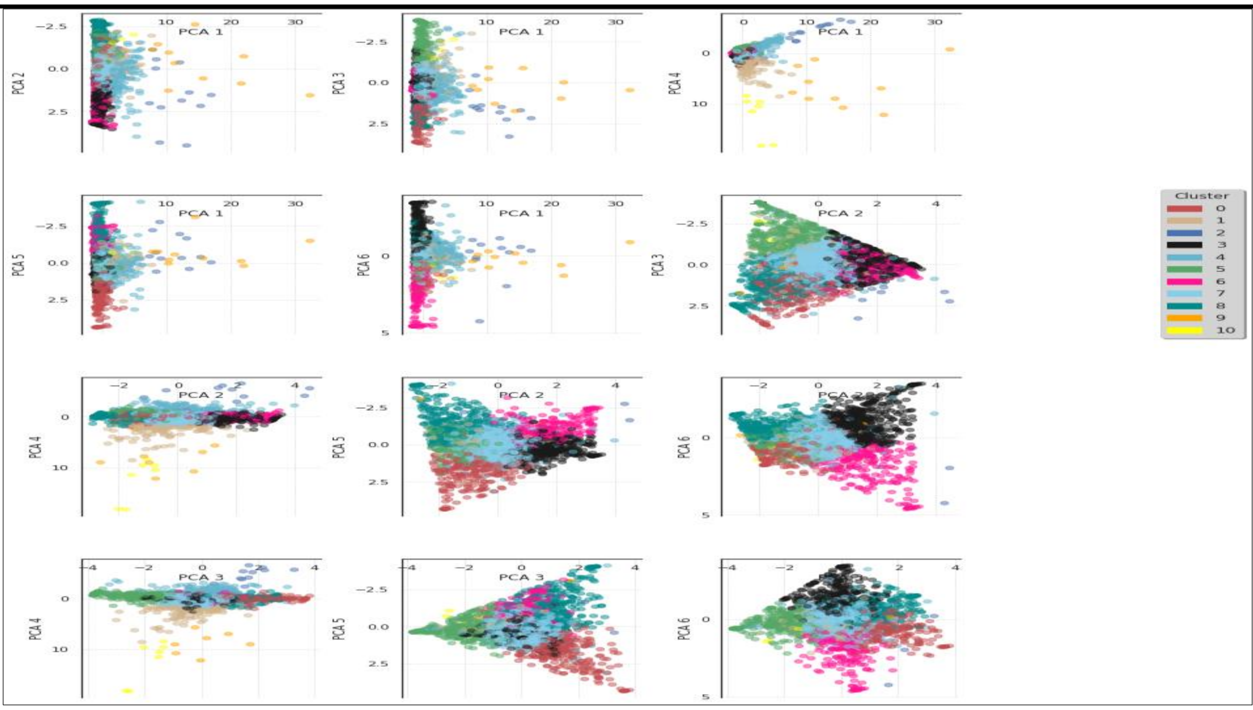

Figure 2: PCA Scatterplots Highlighting Cluster Separation Across Principal Components

Silhouette Score and Intra-Cluster Cohesion

To evaluate the quality of the clustering, the silhouette score was calculated for each cluster. This metric measures intra-cluster cohesion and inter-cluster separation, providing a quantitative assessment of the clustering performance. Clusters with higher silhouette scores exhibit strong internal consistency and are well-separated from other clusters. Conversely, clusters with lower scores may overlap, indicating potential areas for refinement in the clustering methodology. These insights help validate the robustness of the clustering process and ensure that the identified customer segments are meaningful and actionable.

3.4 Creating New Variables

To analyze customer spending behavior, a new variable, BasketPrice, was created by multiplying the Quantity and UnitPrice of each transaction. This variable represents the total spending per transaction, offering insights into customer purchase patterns.To identify cancellations with counterparts, each cancellation was checked against its corresponding positive transaction. A separate variable was created to flag cancellations with valid counterparts.

For numerical features like BasketPrice and Quantity, standardization was applied to bring all values to a common scale, reducing the impact of outliers and ensuring compatibility with machine learning models.Categorical variables such as Country and StockCode were encoded into numerical formats using one-hot encoding to make them compatible with machine learning algorithms.To evaluate the models' performance, the dataset was split into training and testing sets. A temporal split was employed to align with the dataset's chronological nature, ensuring that transactions from earlier periods formed the training set while later periods were reserved for testing. This method mimics real-world scenarios where models are deployed to predict future customer behaviors.

