Sensors and Materials, Vol. 38, No. 3 (2026) 1517–1536
MYU Tokyo

S & M 4391

Multisource Sensing Data and AI Models for Online Commerce Analysis

Xinwei Wu[1] and Pingan Hu[2*]

[1]Wuxi University of Technology, Wuxi 214121, China
[2]Hanyang University, Seoul 15588, Korea

(Received June 5, 2025; accepted March 4, 2026)

Keywords: online commerce, data fusion, AI, IoT, real-time analytics

Online trading increasingly depends on the fusion of heterogeneous data sources to optimize logistics, forecasting, and decision-making. In this study, we developed a multisource sensing and AI-driven framework that integrates physical sensors (radio frequency identification, near field communication, micro-electro mechanical systems, and resistance temperature detectors), environmental sensors (DHT22), and virtual sensors that capture behavioral and sentiment data from social media. A three-layer perception architecture was implemented using a multinode sensor network controlled by Espressif 32 microcontrollers and the Message Queuing Telemetry Transport protocol, achieving low-latency transmission (<100 ms) across a dataset exceeding 10 million records. Advanced AI models, including neural networks and ensemble methods, were applied to fuse IoT, commerce platform, and social media data. The results of this study demonstrated significant improvements in demand forecasting accuracy, fraud detection, and logistics optimization compared with traditional single-source analytics. The optimized routing algorithms increased the on-time shipment rate from 71 to 97%, while the sentiment analysis of more than 30000 monthly mentions provided important market information. The real-time processing capability ensures scalability and resilience, addressing data heterogeneity and computational challenges. This integrated sensing data fusion framework establishes a foundation for more competitive, responsive, and efficient online commerce systems.

1. Introduction

The exchange of products and services using digital technologies has become a significant factor behind economic growth. Traditional online commerce and online services, such as cloud computing, digital content, and banking services, have contributed to the global transformation of business activities and interactions.[(1)] In 2020, online commerce comprised about one quarter of the global trade, which increased particularly in the Indo-Pacific area.[(1)] The transition to online commerce has contributed to economic development as internet connectivity, smart device usage, and AI have accelerated data exchange in online commerce. However, rapid development in online commerce faces challenges such as data privacy, cybersecurity,

*Corresponding author: e-mail: hpa1224@hanyang.ac.kr
https://doi.org/10.18494/SAM5798

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

intellectual property rights, complex compliance, and fragmented data regulations. These challenges obstruct data flows between countries,[2] along with a lack of stable internet access and digital abilities.[3] After the COVID-19 pandemic, digital platforms have been more widely used, but they exhibited weaknesses in digital supply chains and networks. To solve these problems, international cooperation is necessary to formulate laws, support innovations, and protect the stakeholders' interests.

Even in digital trading, stakeholders' individual needs for automation are becoming more important than before. With AI analytics, companies provide customers with personalized services.[4] Mobile platforms, cryptocurrencies, and biometrics have been widely adopted for secure online trading and supply chains, which contribute to reduced expenses and efficient goods delivery.[4] These technologies require diverse data sources and AI to merge data to improve performance and outcomes. However, the related process is complicated and demands sophisticated methods to obtain useful and accurate information. Online trading increasingly relies on data fusion from diverse sources for rapid and accurate decision-making. The e-commerce ecosystem with data fusion requires a complex, heterogeneous sensor network that collects transaction data as well as high-precision physical sensors, such as radio frequency identification (RFID) and micro-electro mechanical systems (MEMS) sensors, to provide real-time information on logistics and warehousing statuses.[5] By integrating the physical sensor data with sensing data from digital platforms, data heterogeneity can be addressed in related models, and real-time responsiveness in online trade environments is enhanced.[6]

Data from IoT devices, electronic marketplaces, social networking, and environmental sensors provide different but useful information.[4] The fused data is analyzed to provide essential information on inventory, shipments, and online commerce records, as well as user activities and behaviors. Social media data also needs to be analyzed to obtain insight into current trends in commercial activities. To fuse data from different sources, disinformation and inconsistencies of the data must be avoided, which necessitates efficient analytics and data collection methods.

Data fusion is essential in online commerce to increase the number of consumers, supply goods, and predict and hedge risks by making informed decisions. By fusing multisource data, suspicious and untrustworthy activities can be detected to ensure stable demand or supply.[3] By using machine learning (ML) technology, we developed a data fusion method to optimize online commerce and ensure efficiency.[4]

In this study, we adopted high-precision technical sensors, such as DHT22 capacitive humidity sensors and near-field communication (NFC)/RFID tags. By treating the trading system as a heterogeneous sensor network, we demonstrated how the technical precision of these sensors directly drives the performance of downstream AI analytics. By embedding decision-making capability into the sensor network, technical sensing mechanisms can be a useful data collection tool and an active enabler of operational, strategic, and consumer-facing decisions in online commerce.

2. Materials and Methods

2.1 Sensing mechanisms and perception layer

In online trading, a three-layer perception architecture is defined to integrate physical, environmental, and virtual sensing data.

