2025 International Conference on Computational Robotics, Testing and Engineering Evaluation (ICCRTEE)

Quick Commerce: Product Price Classification Using Machine Learning Algorithms

Shree Raam Senthilkumar Renuka
Computer Science and Engineering
Sri Ramakrishna Engineering College
Coimbatore, Tamil Nadu
shreeraamsenthilkumarrenuka.2301234@srec.ac.in

Seralathan C A
Computer Science and Engineering
Sri Ramakrishna Engineering College
Coimbatore, Tamil Nadu
seralathan.2301231@srec.ac.in

Sachin D R
Computer Science and Engineering
Sri Ramakrishna Engineering College
Coimbatore, Tamil Nadu
sachin.2301219@srec.ac.in

Maria Liviya L
Computer Science and Engineering
Sri Ramakrishna Engineering College
Coimbatore, Tamil Nadu
marialiviya.2301149@srec.ac.in

Shree Shakthivel Jayakumar
Computer Science and Engineering
Sri Ramakrishna Engineering College
Coimbatore, Tamil Nadu
shreeshakthiveljayakumar.2301274@srec.ac.in

Dr. R. Anuradha
Computer Science and Engineering
Sri Ramakrishna Engineering College
Coimbatore, Tamil Nadu
anuradha.r@srec.ac.in

2025 International Conference on Computational Robotics, Testing and Engineering Evaluation (ICCRTEE) | 979-8-3315-9709-2/25/$31.00 ©2025 IEEE | DOI: 10.1109/ICCRTEE64519.2025.11053083

Abstract -- Quick Commerce (Q-Commerce) is a trend that has led to the revolution of online retail by stressing on ultra-fast delivery within 10-30 minutes. In the current fast-paced market, dynamic prediction of products pricing plays an important role in optimizing pricing strategy and maintaining competitiveness. This paper compares classification methods for predicting products price categories in Q-Commerce, which is based on machine learning. For this, a thorough study is conducted on the data preprocessing and feature engineering and evaluating the classifiers: K-Nearest Neighbors (KNN), Naïve Bayes, Decision Tree, Random Forest, and Support Vector Machine (SVM).

Keyword: Quick Commerce, 10-Minute Delivery, Product Price Prediction, Machine Learning, Data Preprocessing, Feature Engineering.

I. INTRODUCTION

Quick Commerce (Q-Commerce) is the follow-up of e-commerce designed to offer instant deliveries in 10 and 30 minutes. It differs from e-commerce in that it does not work with scheduled delivery, it stresses upon ultra-fast instant delivery on or within predicted time.

A. Importance of Pricing in Q-Commerce

Retail models conventional in their stance vary in the regularity of price swings due to:

1. Dynamic changes in market demand and supply, prices of goods increase or decrease.
2. Real-time competitor price adjustments.
3. The shelf life of a product is a great concern in this industry for products such as groceries.
4. The pricing of the region is set according to the area of purchase and delivery.

B. Q-Commerce Model in India

Q-commerce has emerged as a significant trend in India with well-established food delivery platforms and a few startup companies offering deliveries in 10–15 minutes. The ultra-fast delivery model depends on:

1. Dark Stores: Small, strategically placed warehouses stocked with fast-moving products.
2. AI-Driven Inventory Management: Predicts demand patterns to optimize stock availability.
3. Dynamic Pricing Strategies: Ensures that prices change according to the real time demand.

Proper and effective price forecasting still is one of the biggest challenges in this field. This research helps to solve this problem by using machine learning classification algorithm models to predict the price of a product.

