Available online at
https://mijuitm.com.my

e-ISSN: 2735-0703

Mathematical Sciences and Informatics Journal 6(2) 2025, 203 – 217.

Mathematical Sciences and Informatics Journal

Optimizing E-commerce Inventory Management Using a Machine-Learning Approach

Zhao Ruonan[1*], Doris Hooi-Ten Wong[1]

[1]*Faculty of Artificial Intelligence, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia*

ARTICLE INFO

Article history:
Received 22 August 2025
Revised 21 September 2025
Accepted 30 September 2025
Online first
Published 31 October 2025

Keywords:
E-commerce
Inventory Management
Predictive Analytics
Machine Learning
Demand Forecasting
Customer Segmentation

DOI:
10.24191/mij.v6i2.9612

ABSTRACT

E-commerce inventory management faces persistent challenges such as overstock and stockouts caused by unpredictable demand. Traditional inventory systems often fail to process large-scale transactional data efficiently, limiting accurate forecasting and decision-making. To address this issue, this study proposes an integrated machine-learning framework that combines predictive analytics and customer segmentation to improve forecasting precision and inventory control. Three machine learning models LSTM, XGBoost, and Random Forest were compared for demand forecasting. Among them, LSTM achieved the lowest RMSE (0.799), indicating superior predictive performance for time-dependent data. In addition, clustering algorithms, including DBSCAN and K-means, were applied to segment customers based on purchasing behaviour, with DBSCAN achieving a Silhouette Score of 0.9708, suggesting well-separated clusters. The results were visualised to generate actionable insights, enabling data-driven decisions. The findings provide an added approach for e-commerce businesses by linking sales forecasting and customer clustering to more efficient inventory allocation.

1. INTRODUCTION

E-commerce has fundamentally transformed business operations by enabling firms to reach global markets and operate beyond traditional physical boundaries. Projections by Barthel (2023) indicate that e-commerce is anticipated to account for 41% of all retail sales worldwide by 2027. This digital transformation has had a significant impact on various aspects of operations management, including inventory management. The industry faces numerous inventory challenges, including multi-channel inventory coordination (Wang et al., 2025), shortened product lifecycles, frequent new product updates, and inventory cost pressures. In this context, traditional inventory management methods, such as Economic Order Quantity and periodic review systems, are increasingly inadequate in today's dynamic digital marketplace. Furthermore, current forecasting approaches often ignore customer behavioural information, resulting in reactive rather than

[1*] Corresponding author. *E-mail address*: zhaoruonan@graduate.utm.my
https://doi.org/10.24191/mij.v6i2.9612

proactive stock decisions (Xia et al., 2020). Consequently, there remains a gap in developing approaches that combine demand prediction with customer behavioural insights.

The primary aim of this research is to design and evaluate a machine-learning–based framework that integrates predictive analytics and customer segmentation to enhance e-commerce inventory management. The study pursues three specific objectives: (i) to forecast overall product demand using LSTM, XGBoost, Random Forest models and compare their performances; (ii) to segment customers using clustering algorithms such as DBSCAN and K-means; and (iii) to visualise integrated results through Power BI dashboards. This study contributes by demonstrating an integrated analytical framework that combines machine learning prediction, customer segmentation, and interactive visualisation, providing actionable insights for optimizing inventory decisions and improving operational efficiency in e-commerce management.

2. LITERATURE REVIEW

According to the research aim, the part of literature review concludes three main contents: existing methods including traditional and machine learning methods that applied to inventory management, current application of performance evaluation, predictive and customer analytics in inventory management, reliable algorithms used in time series prediction and clustering.

2.1 Limitation of Traditional Inventory Management

Traditional inventory management methods were primarily rule-based and relied on deterministic models. Its primary objective was to maintain optimal stock levels that satisfy anticipated customer demand while minimizing associated costs such as storage, ordering, and stockouts. At its essence, traditional inventory management involved systematic planning and control of goods as they moved from procurement through storage and ultimately to the point of sale. This approach emphasized ensuring that the right products, in the right quantities, are available at the right time and location to support consistent service delivery and operational efficiency (Khirwar et al., 2023). A typical example of how this traditional method works in stock management is the Economic Order Quantity model, which aims to identify the optimal order quantity that minimizes the total cost of inventory by balancing ordering and holding costs. The forecasts required for Economic Order Quantity calculations are typically based on historical demand data and expected changes in usage rates. Similarly, reorder point systems were used in determining the precise time to replenish stock, calculated based on key parameters such as lead time, average demand, and the desired service level. These models provide a structured and interpretable framework for decision-making, but they are often built on the assumption of stable demand and predictable supply conditions, an assumption that may not hold true in today's volatile and fast-paced commercial environments (Trajanova et al., 2023).

2.2 Metrics for Evaluating Predictive and Clustering Results

Demand forecasting relies on time series prediction, where evaluating model accuracy is essential. Common metrics include Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE). MAPE measures the average absolute percentage difference between predicted and actual values, whereas MSE and RMSE quantify the magnitude of prediction error (Hussain et al., 2025). RMSE, being the square root of MSE, retains the same unit as the predicted variable, making it easier to interpret in practical terms. Accordingly, RMSE is appropriate for this study because it retains the same unit as the predicted variable, which facilitates business interpretation in inventory management.