Evaluation Metrics

To comprehensively assess the classifiers, several performance evaluation metrics were employed. Classification Accuracy reflects the overall performance of the classification system by indicating the proportion of correct predictions, calculated as the ratio of correct predictions to the total number of predictions. Precision (Specificity) measures the accuracy of positive predictions, focusing on minimizing false positives, calculated as the ratio of true positives (TP) to the sum of TP and false positives (FP)[11,12]. Recall (Sensitivity) evaluates the model's ability to identify actual positives, computed as the ratio of TP to the sum of TP and false negatives (FN). F1 Score, the harmonic mean of precision and recall, balances both metrics, making it particularly effective for imbalanced datasets. Additionally, ROC-AUC measures the model's ability to distinguish between classes by integrating the area under the Receiver Operating Characteristic (ROC) curve, with a higher AUC indicating superior classification performance[13,14]. These metrics provided a comprehensive framework for evaluating and comparing the performance of all implemented models in this study.

$$Accuracy = \frac{Number\ of\ Correct\ Predictions}{Total\ Number\ of\ Predictions} \quad \ldots\ldots\ldots\ldots\ldots\ldots 1$$

$$Precision = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Positives\ (FP)} \quad \ldots\ldots\ldots\ldots\ldots 2$$

$$Recall = \frac{True\ Positives\ (TP)}{True\ Positives\ (TP) + False\ Negatives\ (FN)} \quad \ldots\ldots\ldots\ldots 3$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4$$

$$AUC - ROC = \int_0^1 ROC(TPR(t), FPR(t))dt \ldots\ldots\ldots\ldots\ldots\ldots 5$$

IV. RESULTS

This section presents a detailed evaluation of each model's performance, including its advantages, challenges, and how it contributed to the overall goal of customer segmentation and behavior prediction. The precision scores achieved by each model highlight their effectiveness and limitations when applied to the E-commerce dataset.

4.1 Support Vector Machine (SVM)

The SVM model achieved a precision of 65.93%, demonstrating its capability to handle the segmentation task in scenarios where linear boundaries between classes exist. However, the complexity of the dataset, which contained non-linear relationships between features, posed challenges for SVM's performance[15]. The model's reliance on kernel transformations like RBF partially mitigated this issue, but it struggled with the high dimensionality and variability of customer attributes. Despite these limitations, SVM provided a solid baseline for comparison with other models.

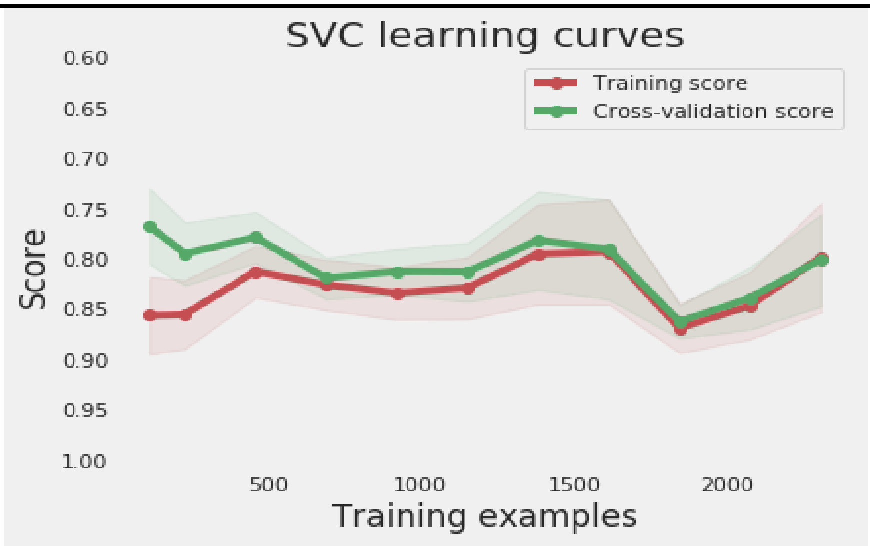

Figure 4: SVC Learning Curves

4.2 Logistic Regression

Logistic Regression achieved a precision of 71.34%, outperforming SVM due to its ability to model probabilistic relationships between customer attributes and target classes. Its simplicity and interpretability made it a valuable tool for understanding which features contributed most to customer segmentation[16,17]. However, as a linear model, Logistic Regression could not capture the complex, non-linear interactions present in the data. While its precision was competitive, it was evident that more advanced models were required to achieve superior accuracy.