1. Physical sensing layer
   RFID tags and NFC devices are used to track inventory and shipments across the supply chain. Operating on the principle of electromagnetic induction, these sensors enable non-line-of-sight identification, ensuring high-fidelity data capture in complex logistics environments. Such mechanisms reduce manual errors, enhance transparency, and provide real-time visibility into product movement, which is critical for fraud detection and demand forecasting.[7]
2. Environmental monitoring layer
   For the storage and transport of perishable or high-value goods, a system that integrates MEMS-based capacitive humidity sensors and resistance temperature detectors (RTDs) is widely used. MEMS sensors are mainly used to monitor changes in humidity, whereas RTDs are used to track temperature variations by delivering precise temperature readings. Humidity and temperature are essential for predictive maintenance, quality assurance, and risk mitigation, ensuring that goods remain under optimal conditions throughout transport along the supply chain.[8]
3. Virtual sensing layer
   Beyond physical sensors, virtual sensors are used to collect data, including user interaction rates, transaction frequencies, and browsing behaviors, which are treated as sensor-derived data. This enables the system to perceive market temperature and consumer sentiment with the same granularity as environmental monitoring. By fusing these virtual signals with IoT data, the architecture enables the holistic monitoring of logistics conditions and consumer dynamics to enhance the adaptability of AI models in real-time commerce environments.[9]

This layered architecture ensures that heterogeneous data sources from tangible sensor readings to intangible behavioral signals are harmonized into a unified analytics pipeline. The integration of physical, environmental, and virtual sensing enhances operational resilience and multisource fusion models to deliver superior accuracy in forecasting, anomaly detection, and logistics optimization.

The sensor network contributes to decision-making related to logistics and supply chain. RFID-based monitoring supports route optimization and shipment scheduling, allowing managers to reroute deliveries in real time when disruptions occur. For inventory management decisions, stock levels are monitored for replenishment planning to prevent stockouts and overstocking. Humidity and temperature sensors provide information on predictive maintenance and storage adjustments for quality assurance, ensuring perishable goods remain within safe thresholds. For fraud detection and security, anomalies in transaction data streams flag suspicious activities, prompting fraud alerts and verification protocols. Virtual sensing of transaction frequencies and social media sentiment is used to guide promotion timing, customer segmentation, and brand positioning strategies for decision-making related to marketing and consumer behavior.

2.2 Data collection and fusion

In this study, data from IoT devices, online commerce platforms, and social media networks were integrated through a three-layer perception architecture to support warehousing, logistics, and market system analysis (Fig. 1).

In the physical sensing layer, sensors embedded in containers, such as RFID tags and NFC devices, transmitted inventory and shipment data to RFID readers and gateways. These data were used for non-line-of-sight identification to reliably monitor the overall process of the cold chain logistics. The collected high-fidelity data streams captured goods location and inventory status, forming the foundation for supply chain monitoring and fraud detection.(10)

In the environmental monitoring layer, MEMS-based capacitive humidity sensors and RTDs were deployed in warehouses and transport systems. These sensors measured temperature and humidity variations, enabling predictive maintenance, quality assurance, and safeguarding of perishable and high-value goods. Real-time monitoring of environmental conditions ensured that fragile items remained within safe thresholds throughout the supply chain.(11)

In the virtual sensing layer, platform-derived data, including transaction frequencies, browsing behaviors, and user interaction rates, were treated as virtual sensor data. This method enabled the system to perceive consumer sentiment with the same granularity as environmental monitoring. Social media posts, user reviews, and comments were analyzed to extract sentiment indicators, providing insights into consumer preferences and brand recognition.(5)

The integration of the three layers formed a collective sensing ecosystem. Physical and environmental sensors captured the tangible aspects of logistics, while virtual sensing extended perception into the digital marketplace. Such heterogeneous data sources were processed to