For clustering evaluation, the Silhouette Score and the Davies-Bouldin Index (DBI) are widely used. The Silhouette Score evaluates clustering quality by measuring cohesion and separation (Ros et al., 2023). It is also computationally efficient, with near-linear time complexity relative to the number of samples and

https://doi.org/10.24191/mij.v6i2.9612

clusters (Ji et al., 2023). However, DBI is highly sensitive to the choice of distance metric, which can significantly affect the index value and thus the evaluation outcome (Rustam et al., 2020). Therefore, this study adopts the Silhouette Score to evaluate DBSCAN results, as its flexibility better captures the non-linear and overlapping structures typical in e-commerce clustering tasks.

2.3 Machine Learning Algorithms for Predictive Analytics

The selection of machine learning algorithms was guided by three key criteria: predictive accuracy in capturing demand fluctuations, computational efficiency for practical deployment, and adaptability to temporal characteristics of sales data.

For capturing long-term dependencies in sequential data, LSTM is well-suited to time series prediction. For instance, Palkar et al. (2020) evaluated multiple algorithms on Iowa liquor sales data and found that LSTM achieved 90% prediction accuracy, outperforming multiple linear regression (70%) and SVM (50–60%). Similarly, Jain et al. (2020) compared SARIMA and LSTM for forecasting product demand in e-commerce. While SARIMA was more effective for long-term seasonal trends, LSTM yielded better short-term predictions due to its memory mechanism and gating structure. Falatouri et al. (2022) likewise found that LSTM outperformed SARIMA for products with stable demand (e.g., potatoes), whereas SARIMA was more accurate for highly seasonal items (e.g., cucumbers). Finally, Shah et al. (2023) applied ARIMA and LSTM to short-term demand for diabetes medications: LSTM performed well in training, but ARIMA produced more consistent test-set results (lower RMSE).

XGBoost has shown strong predictive power. In Pramodhini et al. (2023), XGBoost was used to develop an inventory forecasting model integrated into a customer-facing web application; despite its computational complexity, the model achieved $R^2 = 0.8639$ after training on 80% of local demand data. Random Forest remains a practical choice in e-commerce analytics, for example, Zaghloul et al. (2024) reported 92% accuracy for customer-satisfaction prediction using delivery time and order value, outperforming several deep learning baselines.

2.4 Machine Learning Algorithms for Customer Segmentation

K-means and K-medoids are among the most used algorithms in RFM-based segmentation due to their simplicity and interpretability. Hilmy et al. (2023) applied both algorithms to cluster customers and evaluated the results using Silhouette Coefficient, Davies-Bouldin Index, and Calinski-Harabasz Index. The study concluded that K-means outperformed K-medoids across all metrics, demonstrating better cohesion and separation of customer groups. Similarly, Mufarroha et al. (2022) conducted a comparative analysis using online retail data and found that K-medoids provided more detailed segmentation but required higher computational effort, whereas K-means was faster.

In more specialized domains, Chen et al. (2025) explored customer behaviour in the luxury NFT market by applying K-means to OpenSea transaction data related to Gucci's SUPERGUCCI collection. The study identified three customer segments: Speculators, Casual Collectors, and Cryptocurrency Natives with a Silhouette Score of 0.58 for "Offers Received" and 0.42 for "Offers Made", suggesting generally well-separated clusters. This application illustrates how K-means can also provide business insights in emerging digital economies.

Further empirical evidence is provided by Sahinbas et al. (2022), who analysed customer behaviour from a private e-commerce retailer using several algorithms: OPTICS, BIRCH, Agglomerative Clustering, K-means, and DBSCAN. While most algorithms produced comparable clusters, DBSCAN achieved the highest Silhouette Score 0.206, highlighting its strength in handling noise and discovering non-spherical clusters. Wani et al. (2023) implemented Spectral Clustering by constructing a similarity graph and performing eigenvalue decomposition to uncover customer groups. Its performance was limited in this

https://doi.org/10.24191/mij.v6i2.9612

study, yielding a low Silhouette Score -0.004. Conversely, DBSCAN performed better, especially when applied to datasets with irregular density.

3. METHODOLOGY

To ensure a systematic and structured approach to reach the study objective, this study adopted a multi-phase methodology encompassing data collection, preprocessing, exploratory data analysis, model development, performance evaluation, and visualisation.

3.1 Data Collection

The dataset was taken from Kaggle, an open website providing datasets. The dataset consists of 1,000 transaction records, with each row representing a unique sales transaction. Table 1 presents the data description.

Table 1. Description of dataset

Column Name	Description
Customer ID	An anonymized identifier for each customer
Age	Age of customers
Gender	Gender of the customers
Quantity	Number of items bought
Product Category	Type of product purchased
Total Amount	Total expenditure for the transaction
Transaction ID	Number of the transactions
Date	Dates of transactions
Price per Unit	Unit price of the product

3.2 Data Preprocessing and Exploratory Data Analysis

All data processing and Exploratory Data Analysis (EDA) in this study were conducted using Python within a Google Colab environment. Prior to model development, the dataset underwent several preprocessing steps to ensure data quality, consistency, and suitability for machine learning analysis. These steps included converting the “Date” column to datetime format, data cleaning, one-hot encoding of categorical fields like “Gender” and “Product Category”, scaling and normalize numerical variables.

EDA was conducted to better understand the structure, patterns, and relationships present in the dataset prior to model development. Descriptive statistics is first computed to summarize the central tendencies and variability of key numerical features such as “Quantity”, “Price per Unit”, and “Total Amount”. Time series decomposition of aggregated sales data by date was used to examine seasonality and demand fluctuations, which are critical for forecasting.