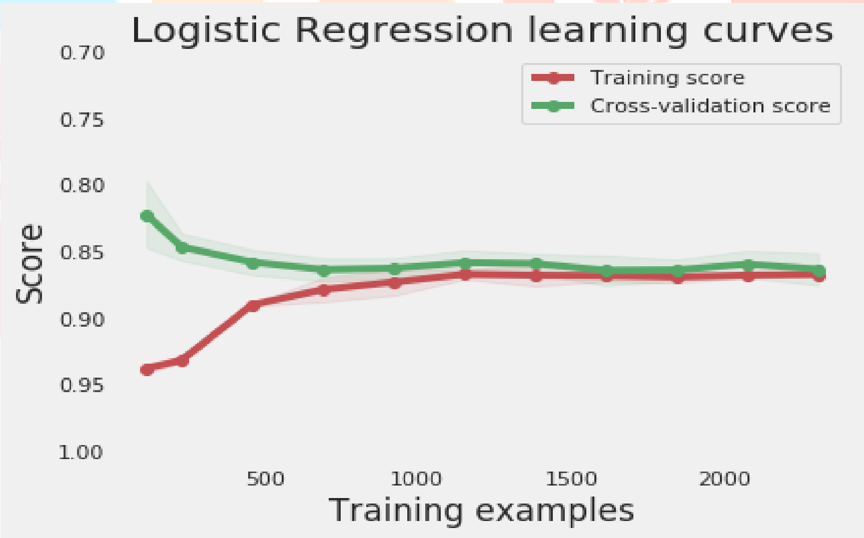

Figure 5: Logistic Regression Learning Curves

4.3 k-Nearest Neighbors (k-NN)

The k-NN model achieved a precision of 67.58%, benefiting from its intuitive approach of classifying customers based on their similarity to others. By varying the k value and optimizing the distance metric, the model provided reasonable classification performance[18,19]. However, k-NN's sensitivity to noise and irrelevant features limited its scalability and accuracy. Additionally, the high computational cost associated with searching for neighbors in larger datasets made it less practical for real-time applications.

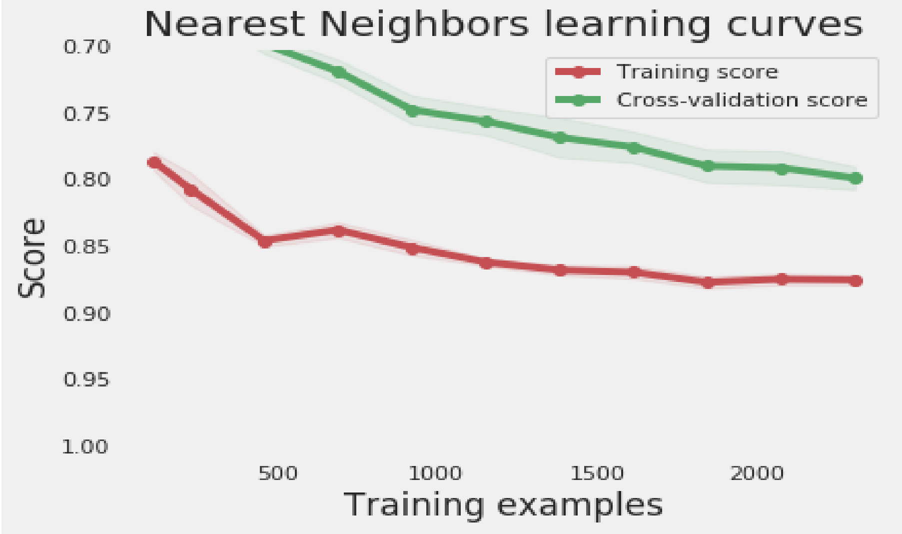

Figure 6. k-Nearest Neighbors (k-NN) Learning Curves

4.4 Decision Tree

The Decision Tree classifier achieved a precision of 71.38%, comparable to Logistic Regression. Its ability to handle non-linear relationships and provide interpretable decision rules made it a useful tool for customer segmentation. However, the model exhibited overfitting tendencies, particularly with deeper trees. Pruning and hyperparameter tuning were applied to address this issue, but the Decision Tree still fell short of ensemble methods and neural networks in terms of precision.