II. LITERATURE REVIEW

Quick Commerce (Q-Commerce) has revolutionized e-commerce by enabling 10–30-minute deliveries, requiring businesses to adopt real-time price prediction and dynamic pricing strategies. Recent advancements in machine learning have significantly improved price forecasting accuracy, optimizing revenue and enhancing customer satisfaction. El Youbi et al. [1] explored machine learning-driven dynamic pricing strategies in e-commerce, emphasizing the role of Gradient Boosting Machines (GBM) in identifying optimal pricing points. Their study demonstrated that GBM outperforms other models in handling complex relationships between pricing factors. Sharma et al. [2] investigated shop price prediction using machine learning, where regression models, decision trees, and neural networks were applied to forecast price variations. Their findings highlighted the importance of historical data and feature engineering in improving price prediction accuracy. Zhu et al. [3] introduced a multi-model fusion strategy for product price prediction, integrating Linear Regression, Decision Trees, and Gradient Boosting. The study found that hybrid models perform better than single-model approaches by effectively managing price fluctuations and improving forecasting reliability. Jiang [4] examined the effectiveness of eight different machine learning models in commodity price prediction. Their study concluded that Random Forest achieved the best performance, while Bayesian optimization enhanced model accuracy. However, sentiment analysis did not significantly improve discounted price predictions. Hennebold et al. [5] analyzed cost prediction in product development using machine learning techniques, highlighting that Decision Trees, K-Nearest Neighbors (KNN), and Support Vector Regression (SVR) can be leveraged for real-time price forecasting. Gu et al. [6] performed a comparative analysis of machine learning methods for asset pricing, showing that tree-based models and neural networks provide better risk premium predictions. Their findings suggest that machine learning can be applied effectively to pricing models in Q-Commerce. Kulshrestha and Saini [7] explored e-commerce market growth prediction using machine learning, applying historical sales data to forecast future income and customer purchase patterns. Their model allowed businesses to optimize inventory management and strategic pricing. Tomitza et al. [8]

conducted a systematic literature review on AI-based dynamic pricing in e-commerce, categorizing approaches based on activity level, application procedure, data foundation, and algorithms. Their research provided a holistic view of AI pricing strategies and highlighted the importance of algorithm selection in different business scenarios. Li [9] investigated e-commerce sales prediction using three machine learning models: Linear Regression, Decision Tree, and Random Forest. The study evaluated model performance using Mean Absolute Error (MAE), Mean Square Error (MSE), and R-squared values, concluding that Random Forest performed the best in forecasting sales. Prakash et al. [10] proposed an ensemble learning-based product price prediction model that analyzes historical pricing trends, sales patterns, and market fluctuations to improve forecasting accuracy. The study emphasized that data-driven pricing strategies can optimize revenue generation and enhance customer satisfaction. Nugraha et al. [11] developed an online price prediction system for consumption commodities, utilizing web scraping to collect data from online retailers. The study demonstrated that automated price tracking can improve price forecasting accuracy and contribute to consumer price index analysis. Prajapat [12] discussed AI-powered e-commerce, outlining how machine learning algorithms are transforming price optimization, demand forecasting, and competitive pricing in online retail. Subbarayudu et al. [13] introduced a novel approach to e-commerce price optimization through machine learning, demonstrating how ML-based algorithms can forecast demand, optimize pricing, and improve revenue generation. Gupta and Pathak [14] examined dynamic pricing in online retail, applying consumer behaviour analytics and demand trends to optimize pricing strategies and maximize revenue. Avinash Kumar Sharma et al. [15] proposed a machine learning-based model for predicting product sales using sentiment analysis and trend forecasting. Their method improved accuracy over traditional regression approaches, helping businesses plan inventory and pricing strategies. [16] shows a method to optimally price a configurable product. [17] Raja Subramanian et al.'s dynamic logistic regression framework for predicting in-play outcomes in cricket can be applied to real-time demand forecasting in quick commerce.

These studies reinforce the critical role of machine learning in Q-Commerce price prediction, showcasing its ability to optimize pricing strategies, enhance profitability, and improve customer satisfaction in real-time.

III. METHODOLOGY

The methodology section showcases the step-by-step process used to process and evaluate the classification algorithms for the price prediction classification. The methodology includes dataset description, data preprocessing, classification models, experimental results

Data Description:

Dataset contains Product Prices. The dataset consists of product details and their historical prices. It includes:

Product Features: Name, Date, Q-Commerce Name, etc.

Pricing Information: Historical price values

Data Collection:

The dataset comprises product price data. The data is sourced various online real-time sources and from the “Comparing Product Prices on various Online platform” dataset.

Data preprocessing:

The preprocessing steps involves:

Removing Irrelevant Columns: Certain columns, such as “SLNO” is removed as they do not contribute to price classification.

Handling Missing Values: Filling missing numerical values with median imputation to avoid bias.

Encoding Categorical Data: Using Label Encoding for non-numeric attributes.

Feature Scaling: Standardizing numerical features for models like SVM and KNN.

Binary Price Classification: Converting price values into two categories (low vs. high).