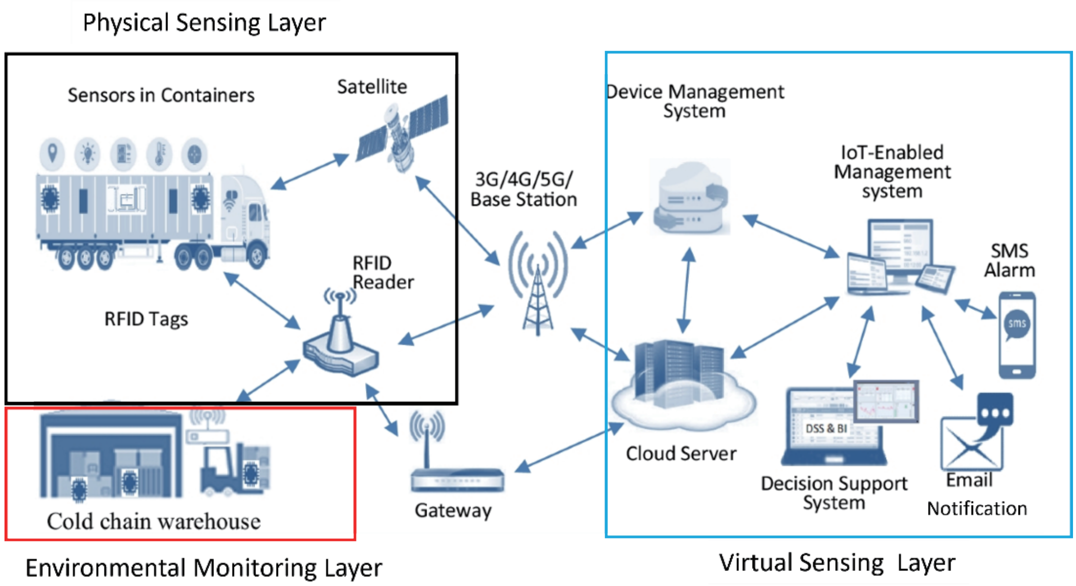

Fig. 1. (Color online) Data collected from IoT sensors.

remove missing details, detect outliers, and normalize values, ensuring consistency across units and formats.[5] Sensitive data from online commerce and social media were protected throughout the process.[12]

To ensure data fidelity and system scalability, we implemented a multinode sensor network. Each sensor was controlled by an Espressif 32 microcontroller [32-bit dual-core, 240 megahertz (MHz)], which was selected for its Wi-Fi and Bluetooth capabilities.[13] The DHT22 (Aosong Electronics, China) humidity and temperature sensor was used to obtain calibrated digital output with a temperature accuracy of <±0.5 °C and relative humidity accuracy of ±2%. The sampling interval was 30 s to balance data granularity with power consumption.[14] MFRC522 RFID modules of NXP Semiconductor, the Netherlands, operating at 13.56 MHz were used at warehouse entry and exit points to read the International Organization for Standardization and the International Electrotechnical Commission 14443A tags, providing event-triggered data streams whenever a package was moved.[15]

Data was encapsulated into JavaScript Object Notation packets and transmitted via the Message Queuing Telemetry Transport (MQTT) protocol to a central gateway.[16] MQTT's lightweight publish–subscribe model ensures low-latency delivery (typically <100 ms) in high-traffic environments, which is essential for the real-time decision-making shown in this study. Digital signals from the sensors are processed to remove noise using a moving average filter before being stored in a time-series database for AI model training. By fusing multisource data streams within this layered architecture, robust predictive analytics were ensured, enabling accurate demand forecasting, fraud detection, and optimized logistics operations.

In this study, we used a dataset containing more than 10 million records and combined the data using a multilayer data fusion method that incorporated statistical and ML algorithms at the data, feature, and decision levels.[5] Data from multiple sources were weighted in accordance with their reliability and high quality using Eq. (1). The data noise was reduced to prevent interference with data analytics and enhance the accuracy of analysis.

$$F = \sum_{i=1}^{n} w_i . x_i \tag{1}$$

Here, F stands for the combined value, x_i stands for single data from each source, and w_i represents weights that are confidence scores of the data from each source.[10]

In feature fusion, features were extracted and concatenated, and then their dimensions were reduced using principal component analysis (PCA).

$$Z = XW \tag{2}$$

Here, X represents the original set of features, W represents the PCA, and Z is the matrix for the reduced number of features. Through feature fusion, the data without flaws can be used in AI models to obtain better information.

2.3 AI analytics

Decision trees, artificial neural networks (ANNs), and support vector machines (SVMs) were used in this study to process multiple-source data, which can be nonlinear.[5] Backpropagation was also used to optimize ANNs, which contained hidden layers in the feedforward structure. The network's final output was obtained using

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right). \tag{3}$$

Here, x_i is the input feature, w_i is the weight, b is the bias, and f is the activation function, such as rectified linear unit or sigmoid. As a result, the connections between features and variables in fused data were identified for forecasting outcomes or fraud detection.

2.4 Evaluation metrics

In this study, we define the on-time shipment rate (OTSR) as a key performance indicator of logistics efficiency. OTSR is calculated using[17]

$$OTSR = \frac{S_{on\text{-}time}}{S_{total}} \times 100\ (\%), \tag{4}$$

where $S_{on\text{-}time}$ represents the number of orders successfully dispatched from the warehouse within the predefined target window (24 h from order placement), and S_{total} is the total number of orders processed during the same period. This metric is derived by cross-referencing the real-time goods location data from RFID sensors with the transaction timestamps from the e-commerce platform.

To evaluate the efficiency of data fusion and ML models, accuracy, precision, recall, and F1-score were calculated as the evaluation metrics of classification and regression. To determine the accuracy of the model, the ratio of true positives to all predictions was calculated. Fraud detection involves outliers caused by fraudulent actions and accuracy in prediction. Precision is the proportion of true positives in all positive predictions, and recall is the percentage of true positives to the sum of true positives and false negatives. F1 score is a measure of a model's accuracy in prediction evaluation, and is a balance between precision and recall.