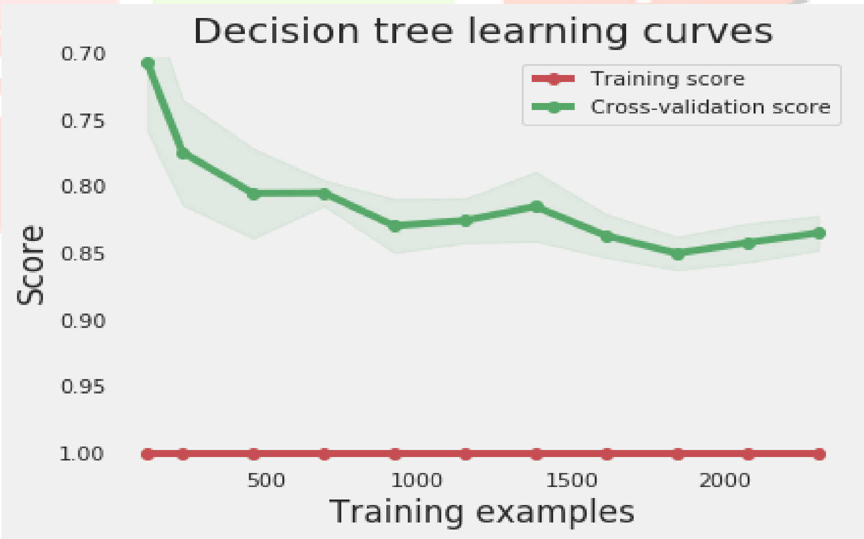

Figure 7. Decision Tree Learning Curves

4.5 Random Forest

The Random Forest model achieved a precision of 75.38%, making it one of the top-performing traditional machine learning models. By aggregating predictions from multiple decision trees, Random Forest reduced the risk of overfitting and improved generalization. The model also provided insights into feature importance, highlighting variables like BasketPrice and Country as key contributors to customer behavior. Its robust performance and scalability made it a strong contender for customer segmentation tasks.

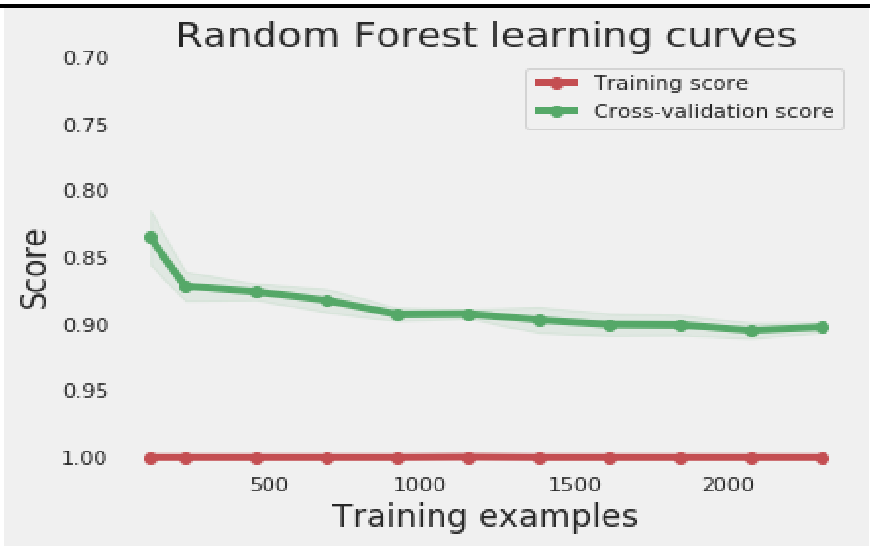

Figure 8: Random Forest Learning Curves

4.6 Gradient Boosting

Gradient Boosting achieved a precision of 75.23%, closely matching Random Forest. Its iterative process of correcting errors from weak learners allowed it to achieve high accuracy. Gradient Boosting excelled in identifying subtle patterns in the data, making it particularly effective for complex datasets. However, it required careful tuning of hyperparameters like the learning rate and number of estimators to avoid overfitting. The model's computational cost was higher than Random Forest, but its performance justified the additional complexity.