The dataset is carefully examined and split into 80% for testing set and 20% for training test.

IV. CLASSIFICATION ALGORITHMS

A. K-Nearest Neighbour (KNN)

K-Nearest Neighbors (KNN) is a simple algorithm used for classification. It predicts outcomes based on the majority class (classification) or average (regression) of the K closest data points in feature space. KNN is intuitive, but computationally expensive for large datasets. K-Nearest Neighbors (KNN) is a supervised machine learning algorithm used for classification and regression tasks. It operates by identifying the 'k' closest data points (neighbors) to a specified point based on a distance metric, typically Euclidean distance formula. KNN is non-parametric, it makes no assumptions about the underlying data distribution, and it is also instance-based, storing the entire training dataset for prediction. Its simplicity and effectiveness make KNN popular in pattern recognition, anomaly detection, and recommendation systems, though it can be computationally expensive for large datasets.

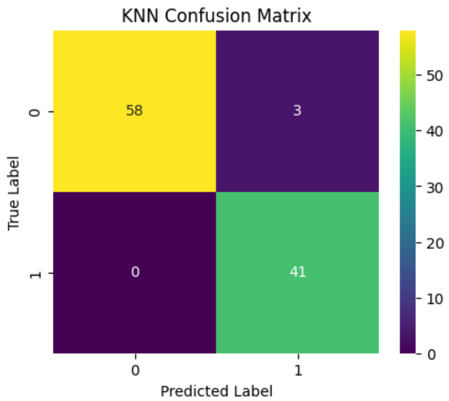

Fig. 1. KNN Confusion Matrix

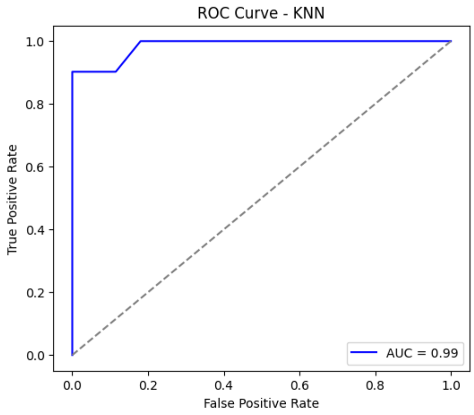

Fig. 2. KNN ROC Curve

Analysis of Fig. 1 and Fig. 2

The KNN confusion matrix (Fig.1) indicates high accuracy, with 58 true negatives and 41 true positives, while only three instances were misclassified. The recall (sensitivity) is perfect, as there are no false negatives, ensuring that all positive cases are correctly identified. Specificity is also high, with only a few false positives.

The ROC curve (Fig.2) further supports the model's strong performance, with an AUC value of 0.99, indicating near-perfect classification ability. This suggests that the model is highly effective at distinguishing between classes.

B. Naïve Bayes

Naïve Bayes' Theorem classification is a probabilistic approach used in machine learning to predict class labels based on prior knowledge. It calculates the probability of a class given input features using Bayes' Theorem. It is a common classifier assuming feature independence, making it efficient for text classification and spam filtering.

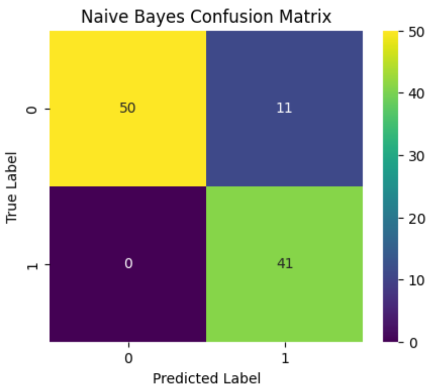

Fig. 3. Naïve Bayes Confusion Matrix

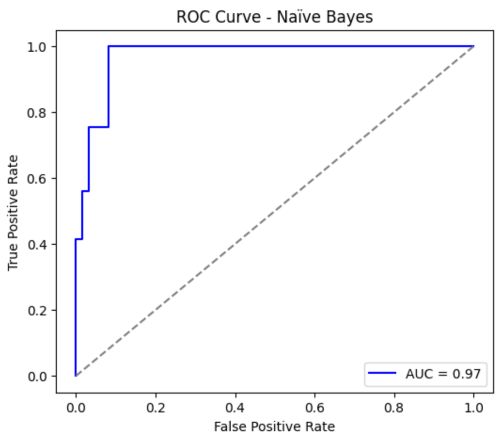

Fig. 4. Naïve Bayes' ROC Curve

Analysis of Fig. 3 and Fig. 4:

The Naïve Bayes confusion matrix (Fig. 3) shows 50 correctly classified class 0 instances and 41 for class 1, with 11 false positives but no false negatives, indicating a slight bias toward class 1.