Mean absolute error (MAE), root mean squared error ($RMSE$), and R^2 were used to evaluate the performance and fit of models in forecasting demand and predicting prices. MAE is calculated to measure the average absolute difference and the extent of error in prediction. $RMSE$ is used to penalize errors exceeding MAE, which is useful in detecting big mistakes. When the coefficient of determination (R^2) is close to 1, the model accounts for a large proportion of the variance in the dependent variable. R^2 is commonly used to assess how well a regression model fits the data, but it does not imply causation or guarantee that the model is the correct one.

Since online commerce happens in real time, it is necessary to ensure that the ML model efficiently processes a large dataset. The amounts of latency, memory required, and throughput need to be estimated to judge whether data fusion and AI models are adequate for analyzing a high volume of data. The appropriate algorithms might be slow, although they are accurate, which is not always appropriate for online commerce data analysis. As various datasets are used in cross-validation to check if the generalizability of model is high enough to handle different data, the evaluation metrics need to be measured to evaluate the outcomes of the ML models in online commerce.

2.5 Algorithm optimization

ML algorithms need to be optimized to enhance accuracy, efficiency, and scale for multisource data in online commerce. ML algorithms are optimized through data fusion by evolutionary and heuristic methods.[11] Genetic algorithms (GAs) are widely used to enhance the model's hyperparameters (learning rate) and structures (feature subsets and model architecture) through a similar process of evolution to find optimal solutions to complex problems. They find solutions by creating chromosomes that represent answers, using crossover and mutation, and choosing the best individual by scoring each one on the basis of its accuracy. GAs are used for optimizing neural networks and similar models in online commerce.

We also employed recursive feature elimination (RFE) and Lasso regression to select the most important features and improve the model's learning ability. These methods help identify and remove unhelpful features and prevent the model from memorizing training data. Boosting and bagging methods were also conducted to increase the effectiveness of models by helping weak learners resist noise and enhance accuracy. In addition, reinforcement-learning-based optimization was used to modify the settings in response to changes in the data. Reinforcement-learning-based optimization is widely used for predicting unexpected changes in traffic and consumer demands.

2.6 Real-time data processing

In online commerce data analysis, the immediate and fast analysis of real-time data from various sources is essential. Therefore, we used Apache Kafka and Apache Flink to design a data processing system as they handle many events and faults and diverse types of data from IoT, online commerce, and social networks. Such quickly running models were installed and activated at the edge and cloud server in this study. Various compression methods, such as pruning and quantization cuts, were used to run them on low-cost devices. Reliable and scalable analysis was conducted by using load balancing and fault tolerance when the real-time data collection was not stable. Kalman filtering was applied in data fusion to improve the real-time data quality. The Kalman filter repeats calculations and updates the results as streamed data. By these data processing methods, online commerce platforms had improved operational efficiency and reacted quickly to changes in the market, supply chain, and consumers.

3. Data Fusion

Data fusion is used to merge different types of data and increase the accuracy and reliability of ML model predictions. Weighted averaging and Kalman filtering were once widely used fusion methods, but they have recently been replaced by Bayesian inference and Dempster–Shafer theory. These enable independent networks to use local details and share data, increasing the system's smartness and capability to function even during emergencies.[18] Bayesian fusion is used to build a framework that continuously updates sensor datasets as new data becomes available.[19] Bayesian fusion is efficient in online commerce because the data becomes unreliable regardless of collection time.

Features are fused in several stages, which enables the discovery of complex patterns hidden in different data types. Convolutional neural networks and recurrent neural networks with long short-term memory (LSTM) are often employed to fuse information collected from IoT devices, such as transaction and social media data.[20] Unlike traditional models, these neural networks record nonlinear associations and time-changing patterns. They are trained to adapt to newly added data from different sources to improve accuracy, especially when the data has noise or is incomplete.[19] They are significant in online commerce data analysis since the diverse forms and quantities of data present huge challenges.

Multifusion methods have been proposed at the data, feature, and decision levels to offer different benefits. Early removal of noise and redundancy in data fusion leads to effective feature extraction and decision-making by integrating individual models' predictions.[18,20] The integration of multiple methods ensures forecasting in smart logistics and online commerce. Improved fusion methods, such as ensemble and metalearning, make the trade process efficient and enable large data management.

3.1 Optimization of ML algorithms

Optimizing ML algorithms enables the best results in predicting and processing. To determine the optimal learning rates, hyperparameter tuning methods, such as grid search, random search, and Bayesian optimization, are used.[21] These methods allow models to avoid biases and fit data better in online commerce. GAs and particle swarm optimization (PSO) play an important role in optimizing neural networks (Fig. 2) and solving multidimensional problems.[22]

Reducing dimensions and selecting significant variables significantly improve the effectiveness of ML models. RFE, Lasso regression, and PCA are used to preserve essential data, which makes it possible to use simpler computations. This is important in analyzing data from IoT sensors, online commerce, and social media, as it contains numerous attributes. RFE, Lasso regression, and PCA are integrated to make weak learners develop sturdy, accurate ML models that can cope with noisy information. Many online commerce companies use extreme gradient boosting (XGBoost) and similar algorithms because they perform better in online commerce.