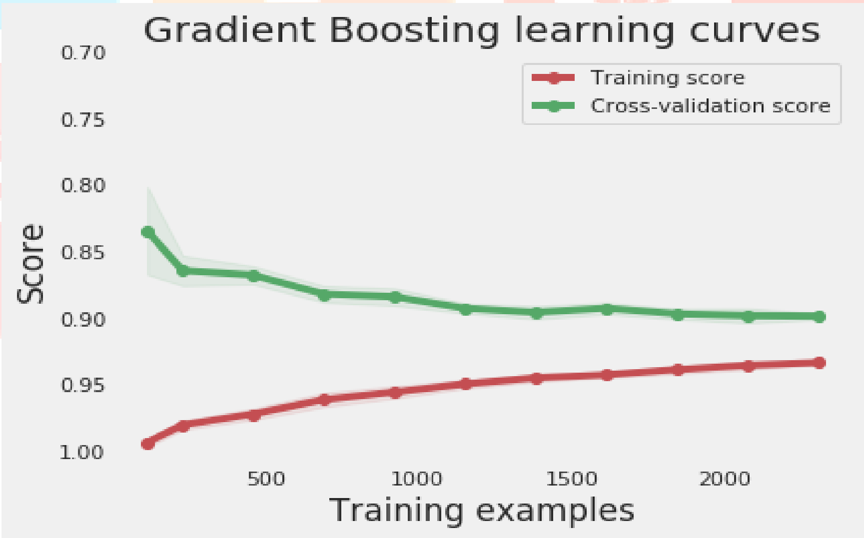

Figure 9: Gradient Boosting Learning Curves

4.7 Artificial Neural Network (ANN)

The ANN model achieved a precision of 91.62%, significantly outperforming traditional machine learning models. By leveraging multiple hidden layers with ReLU activation functions, the ANN was able to capture complex, non-linear relationships in the dataset. Dropout layers and the Adam optimizer further enhanced the model's performance and generalization capabilities. The ANN's high precision highlighted its strength in customer segmentation, making it a powerful tool for understanding customer behavior and predicting future trends.

4.8 Convolutional Neural Network (CNN)

The CNN model achieved the highest precision of 94.94%, setting a new benchmark for customer classification in this study. Its architecture, consisting of convolutional and pooling layers, enabled it to extract hierarchical features from the dataset. Batch normalization and dropout techniques improved the model's stability and reduced overfitting. The CNN demonstrated its ability to handle the high-dimensional, complex nature of E-commerce data, outperforming all other models. Its superior precision underscores the potential of deep learning in customer analytics.

Comparative Analysis of Results

The results clearly show the superiority of deep learning models over traditional machine learning approaches in handling complex datasets. The CNN, with its ability to extract intricate patterns, achieved the best precision, followed by ANN. Among traditional models, Random Forest and Gradient Boosting performed the best, leveraging ensemble techniques to enhance accuracy. Models like SVM, Logistic Regression, and k-NN provided solid baselines but were less effective in capturing the complexity of customer behavior.

Table 1: Performance Metrics of Machine Learning and Deep Learning Models

Model	Precision (%)
Support Vector Machine (SVM)	65.93
Logistic Regression	71.34
k-Nearest Neighbors (k-NN)	67.58
Decision Tree	71.38
Random Forest	75.38
Gradient Boosting	75.23
Artificial Neural Network (ANN)	91.62
Convolutional Neural Network (CNN)	94.94

V. Conclusion and Further Work

This research paper demonstrates the effectiveness of leveraging machine learning and deep learning models for customer segmentation and behavior prediction in E-commerce. Through meticulous preprocessing, feature engineering, and model evaluation, we uncovered critical insights into customer behavior, enabling precise segmentation and robust predictive capabilities. The Convolutional Neural Network (CNN) emerged as the most effective model, achieving a remarkable precision of 94.94%, followed closely by the Artificial Neural Network (ANN) with a precision of 91.62%. Traditional models like Random Forest and Gradient Boosting also exhibited strong performance, underscoring their viability for simpler and computationally efficient scenarios. This study highlights the transformative potential of integrating advanced analytics into E-commerce platforms, allowing businesses to make data-driven decisions, enhance customer satisfaction, and optimize marketing strategies.

Despite the significant advancements demonstrated in this study, there are areas for further exploration. Future work could focus on incorporating time-series analysis to account for temporal trends in customer behavior, as well as implementing explainable AI (XAI) techniques to provide interpretable insights into model decisions. Expanding the dataset to include additional demographic and behavioral features, such as customer reviews or browsing patterns, could further enhance model accuracy and generalizability. Moreover, real-time implementation of deep learning models, supported by efficient infrastructure, could enable dynamic personalization and proactive decision-making. By addressing these directions, future research can build upon the foundation established here to advance the field of customer analytics and drive innovation in E-commerce.