3.3 Predictive and Clustering Model Building

For forecasting future product demand, three models were trained and compared: LSTM, Random Forest, and XGBoost. The selection of models was based on their suitability for time-series forecasting and customer segmentation. LSTM was chosen for its strength in capturing nonlinear temporal dependencies, XGBoost for its high predictive accuracy and robustness against overfitting, and Random Forest for its stability, interpretability, and ability to handle high-dimensional data. ARIMA are also widely used for time-series forecasting, they are more effective for linear and stationary data. However, e-commerce demand patterns are often nonlinear and influenced by multiple features. Therefore, ARIMA is less suitable for this study. Variables “Date” and “Quantity” were aggregated daily to construct a time series, which was then used to train predictive models. The best-performing model is then used to forecast the total sales

https://doi.org/10.24191/mij.v6i2.9612

quantity for the following 30 days. For all predictive models, the dataset was divided approximately 80% of the earliest records were used for model training and the remaining 20% for testing.

For segmenting customers, two models were implemented: K-means and DBSCAN. K-means was used for its efficiency and ease of interpretation, while DBSCAN was applied for its ability to identify clusters of arbitrary shapes and detect noise without requiring a predefined number of clusters. The features used for clustering include “Price per Unit” and “Product Category”. The clustering algorithm with the higher Silhouette Score was selected for the final segmentation analysis.

3.4 Performance Evaluation

For predictive modelling, RMSE was used to evaluate accuracy. As RMSE penalizes larger errors and is expressed in the same unit as “Quantity”, it offers an interpretable measure of forecasting performance. The model with the lowest RMSE on the test set was selected to predict the next 30 days of sales. For clustering, the Silhouette Score measured how well-defined each cluster was, with values near 1 indicating strong separation and values near 0 or negative suggesting poor clustering. K-means with cluster numbers from 3 to 9 was compared against DBSCAN, and the model with the highest Silhouette Score was selected.

3.5 Data Visualization

To enhance the interpretability of the customer segmentation results, the final clustering results were visualised using Microsoft Power BI, a widely accessible business intelligence tool. The visualisation primarily focused on uncovering which product categories were most popular among different customer segments.

4. RESULTS AND DISCUSSION

This section presents and discusses the results obtained from the implementation of predictive and clustering models, executed in Python using the Google Colab environment, to evaluate the performance of different machine learning techniques and interpret their outputs in the context of inventory optimisation.

4.1 Exploratory Data Analysis

As shown in Table 2, the average customer age is 41.4 years, with a wide range from 18 to 64, indicating a diverse customer base. The mean quantity per transaction is 2.51, suggesting most purchases involve 1 to 4 items. “Price per Unit” and “Total Amount” show wide variation, reflecting heterogeneous pricing strategies and customer spending levels across product categories. Such dispersion highlights the need for segmentation to better understand different purchasing behaviours.

Table 2. Summary statistics of numerical features

index	Age	Quantity	Price per Unit	Total Amount
count	1000.0	1000.0	1000.0	1000.0
mean	41.392	2.514	179.89	456.0
std	13.681429659	1.1327343409	189.68135627	559.997631
min	18.0	1.0	25.0	25.0
25%	29.0	1.0	30.0	60.0
50%	42.0	3.0	50.0	135.0
75%	53.0	4.0	300.0	900.0
max	64.0	4.0	500.0	2000.

As shown in Fig. 1, daily sales quantity in 2023 fluctuated significantly, with notable spikes, especially between May and July. These peaks may correspond to promotional campaigns or external seasonal effects,

https://doi.org/10.24191/mij.v6i2.9612

indicating short-term influences on purchasing behaviour. The absence of a clear long-term seasonal pattern suggests that demand is irregular, reinforcing the need for robust predictive models capable of capturing non-linear and event-driven variations.

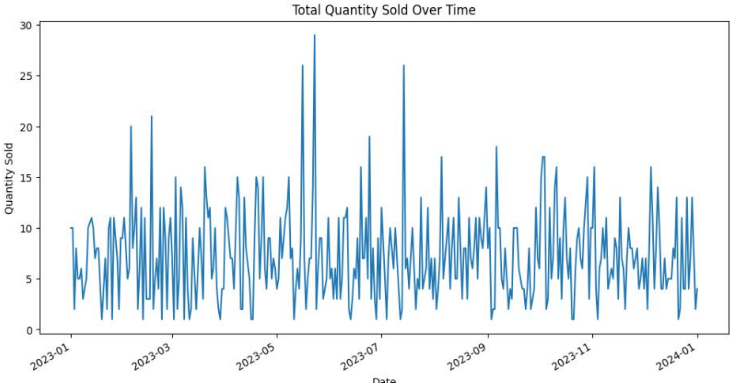

Fig. 1. Total quantity sold daily throughout the year 2023

4.2 Predictive Results of LSTM, XGBoost and Random Forest

Fig.2 shows the comparison between the actual and predicted quantity sold using the LSTM model. The blue line represents the actual daily sales, while the orange line represents the model's predictions. While the predicted values generally follow the trend of the actual data, the LSTM model tends to smooth out extreme fluctuations. This is a common outcome in time series models trained on highly volatile data, indicating that the model captures overall patterns well but underestimates sharp peaks in sales. The model achieved RMSE of 0.7993, indicating relatively good forecasting accuracy given the inherent noise and variability in the sales data. The main parameters influencing the LSTM model include the window size, number of epochs, and batch size. Several configurations were tested by adjusting these values, but the performance differences were minimal.