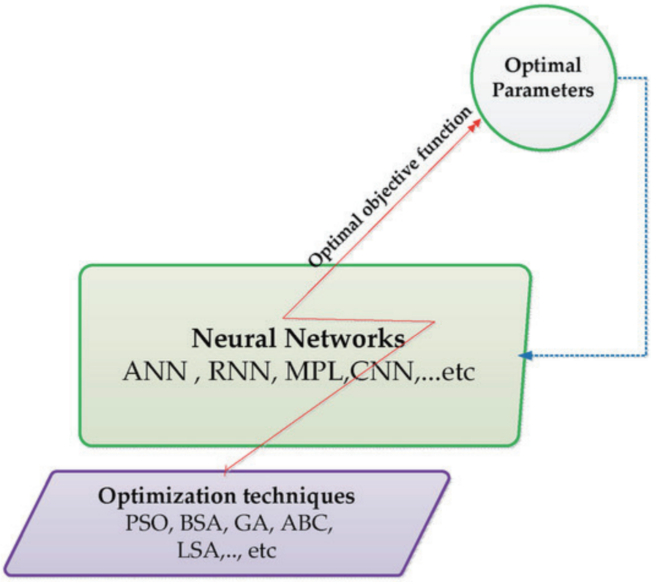

Fig. 2. (Color online) Method of optimizing neural networks (MPL: multi-protocol label switching, RNN: recurrent neural network, CNN: convolutional neural network, BSA: bee swarm algorithm, LSA: latent semantic analysis, and ABC: artificial bee colony).

Recently, reinforcement learning and metalearning have gained popularity in optimizing ML models. Reinforcement learning allows online commerce applications to learn the best parameters as they adapt quickly to unexpected changes in the data. Metalearning is used to adapt to new situations using the knowledge previously learned and is mainly applied to online commerce platforms.[21] Such learning methods were used in this study to increase the model's ability to fuse data in online commerce.

3.2 Data fusion strategy in online commerce

Data fusion methods were used to integrate data from multiple sources and produce optimal outcomes in online commerce in this study. In weighted data fusion, the reliability and relevance of each data type are assessed, and weights are determined on the basis of updated data using adaptive algorithms.[18] In adaptive weighted data fusion, data from feedback and error measurements are used to revise weights.[20] In hierarchical fusion, data streams are fused to lower the workload involved.[18] In ensemble fusion, outputs from various models for using various data types are fused. Accuracy and bias are decreased as majority voting is adopted with predictions by stacking or blending the results.[20] These fusion methods are effective in processing unpredictable data in online commerce. The data fusion is conducted in a layered structure in the methods. The lower layers of a neural network are responsible for sharing information, while the upper layers classify data types for the final output.[10] Such a modular structure enables the easy and quick processing of trade data on a large scale.

Deep learning and XGBoost were also used to predict buyers in models, improving the model's performance in individual approaches. For efficient data fusion, we decided on how to prioritize accuracy, time, and computation load for real-time online commerce data analysis. To ensure efficient and effective fusion, multi-objective optimization methods with different data fusion methods were used.[20]

3.3 Real-time data processing

The significant feature of online commerce is its reliance on real-time data for immediate use. New features identified in continuously added data using Apache Kafka and Apache Flink allow companies to process data from multiple sources at high speed,[21] which enables microbatch processing and event-driven modeling to control the speed and number of data processed in various applications. By using Apache Kafka and Apache Flink, asynchronous and inconsistent data from various sources are simplified.[10] The ML models process and fuse data quickly in real time. By using methods such as pruning, quantization, and knowledge distillation, the size and the need for computational resources of a model are reduced, maintaining accuracy.[21] These methods contribute to the improvement of the architecture designs of the edge server since they enable calculations to be faster with less data transfer. They also enable parallel and distributed computing to fuse a significant amount of online commerce data. This approach enables the models to perform better with inconsistent data. The models can use workload-based resource allocation to maintain a good quality of analysis,[20] being reliable with increased fault tolerance and appropriate recovery methods to prevent interruptions caused by hardware trouble or network issues. Using such methods, digital trading platforms easily manage data fusion.

Data fusion and optimization lead to accurate forecasting of demands, organizing of shipments, and prediction of customer habits. By integrating long- and short-term data through various fusion methods, the ML models effectively address the limitations of traditional RNNs. For example, metalearning improves the accuracy of forecasting algorithms in offline and online applications. By using deep learning and XGBoost, ML models with ensemble data fusion predict buyers of stocks and predict prices accurately.[5] In logistics, more goods are delivered faster at reduced costs. In particular, scheduling and route selection can be optimized by using random forests (RFs), XGBoost, and ANNs with feature-level data from diverse data sources (Fig. 3). GAs with data fusion enable cross-border online commerce logistics to be more efficient by improving routes and detecting potential hazards.[5] Issues in the digital global trade route are predicted, which lead to effective customer relations management.