The ROC curve (Fig.4), with an AUC of 0.97, confirms strong class separation. While accuracy is high, the false positives suggest a trade-off between precision and recall, requiring fine-tuning for specific applications.

C. Decision Tree

A decision tree is a flowchart-like structure used for decision-making and predictive modeling. It splits data into branches based on feature conditions, forming a tree with nodes representing decisions and outcomes. It is widely used in classification and regression tasks due to its simplicity and interpretability.

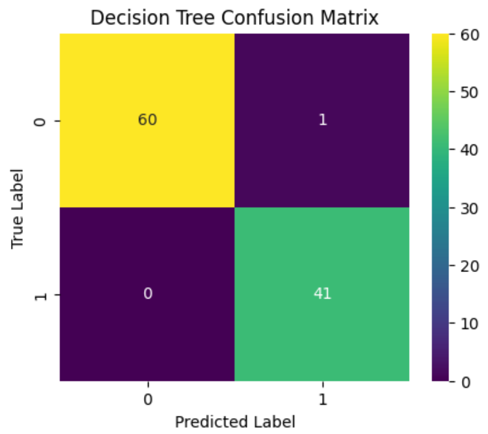

Fig. 5. Decision Tree Confusion Matrix

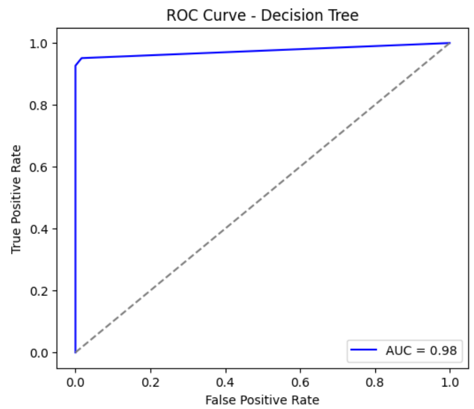

Fig. 6. Decision Tree ROC Curve

Analysis of Fig. 5 and Fig. 6:

The Decision Tree confusion matrix (Fig.5) shows strong performance, correctly classifying 60 instances of class 0 and 41 of class 1, with only one false positive and no false negatives. This indicates high precision and recall.

The ROC curve (Fig.6) confirms this, with an AUC of 0.98, suggesting excellent discrimination between classes. The steep rise in the curve reflects minimal false positive rates, making the Decision Tree a reliable model for classification.

D. Random Forest

Random Forest is an ensemble learning method that builds multiple decision trees using random data subsets. It improves accuracy by combining tree predictions, reducing overfitting, and handling large datasets efficiently. It is widely used for classification and regression tasks due to its robustness and ability to handle missing data.