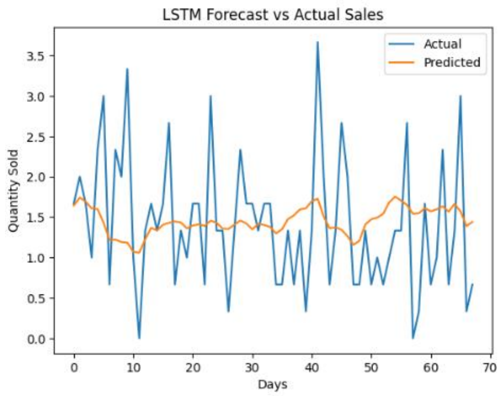

Fig. 2. Actual vs predicted daily quantity sold using LSTM

Fig. 3 shows the comparison between actual and predicted daily sales quantities using the XGBoost regression model. The blue line represents the actual sales data, while the orange line shows the model's predicted values across the test set. The model captures the general trend and fluctuations reasonably well, though some of the sharper spikes and dips are smoothed out or slightly shifted. The RMSE of 4.455

https://doi.org/10.24191/mij.v6i2.9612

indicates a moderate level of prediction error, which is expected given the high volatility in the actual sales data. The XGBoost regressor was applied using its default configuration, which includes 100 boosting rounds that $n\_estimators = 100$, a learning rate of 0.1 that $learning\_rate = 0.1$, and a maximum tree depth of 3, $max\_depth = 3$. These default settings are widely adopted. The outcome provides stable and reliable performance.

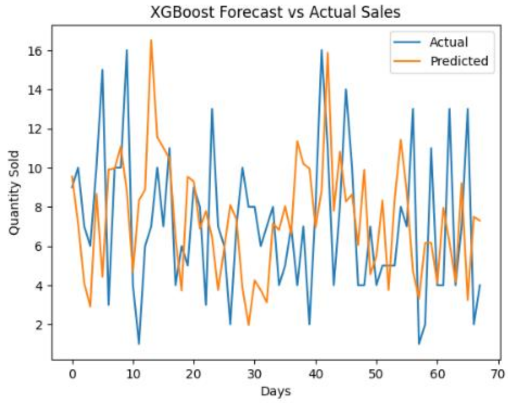

Fig. 3. Actual vs predicted daily quantity sold using XGBoost

The model of Random Forest achieved a RMSE of 3.995, slightly outperforming XGBoost with RMSE value equal to 4.455. As shown in Fig. 4, the predicted values, orange line closely follow the overall trend of actual sales, blue line, though some of the more extreme fluctuations are smoothed out. The results suggest that Random Forest captures mid-level patterns and general demand movement effectively, though like other models, it is less sensitive to sharp, short-term sales spikes. Furthermore, the Random Forest model was implemented with 100 decision trees that $n\_estimators = 100$ and a fixed random seed $random\_state = 42$ to ensure reproducibility. Increasing the number of trees beyond this point did not produce noticeable performance improvements, indicating that the model had reached a stable level of generalization.

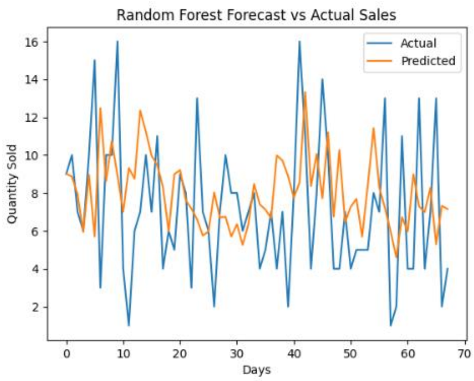

Fig. 4. Actual vs predicted daily quantity sold using random forest

Table 3 presents the predictive results evaluated by RMSE, which illustrates that LSTM achieved the lowest RMSE value in a straightforward way. Compared with Random Forest, LSTM reduced average forecast error by approximately 80%, indicating that its predictions are significantly closer to actual sales

https://doi.org/10.24191/mij.v6i2.9612

quantities. Furthermore, unlike XGBoost and Random Forest, LSTM learns temporal dependencies automatically from the sequence of past sales data. This ability enables it to anticipate short-term demand shifts without manual feature engineering.

Table 3. Performance of predictive algorithms

Algorithm	RMSE
LSTM	0.799
XGBoost	4.455
Random Forest	3.995

In this study, LSTM model was employed to forecast the total quantity of sales over the next 30 days. Before feeding the data into the model, the target variable Quantity was normalized using the MinMaxScaler technique to transform all values into a [0,1] range. This preprocessing step is essential for improving model convergence and stability, as LSTM networks are sensitive to the scale of input data. After training, the predicted outputs were inverse transformed to convert them back to their original scale. This allows the final forecast results to be interpreted in real-world quantities.

To improve the model's temporal learning capacity, the "window size" was set to 21, meaning the model uses the sales data of the past 21 days to predict the next day's quantity. This long input sequence enables the model to capture short-term trends and seasonal effects. In addition, the number of training "epochs" was increased to 200, ensuring the model has sufficient opportunity to learn patterns from the data without underfitting. As illustrated in Fig. 5, the blue line represents the historical sales quantity for the past 60 days, and the orange line shows the predicted quantity for the upcoming 30 days. While the historical data demonstrates significant daily fluctuations, the LSTM forecast appears relatively smooth and stable, the future sales was predicted around 1.5 in average. This phenomenon can be attributed to the limited feature set, small dataset size, and the model's inherent tendency to minimize error by averaging predictions when uncertainty is high. Despite the smoothness, the forecast provides a useful reference for inventory planning, as it suggests a steady demand trend with minimal risk of extreme fluctuations and avoiding both overstock and stockouts.