Customers' online purchase behavior is also predicted by combining different learning techniques. Logistic regression, SVMs, and XGBoost are used to predict purchase behavior and preferences more effectively than individual algorithms. By analyzing user behavior, the ML models predict the optimal place and time for purchases on the basis of various criteria, including social and environmental factors. Additionally, companies can identify and select reliable partners for their online commerce operations.[5]

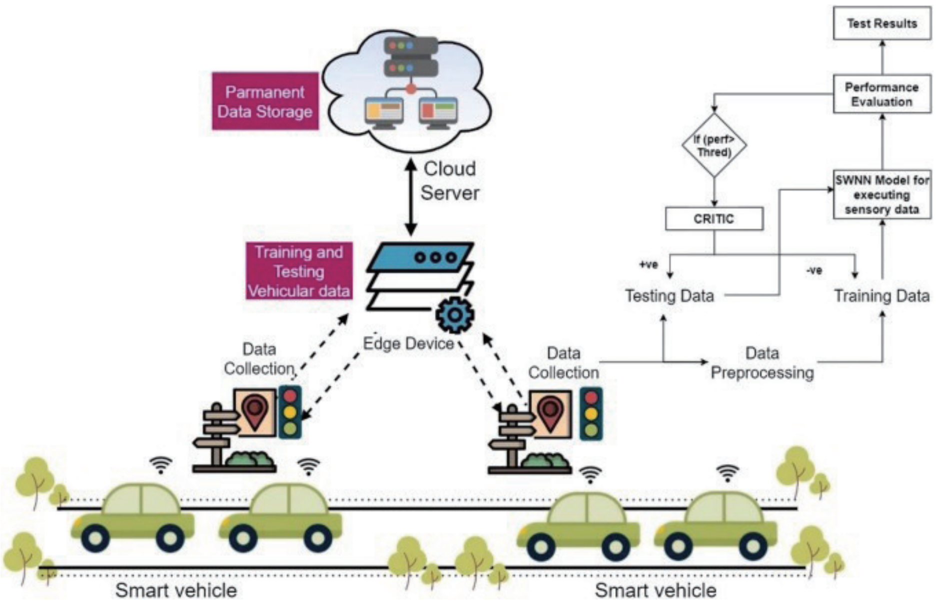

Fig. 3. (Color online) ML models used in logistics industry.

4. Results and Discussion

In this study, we examined how sensor data and ML algorithms are used in online commerce and how to fuse data from IoT, online commerce, social media networks, and the environment, which were collected over one year. ML models were used to optimize data fusion strategies to define spatial and temporal trends. The results can be used to benefit stakeholders by improving the supply chain, forecasting customer behavior, and identifying fraud. The developed method, based on fused data and AI analytics, is expected to improve the efficiency of online commerce operations.

4.1 IoT data

The data were gathered from IoT sensors installed in logistics and warehousing facilities from January to May 2024. Table 1 presents that the average daily temperature was around 21 °C and increased by 0.7 °C from January to March, but dropped to a similar level in January. Daily temperature fluctuated from 14.1 to 29.4 °C. The average humidity was in the range from 52% (in March) to 57.9% (in February). April was when humidity levels were lowest, at only 32% near midnight, and as high as 85.8% at noon. Because of such results, monitoring the environment helps maintain good quality, especially for fragile items. The inventory level was maintained at around 500 units. The minimum was 444 in February, while the maximum reached 567 units in May. The consistency in inventory indicated that the company managed its stock effectively through the use of IoT technology to respond swiftly to any changes in demand. Maximum and minimum levels in each month are important in organizing the supply chain and

Table 1
IoT sensor data from January to May 2024.

Month	Temperature			Humidity			Inventory level			OTSR
	Ave.	Min.	Max.	Ave.	Min.	Max.	Ave.	Min.	Max.	
January	21.4	16.3	26.7	56.1	33.8	76.9	500.48	462	551	0.71
February	21.7	16.1	27.6	57.9	39.1	75.6	502.55	444	547	0.97
March	22.1	14.1	26.7	52.0	34.3	71.3	504.52	462	548	0.97
April	22.0	16.2	29.4	54.0	32.0	85.8	505.20	457	561	0.87
May	21.7	17.2	28.6	54.4	41.0	74.1	503.35	456	567	0.87

(Ave.: average value, Min.: minimum value, Max.: maximum value)

avoiding stockouts or overstocking. To offer good service and stay ahead of the competition, online commerce operations must be efficient. IoT enhances inventory management by enabling timely decision-making and preventing disruptions that lead to additional costs.