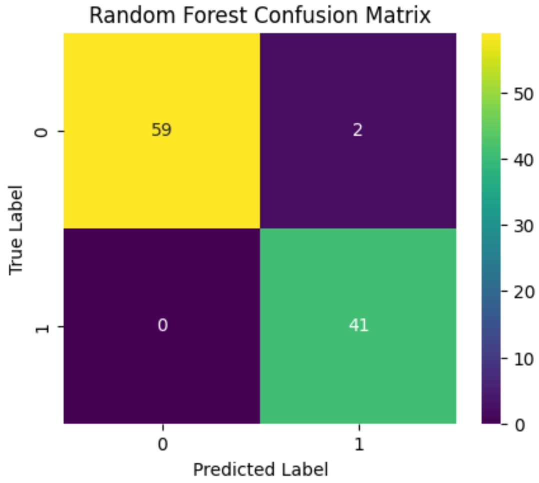

Fig. 7. Random Forest Confusion Matrix

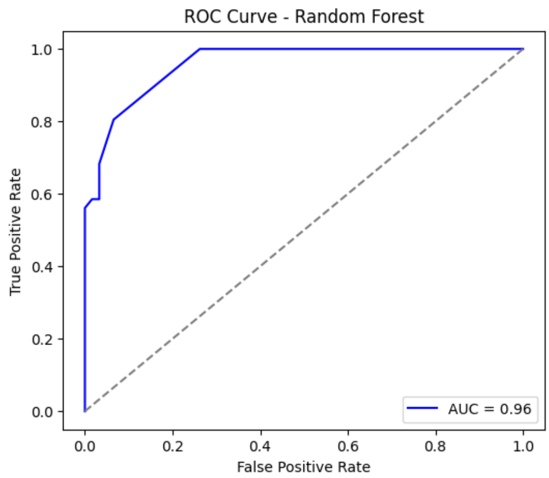

Fig. 8. Random Forest ROC Curve

Analysis of Fig. 7 and Fig. 8:

The Random Forest confusion matrix (Fig.7) shows that the model correctly classified 59 instances of class 0 and 41 instances of class 1, with only 2 misclassifications of class 0 as class 1. No class 1 instances were misclassified, indicating strong sensitivity.

The ROC curve (Fig.8) has an AUC of 0.96, signifying high classification performance. The steep curve suggests the model maintains a low false positive rate while effectively distinguishing between classes. Overall, the Random Forest model demonstrates strong accuracy, with minimal misclassification and a high AUC score, making it a reliable choice for classification.

E. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised learning algorithm that uses an optimal hyperplane to separate classes in a dataset. It maximizes the margin between different the classes, improving classification accuracy and generalization. SVM is effective for both linear and non-linear classification by using kernel functions to transform data into a higher-dimensional space.

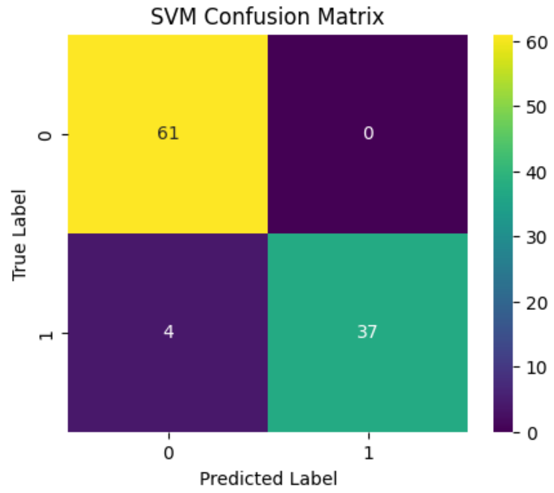

Fig. 9. SVM Confusion Matrix

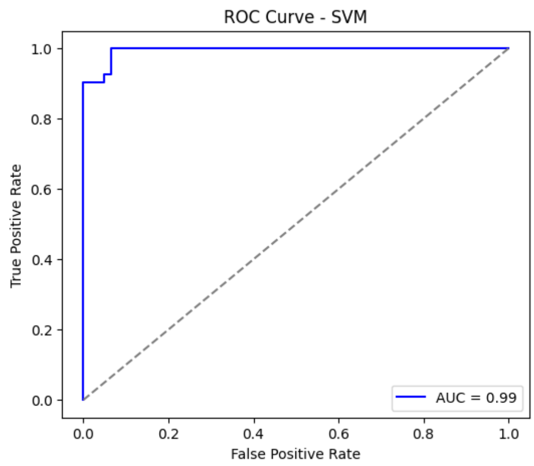

Fig. 10. SVM ROC Curve

Analysis of Fig. 9 and Fig. 10:

The SVM confusion matrix (Fig.9) shows that the model correctly classified all 61 instances of class 0 and 37 instances of class 1. However, 4 instances of class 1 were misclassified as class 0, leading to some false negatives.

The ROC curve (Fig.10) indicates a strong performance with an AUC of 0.99, suggesting excellent discrimination between classes. The model achieves a low false positive rate while maintaining a high true positive rate.

SVM performs well, though the false negatives may need attention, especially if missing positive cases is critical for the application.

V. Experimental Results

Performance metrics are quantitative measures used to evaluate the effectiveness of a machine learning model, particularly in classification and regression tasks. These metrics help assess how well a model predicts outcomes and whether it meets the desired accuracy and reliability.