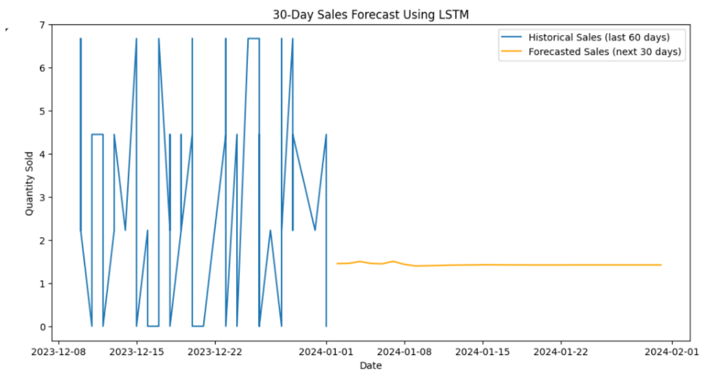

Fig. 5. Sales prediction for the future 30 days

4.3 Clustering Results of DBSCAN and K-means

DBSCAN was applied to two standardized features: "Price per Unit" and "Product Category". The model was configured with an eps value of 0.5 and *min_samples* set to 5. These hyperparameters were selected based on experimentation to balance sensitivity and specificity in cluster formation. The resulting DBSCAN clustering output, visualized in Fig. 6 shows that the algorithm successfully identified nine distinct clusters along with a few noise points. Each coloured group corresponds to a cluster of products

https://doi.org/10.24191/mij.v6i2.9612

with similar pricing patterns and categorical similarities. The performance of the DBSCAN model was evaluated using the Silhouette Score, which yielded a remarkably high value of 0.9708. This indicates excellent clustering quality in terms of both cohesion and separation.

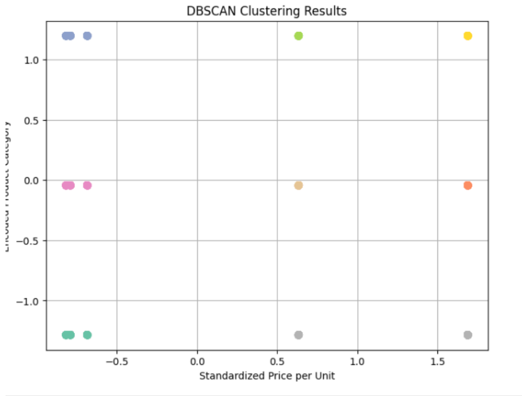

Fig. 6. DBSCAN clustering result

The K-means algorithm was also applied to segment the dataset based on two features: “Price per Unit” and “Product Category”. K-means is a centroid-based algorithm that requires the number of clusters k to be predefined. To determine the optimal number of clusters, two evaluation techniques were used: the Elbow method and the Silhouette Score. As shown in Fig. 7, the Elbow method indicated that the Elbow curve shows a gradual and smooth decline without a clear "elbow point”. $K = 4$ is merely a relatively reasonable estimation.

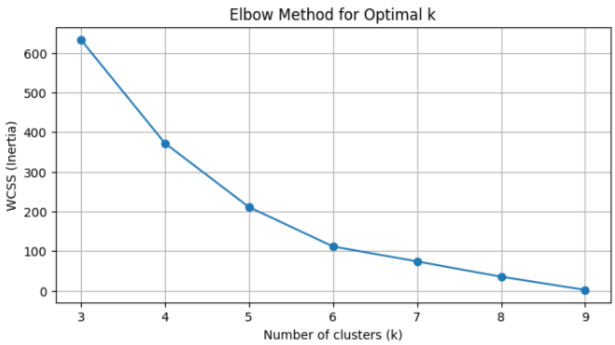

Fig. 7. Elbow method of K-means

Meanwhile, Fig. 8 demonstrates that the Silhouette Score continued to increase up to k = 9, where it peaked at 0.9708. This the same Silhouette Score as DBSCAN, then selecting nine clusters for K-means allows for a direct, fair comparison between the two algorithms. Customers are also divided into nine distinct clusters clearly.