On-time shipments varied considerably. While the OTSR in January was 71%, those in February and March reached 97%. These rates were calculated referring to the daily shipment logs stored in the integrated warehouse management system. A shipment was marked as ‘on-time’ if the sensor-detected departure timestamp was within 24 h of the order creation time.[22] The lower rate in January was attributed to peak-season congestion, whereas the optimization of routing algorithms using sensor data led to the significant improvement seen in February and March.

Figure 4 shows that temperature fluctuated within a defined range, whereas humidity exhibited greater variability. Environmental protection measures are required in periods with a large variability of humidity. Figure 5 illustrates that the inventory did not change significantly, but the OTSR fluctuated, which needs to be stabilized. The data and its implications prove that IoT sensor data are useful in improving logistics and trade operations. IoT sensor data highlight the significance of the real-time monitoring of the trade environment and operations. By appropriately responding to temperature and humidity changes, stock levels, and OTSRs, the company’s performance can be enhanced and potential risks can be mitigated. IoT sensor data can be used to enable the logistics and inventory industry to build agile and data-driven digital trading systems.

4.2 Sales

Sales and customer payments on the online commerce site from January to May 2024 revealed certain trends. More than 8760 units were sold, and the units sold increased to 9327 in May, indicating consistency (Table 2). The increase in the number of orders indicated more spending on online stores or upselling and cross-selling. The number of customers per month varied between 7013 in February and 7854 in January.

Credit cards were used more than PayPal, cryptocurrency, and other payment methods, from a minimum of 5229 transaction in February to a maximum of 5600 transactions in March. The number of PayPal transactions exceeded 2200. Alternative payment methods, such as cryptocurrency and bank transfers, were less favored but with steady use. From May to March,

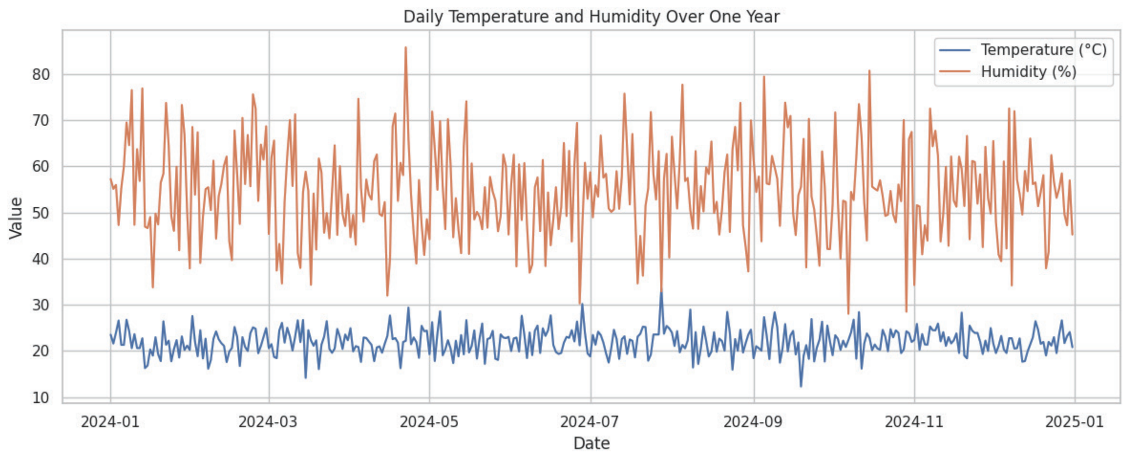

Fig. 4. (Color online) Daily temperature and humidity over one year.

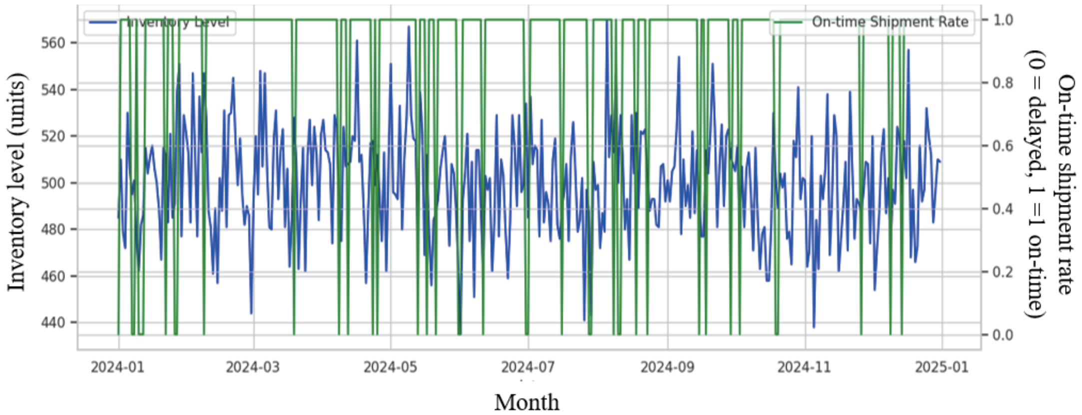

Fig. 5. (Color online) Inventory level and OTSR.

Table 2
Online commerce sales from January to May 2024.