In classification problems, key performance metrics include:

A. Accuracy

Measures the overall correctness of the model. It is given by:

Accuracy = Correct predictions / Total predictions

B. Precision

Measures how many of the predicted positive instances are actually positive. It is given by:

Precision = TP / (TP + FP)

C. Recall (Sensitivity)

Measures how many actual positive instances were correctly identified. It is given by:

Recall = TP / (TP + FN)

D. Specificity

Measures how well the model identifies negative instances correctly. It is given by:

Specificity = TN / (TN + FP)

E. F1-score

The harmonic mean of precision and recall, balancing both metrics. It is given by:

F1 Score = 2 × (Precision × Recall) / (Precision + Recall)

TABLE I. PERFORMANCE METRICS

The five classification models are evaluated on the basis of the performance metrics and their confusion matrix is plotted successfully and the results are compared.

Performance Analysis:

The performance comparison of classification algorithms highlights Decision Tree and Random Forest as the top performers, achieving high accuracy, precision, recall, and specificity. Decision Tree leads with an accuracy of 0.99, making it the most balanced model. Random Forest follows closely with 0.98 accuracy and strong overall metrics. K-Nearest Neighbors (KNN) also performs well with 0.97 accuracy, though it has slightly lower specificity. Support Vector Machine (SVM) excels in specificity (1.00), meaning it perfectly identifies negative instances but has a slightly lower recall (0.95). Naïve Bayes, with the lowest accuracy (0.89), struggles due to its simplifying assumptions, making it less reliable compared to other models. Overall, Decision Tree and Random Forest are the best choices for this dataset, while SVM is suitable when minimizing false positives is a priority.

Error Analysis:

In the Quick Commerce price classification, error analysis helps identify where the model fails to generalize effectively to unseen products. Misclassifications often highlight gaps in the model's ability to distinguish between overlapping or ambiguous price ranges, especially for products with similar features. These errors suggest that the model may be overfitting patterns in the training data, limiting its generalization capability. By analyzing such errors, we can improve the model's robustness by adjusting feature selection, addressing class imbalance, or introducing regularization technique such as to enhance its performance on real-world, diverse product data.

VI. CONCLUSION

Quick Commerce is changing the way people shop, with 10-minute. To keep up this new trend, businesses need smart pricing models that adjust quickly to demand, competition, and market trends. This study explored machine learning models for price prediction and found that Decision Tree performed best.

This study analyzed multiple machine learning models to optimize Q-Commerce pricing. Among the evaluated models, Decision Tree and Random Forest demonstrated the highest accuracy (0.99 and 0.98, respectively), making them ideal for predictive tasks in a fast-paced commerce environment. SVM, with its perfect specificity (1.00), effectively reduced false positives, ensuring that incorrect classifications were minimized. Meanwhile, Naïve Bayes, with the lowest accuracy (0.89), struggled due to its probabilistic assumptions, which may not align well with real-world transactional data in Q-Commerce.

ALGORITHM NAME	ACCURACY	PRECISION	RECALL (SENSTIVITY)	SPECIFICITY	F1 SCORE
KNN	0.97	0.97	0.98	0.95	0.97
BAYES	0.89	0.89	0.91	0.82	0.89
DECISION TREE	0.99	0.99	0.99	0.98	0.99
RANDOM FOREST	0.98	0.98	0.98	0.97	0.98
SVM	0.96	0.97	0.95	1.00	0.96

These findings show the potential of machine learning in enhancing operational efficiency and decision-making within the rapidly evolving Q-Commerce industry. Implementing the right models can significantly improve order fulfillment, fraud detection, and overall customer experience.

V. FUTURE ENHANCEMENT

This current study focused on binary price classification, future research can explore multi-class price prediction models, Real-time adaptive pricing strategies, Deep learning-based models (LSTMs, Transformers), Integration of external factors such as inflation, supply chain disruptions, and competitor pricing to improve the model capabilities.

While the current study highlights the effectiveness of machine learning in Q-Commerce, several future enhancements can further improve its performance and scalability. Deep learning models, such as neural networks, can be integrated for more precise demand forecasting by capturing complex patterns in customer behavior. Real-time analytics using big data and AI can help businesses make instant decisions on inventory management and order processing, reducing delays.