https://doi.org/10.24191/mij.v6i2.9612

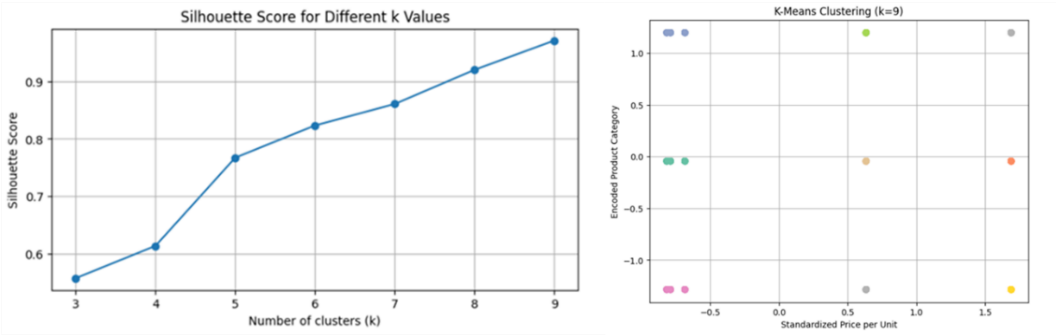

Fig. 8. Clustering result of K-means

The clustering analysis using both DBSCAN and K-means yielded a very high Silhouette Score of 0.9708, which indicates near-perfect separation between clusters. However, such an unusually high value warrants critical interpretation rather than being viewed as purely positive. In real-world customer or product segmentation tasks, Silhouette Score above 0.8 are rare, as behavioural and pricing data typically contain overlaps and noise. The result likely stems from the limited feature set, only "Price per Unit" and "Product Category." Since these two variables already separate products clearly, the algorithms produced clusters that merely reflect existing categories rather than uncovering new behavioural patterns. Thus, the high score reflects structural simplicity, not true clustering quality. Similar e-commerce studies using richer variables typically report Silhouette Score between 0.3 and 0.8. Future research should include additional behavioural or demographic features to obtain more realistic and meaningful segmentation results. Despite the high Silhouette Score, DBSCAN was retained as the primary clustering method because it automatically determines the number of clusters and detects outliers effectively.

4.4 Data Visualization Analysis

For further interpreting the results of clustering and support data-driven decision-making, a series of visualizations were created using Power BI where product categories and pricing segments were visualized across clusters. This allowed for a clear identification of which clusters contained high value products, frequent buyers, or price-sensitive categories.

As shown in Fig. 9, which displays the total purchase quantity per cluster. Notably, Cluster 3 recorded the highest purchase quantity with 544 units, followed by Clusters 2 and 0 with 497 and 460 units respectively. These clusters likely represent the most active or high demand product groups and should be prioritized in inventory restocking and promotional planning.

https://doi.org/10.24191/mij.v6i2.9612

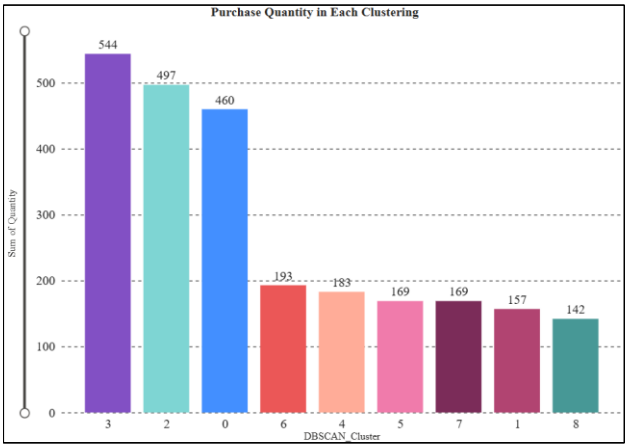

Fig. 9. Purchase quantity in each cluster

Fig. 10 shows that Cluster 3 stands out with the largest portion of total purchases: 21.64%, and clothing is the most popular product, suggesting this segment is heavily interested in fashion or apparel. Clusters 2 is strongly associated with electronics, indicating a tech-focused customer group. Cluster 0 shows a clear preference for beauty products.

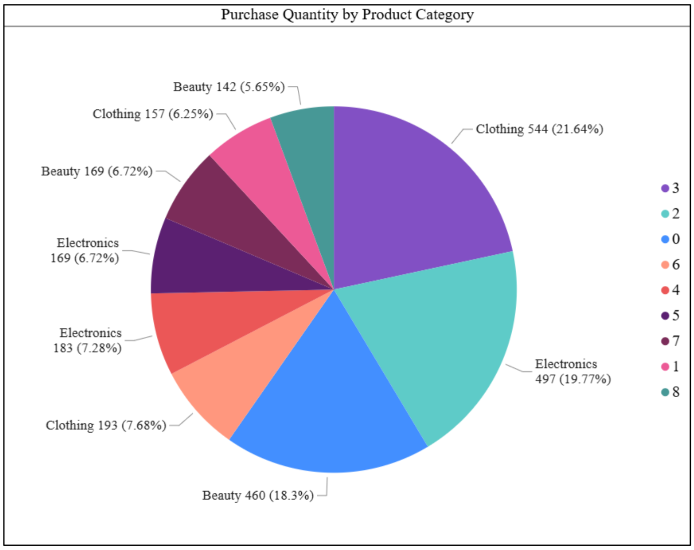

Fig. 10. Pie chart of product categories dominates each cluster

In addition to examining purchase quantity, the total purchase amount per cluster was visualized to better understand the value contribution of each group. As shown in Fig. 11, Clusters 5 and 7 had the highest total spending, each exceeding 85,000 dollars, suggesting that these groups include high value or premium customers. Interestingly, while Cluster 3 had the highest purchase quantity earlier, its total amount spent ranks much lower, indicating that this group tends to purchase lower priced items more frequently.