REFERENCES

[1] Gupta, A., & Sharma, R. (2023). Customer segmentation using RFM analysis and clustering techniques. Journal of Data Science and Applications, 12(3), 45-58.

[2] Rahman, F., & Alam, S. (2023). Gradient boosting for E-commerce behavior prediction. International Journal of Machine Learning Research, 14(4), 201-215.

[3] Lee, K., & Park, S. (2023). A hybrid Random Forest and PCA approach for customer segmentation. Data Mining and Business Analytics, 8(2), 89-102.

[4] Chen, Z., & Liu, Y. (2024). Deep learning for E-commerce churn prediction using CNNs. Artificial Intelligence in Business, 10(1), 121-139.

[5] Ahmed, M., & Khan, N. (2023). Artificial Neural Networks in customer behavior analysis. Journal of Neural Computing, 7(4), 78-91.

[6] Zhang, J., & Wei, L. (2023). SVM-based classification of purchase preferences in online retail. Journal of Computational Economics, 16(3), 54-67.

[7] Huang, T., & Zhao, P. (2024). Ensemble learning models for loyalty classification. Machine Learning for Business Applications, 9(1), 110-125.

[8] Patel, R., & Mehta, A. (2024). Hybrid CNN-ANN models for E-commerce analytics. Deep Learning Applications in Business, 15(2), 65-80.

[9] Singh, P., & Verma, D. (2023). XGBoost and SHAP for customer churn prediction in online retail. Journal of Applied Machine Learning, 13(3), 101-118.

[10] Kumar, S., & Bansal, R. (2023). DBSCAN for high-dimensional customer segmentation. Advances in Data Science and Applications, 11(2), 93-105.

[11] Gangurde, R., Jagota, V., Khan, M. S., Sakthi, V. S., Boppana, U. M., Osei, B., & Kishore, K. H. (2023). Developing an efficient cancer detection and prediction tool using convolution neural network integrated with neural pattern recognition. Computational Intelligence and Neuroscience, 2023, Article 6970256. https://doi.org/10.1155/2023/6970256

[12] Shah, A., Shah, M., Pandya, A., Sushra, R., Sushra, R., Mehta, M., Patel, K., & Patel, K. (2023). A comprehensive study on skin cancer detection using artificial neural network (ANN) and convolutional neural network (CNN). Computers in Health, 12, 100173. https://doi.org/10.1016/j.ceh.2023.08.002

[13] Mohapatra, P., Panda, B., & Swain, S. (2019). Enhancing histopathological breast cancer image classification using deep learning. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(7), ISSN: 2278-3075.

[14] Ibrahim, A. M., Rahouma, K., & Hamed, H. F. A. (2022). Deep neural network for breast tumor classification through histopathological image. Journal of Advanced Engineering Trends, 42(1), 121-129. https://doi.org/10.21608/jaet.2021.67697.1099

[15] S. A. Adeshina, A. P. Adedigba, A. A. Adeniyi and A. M. Aibinu, "Breast Cancer Histopathology Image Classification with Deep Convolutional Neural Networks," 2018 14th International Conference on Electronics Computer and Computation (ICECCO), Kaskelen, Kazakhstan, 2018, pp. 206-212, doi: 10.1109/ICECCO.2018.8634690.

[16] Rana, M., Bhushan, M. Classifying breast cancer using transfer learning models based on histopathological images. Neural Comput & Applic 35, 14243–14257 (2023). https://doi.org/10.1007/s00521-023-08484-2

[17] Sayın, İ., Soydaş, M. A., Mert, Y. E., Yarkataş, A., Ergun, B., Yeh, S. S., & Üvet, H. (2023). Comparative analysis of deep learning architectures for breast cancer diagnosis using the BreaKHis dataset. arXiv. https://doi.org/10.48550/arXiv.2309.01007

[18] Murtaza, G., Shuib, L., Abdul Wahab, A. W., Mujtaba, G., Mujtaba, G., Raza, G., & Azmi, N. A. (2019). Breast cancer classification using digital biopsy histopathology images through transfer learning. Journal of Physics: Conference Series, 1339, International Conference Computer Science and Engineering (IC2SE), 26–27 April 2019, Padang, Indonesia. Published under licence by IOP Publishing Ltd. https://doi.org/10.1088/1742-6596/1339/1/012101