Reinforcement learning techniques can be employed to optimize last-mile delivery logistics, dynamically adjusting routes and resource allocation based on live traffic data. Additionally, personalized customer experiences can be enhanced using AI-powered recommendation systems to improve engagement and retention. Lastly, blockchain technology can be introduced to enhance transparency and security in transactions, ensuring a reliable and fraud-resistant supply chain. Implementing these advancements will make Q-Commerce more efficient, scalable, and customer-centric, further driving its growth in the competitive digital marketplace.

REFERENCES

[1] R. El Youbi, F. Messaoudi, and M. Loukili, "Machine Learning-driven Dynamic Pricing Strategies in E-Commerce", *2023 14th International Conference on Information and Communication Systems (ICICS)*, IEEE, Nov. 2023.

[2] A. K. Sharma, D. Sharma, M. Yadav, and P. Singh, "Shop

Price Prediction Using Machine Learning", *International Journal of Innovative Research in Technology*, vol. 11, no. 1, pp. 1068–1073, June 2024.

[3] D. Zhu, X. Chen, and Y. Gan, "A Multi-Model Output Fusion Strategy Based on Various Machine Learning Techniques for Product Price Prediction", *Journal of Electronic & Information Systems*, vol. 4, no. 1, pp. 42–51, Apr. 2022.

[4] K. Jiang, "Effective E-Commerce Price Prediction with Machine Learning Technologies", *ACM Digital Library*, Aug. 2024.

[5] C. Hennebold, K. Klöpfer, P. Lettenbauer, and M. Huber, "Machine Learning-Based Cost Prediction for Product Development in Mechanical Engineering", *Procedia CIRP*, vol. 107, pp. 264–269, 2022.

[6] S. Gu, B. Kelly, and D. Xiu, "Empirical Asset Pricing via Machine Learning," *The Review of Financial Studies*, vol. 33, no. 5, pp. 2223–2273, May 2020.

[7] S. Kulshrestha and M. L. Saini, "Study for the Prediction of E-Commerce Business Market Growth Using Machine Learning Algorithm", *2020 5th IEEE International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, IEEE, Dec. 2020.

[8] C. Tomitza, U. Ibrahimli, and L.-V. Herm, "AI-Based Methods of Dynamic Pricing in E-Commerce: A Systematization of Literature", *The 32nd European Conference on Information Systems (ECIS 2024)*, June 2024.

[9] X. Li, "Predicting E-Commerce Sales with Three Machine Learning Models", *Atlantis Press*, Oct. 2024.

[10] V. Prakash, S. Sahil, H. Rayi, P. N. Reddy, and B. Maridu, "Product Price Prediction Using Ensemble Learning Techniques," *2024 8th International Conference on Inventive Systems and Control (ICISC)*, IEEE, July 2024.

[11] N. S. Sutiawan and I. G. B. Nugraha, "Online Price Prediction System of Consumption Commodities," *2017 International Conference on Information Technology Systems and Innovation (ICITSI)*, IEEE, Oct. 2017.

[12] R. Prajapat, *AI-Powered E-Commerce: How Machine Learning Is Transforming Online Shopping*, Springer, 2024.

[13] Y. Subbarayudu, G. Vijendar Reddy, M. Vamsi Krishna Raj, and K. Uday, "An Efficient Novel Approach to E-Commerce Retail Price Optimization Through Machine Learning," *E3S Web of Conferences*, June 2023.

[14] R. Gupta and C. Pathak, "A Machine Learning Framework for Predicting Purchase by Online Customers Based on Dynamic Pricing, "*Procedia Computer Science*", vol. 36, pp. 599-605, 2014.

[15] A. K. Sharma, N. Goel, J. Rajput, and M. Bilal, "An Intelligent Model for Predicting the Sales of a Product", *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, IEEE, Jan. 2020.

[16] Angel C.Hernandez; David Masaryk, Juraj Mecir and Siamak Saliminejad, "Optimal Pricing of Configurable Products using Machine Learning", *2024 11th IEEE Swiss Conference on Data Science (SDS),* IEEE, September 2024.

[17] R. Raja Subramanian, P. Vijaya Karthick, S. Dhanasekaran, R. Raja Sudharsan, S. Hariharasitaraman, S. Rajasekaran and B. S. Murugan, "Regression to Forecast: An In-Play Outcome Prediction for One-Day Cricket Matches", "*International Conference on Cognitive and Intelligent Computing*", pp 43–55, Springer, Jan 2023.