https://doi.org/10.24191/mij.v6i2.9612

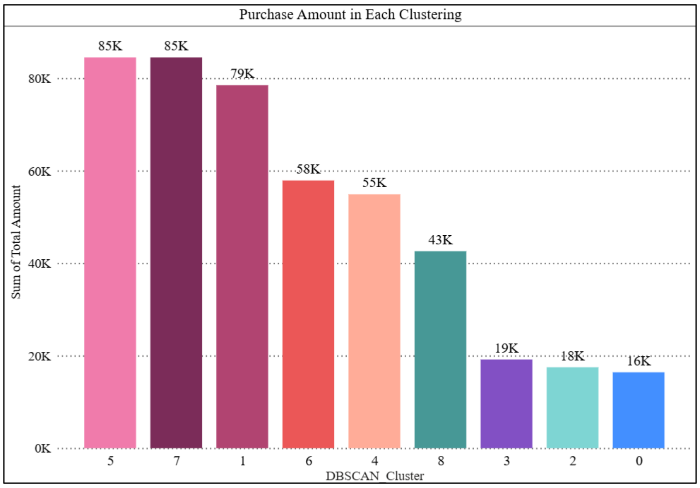

Fig. 11. Purchase amount in each cluster

Fig. 12 shows how product categories and price levels vary across different clusters. Clusters 0, 2, and 3 mostly purchased lower priced items: 25-50, indicating price-sensitive segments. These clusters primarily bought beauty products, electronics, and clothing, respectively. Clusters 4, 5, 6, 7, 1 and 8 were dominated by high priced purchases: 300-500, showing a clear preference for premium items, especially in clothing and electronics.

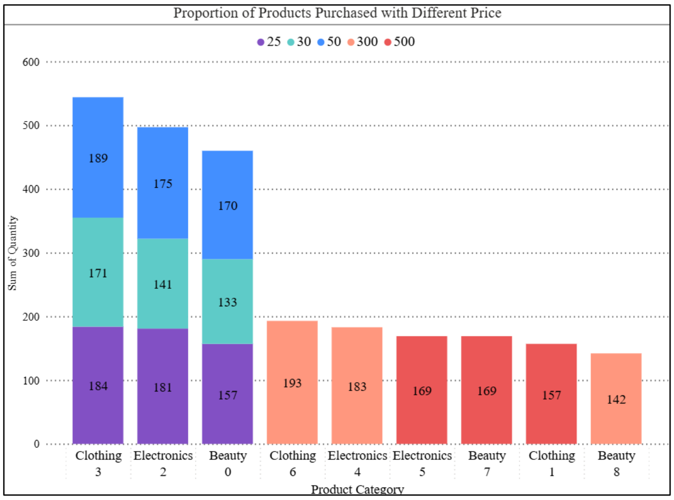

Fig. 12. Proportion of products purchased with different price

In summary, Clusters 5 and 7 exhibited high spending and should be prioritized in inventory planning to maximize value. By focusing on these premium groups, businesses can better align stock with purchasing power and enhance customer loyalty. Insights into product categories further support a segmented stocking strategy. Clothing and electronics, frequently purchased in large volumes, should be stocked more generously, while beauty products with lower demand can be managed through leaner inventories and frequent restocking. This ensures inventory aligns with actual consumption patterns. Additionally, price sensitivity analysis shows that lower-priced items: $25–$50 had higher sales, highlighting their fast-moving

https://doi.org/10.24191/mij.v6i2.9612

nature. In contrast, high-priced products \$300–\$500 were less frequently purchased, suggesting a selective stocking approach. Aligning inventory with both price range and demand helps improve cost efficiency and meet diverse customer needs.

5. CONCLUSION

This section summarizes the main findings of the study, focusing on how clustering and predictive models contribute to inventory optimization.

5.1 Limitation

The predictive model was trained on only one year of historical sales data, which significantly restricted its ability to capture long-term seasonal patterns or trends. As a result, the future sales forecast could only extend 30 days ahead. Furthermore, the predicted values for the latter part of this period appeared unusually flat, with no significant variation in the final 20 days. This suggests that the model may have been unable to detect deeper temporal patterns due to data insufficiency, or that the input features lacked the diversity required to generate realistic, dynamic forecasts. Additionally, the high Silhouette Score (0.97) in the clustering results may indicate overfitting caused by the limited number of features, as only Price and Category were used in the analysis. These limitations indicate the need for larger, multi-year datasets and more feature-rich input to support more reliable and actionable forecasting.

5.2 Future Work

Although the LSTM model exhibited smoothing and the clustering analysis showed a high Silhouette Score, these results are primarily attributed to data limitations rather than methodological flaws. The dataset was relatively small and feature-constrained, which restricted the models' ability to learn complex temporal and behavioural patterns. With access to richer and more diverse data, spanning multiple time periods, customer attributes, and product features, the combined forecasting and clustering approach could be further refined to capture complex demand behaviours and customer heterogeneity.

6. ACKNOWLEDGEMENT/FUNDING

We would like to express my sincere gratitude to my supervisor, Dr. Doris Wong Hooi Ten, for her valuable guidance and continuous support throughout this research. My appreciation also goes to the Faculty of Artificial Intelligence (FAI), Universiti Teknologi Malaysia, for providing the necessary resources and academic environment to complete this project.

7. CONFLICT OF INTEREST STATEMENT

The authors agree that this research was conducted in the absence of any self-benefits, commercial or financial conflicts and declare the absence of conflicting interests with the funders.

8. AUTHORS' CONTRIBUTIONS

Authors contributed equally to conceptualisation, methodology, software and data curation, analysis and interpretation, visualisation, and writing.

https://doi.org/10.24191/mij.v6i2.9612

REFERENCES

Barthel, M., Faraldi, A., Robnett, S., Darpö, O., Lellouche Tordjman, K., Derow, R., & Ernst, C. (2023). *Winning formulas for e-commerce growth*. Boston Consulting Group. https://www.bcg.com/publications/2023/winning-formulas-for-ecommerce-growthh

Chen, Q., Choi, B. J., & Lee, S. J. (2025). Tailoring customer segmentation strategies for luxury brands in the NFT market – The case of SUPERGUCCI. *Journal of Retailing and Consumer Services*, *82*. https://doi.org/10.1016/j.jretconser.2024.104121

Falatouri, T., Darbanian, F., Brandtner, P., & Udokwu, C. (2022). Predictive Analytics for Demand Forecasting - A Comparison of SARIMA and LSTM in Retail SCM. *Procedia Computer Science*, *200*, 993–1003. https://doi.org/10.1016/j.procs.2022.01.298

Hilmy, F. M., Nurhaliza, R. A., Huzyan Octava, M. Q., & Alfian, G. (2023). Web-based E-Commerce Customer Segmentation System Using RFM and K-means Model. *2023 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, 3ICT 2023*, 83–87. https://doi.org/10.1109/3ICT60104.2023.10391650

Hussain, F., Hasanuzzaman, M., & Rahim, N. A. (2025). Multivariate machine learning algorithms for energy demand forecasting and load behavior analysis. *Energy Conversion and Management: X*, *26*. https://doi.org/10.1016/j.ecmx.2025.100903

Jain, A., Shambhavi, B. R., Karthikeyan, V., Sindhu, K., Sahana, B., & Balaji, S. (2020). *Demand forecasting for e-commerce platforms*. In *Proceedings of the 2020 International Conference on Data Science and Network Security (ICDSNS 2020)*. IEEE. https://doi.org/10.1109/ICDSNS58469.2023.10245677

Ji, Y., Dutta, P., & Davuluri, R. (2023). Deep multi-omics integration by learning correlation-maximizing representation identifies prognostically stratified cancer subtypes. *Bioinformatics Advances*, *3*(1). https://doi.org/10.1093/bioadv/vbad075

Khirwar, M., Gurumoorthy, K. S., Jain, A. A., & Manchenahally, S. (2023). *Cooperative Multi-Agent Reinforcement Learning for Inventory Management*.

Mufarroha, F. A., Suzanti, I. O., Satoto, B. D., Syarief, M., Husni, & Yunita, I. (2022). K-means and K-Medoids Clustering Methods for Customer Segmentation in Online Retail Datasets. In *Proceeding – IEEE. 8th Information Technology International Seminar, ITIS 2022*, 223–228. https://doi.org/10.1109/ITIS57155.2022.10010135

Palkar, A., Deshpande, M., Kalekar, S., & Jaswal, S. (2020). *Demand forecasting in retail industry for liquor consumption using LSTM.* In *Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020)* (pp. 521–525). IEEE. https://doi.org/10.1109/ICESC48915.2020.9155712

Pramodhini, R., Kumar, S., Bhardwaj, S., Agrahari, N., Pandey, S., & Harakannanavar, S. S. (2023). E-Commerce Inventory Management System Using Machine Learning Approach. *2023 International Conference on Data Science and Network Security, ICDSNS 2023*. https://doi.org/10.1109/ICDSNS58469.2023.10245500

https://doi.org/10.24191/mij.v6i2.9612

Ros, F., Riad, R., & Guillaume, S. (2023). PDBI: A partitioning Davies-Bouldin index for clustering evaluation. *Neurocomputing*, *528*, 178–199. https://doi.org/10.1016/j.neucom.2023.01.043

Rustam, Usman, K., Kamaruddin, M., Chamidah, D., Nopendri, Saleh, K., Eliskar, Y., & Marzuki, I. (2020). *Modified Possibilistic Fuzzy C-Means Algorithm for Clustering Incomplete Data Sets*. http://arxiv.org/abs/2007.04908

Sahinbas, K., & Catak, F. O. (2022). Customer Segmentation in the Retail Sector: A Data Analytics Approach. In *Proceedings - 2022 14th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2022*, 174–178. https://doi.org/10.1109/IHMSC55436.2022.00048

Shah, K., Rasal, N., & Mhatre, S. (2023). *Demand forecasting in retail for diabetes medicine*. In *Proceedings of the 2023 International Conference on Data Science and Network Security (ICDSNS 2023)*. IEEE. https://doi.org/10.1109/ICDSNS58469.2023.10245677

Trajanova, K., & Dimitrova, J. (2023). *Methods and policies for inventory management. Journal of Sustainable Business and Management Solutions in Emerging Economies*, 28(1), 45–52

Wang, Y., & Minner, S. (2025). Data-driven multi-location inventory placement in digital commerce. *Computers and Industrial Engineering*, *200*. https://doi.org/10.1016/j.cie.2024.110842

Wani, A., Priyanka, M., & Prasath, R. (2023). Unleashing Customer Insights: Segmentation Through Machine Learning. *2023 World Conference on Communication and Computing, WCONF 2023*. https://doi.org/10.1109/WCONF58270.2023.10235136

Xia, H., Pan, X., Zhou, Y., & Zhang, Z. (Justin). (2020). Creating the best first impression: Designing online product photos to increase sales. *Decision Support Systems*, *131*. https://doi.org/10.1016/j.dss.2019.113235

Zaghloul, M., Barakat, S., & Rezk, A. (2024). Predicting E-commerce customer satisfaction: Traditional machine learning vs. deep learning approaches. *Journal of Retailing and Consumer Services*, *79*. https://doi.org/10.1016/j.jretconser.2024.103865