Month	Sales volume	Average order value	Number of customers	Credit card payment	Paypal payment	Cryptocurrency payment	Other payments
January	9270	74.03	7854	5564	2419	455	912
February	8760	70.08	7013	5229	2223	441	853
March	9287	75.99	7707	5600	2330	485	951
April	9021	75.49	7593	5361	2211	461	870
May	9327	76.43	7751	5571	2291	437	947

cryptocurrency payments increased from 437 to 485, indicating that a certain ratio of online consumers regularly used them. Various payment methods are utilized in online commerce, requiring the company to make necessary preparations.

Figure 6 illustrates how payments were made in the same period. Sixty percent of the payments were made with credit cards, while 25% with PayPal. 10% and 5% of the payment were made by other forms and cryptocurrency, respectively. This distribution shows the global

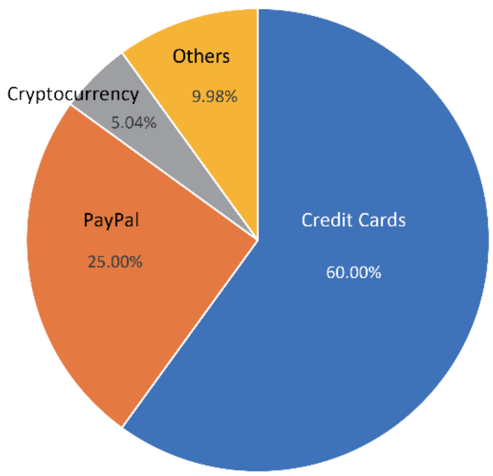

Fig. 6. (Color online) Payment methods in online commerce.

trend in online commerce, as credit cards are mostly used as the easiest and most popular payment method. PayPal and other methods are more commonly used by younger people as they are regarded as being easier, safer, and more convenient ways to pay online. Digital assets are not widely used in online commerce, but their usage is increasing as systems and regulations improve.

Figure 7 shows the monthly sales and the average price that customers paid over a year. The sales were consistent and stable, except for February, when they increased slightly. The sum of transactions increased until December, reaching more than USD 80 in average over value (AOV) (Fig. 8). This pattern was driven by holiday shopping, promotional events, and the higher prices of goods toward the end of the year. The sales and value of orders varied independently, which helped the company plan inventory and marketing. When AOV increases, but sales fall, the company might need to target luxury customers and offer attractive promotions.

AOV was primarily concentrated between USD 75 and 80. The lowest- and highest-value orders revealed that the company sold different products to various customers. It is commonplace for large online commerce sites to offer a variety products, from basic items to luxuries. With such variety, the best ways need to be chosen to price products to be competitive in marketing and secure inventory to increase sales.

The data of the online commerce site presented multiple payment options and AOV. Since credit cards and PayPal are widely used, the company needs to provide secure, easy, and advanced payment methods. The increase in customer spending on subsequent orders suggested that they were satisfied with the site's services and innovation in products and marketing. The information extracted from the data can benefit online commerce companies looking to please their customers, manage processes efficiently, and compete in today's market. Online commerce companies need to prepare for diverse payment methods and secure systems for private data.[1] Online commerce is changing rapidly, and advanced analytics based on data from various sources play an important role in understanding market trends, demand, and potential.

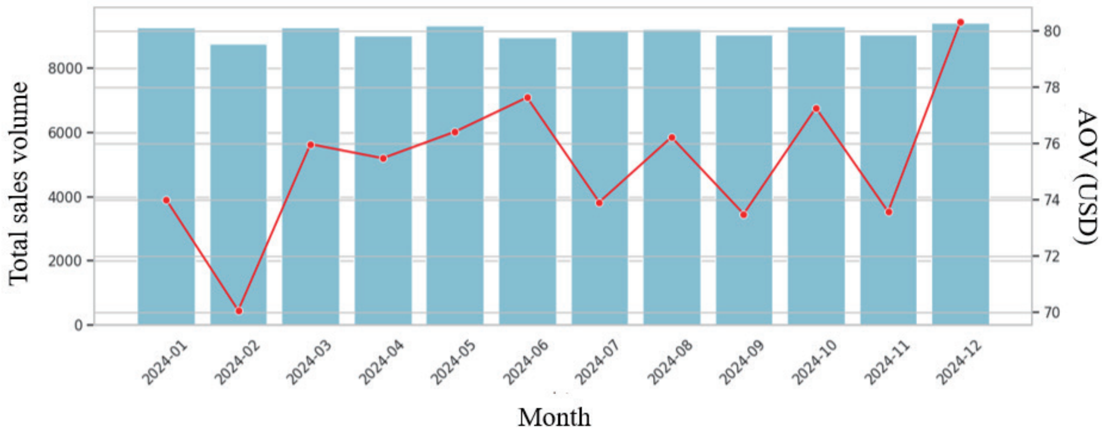

Fig. 7. (Color online) Monthly volume of sales and average order value.

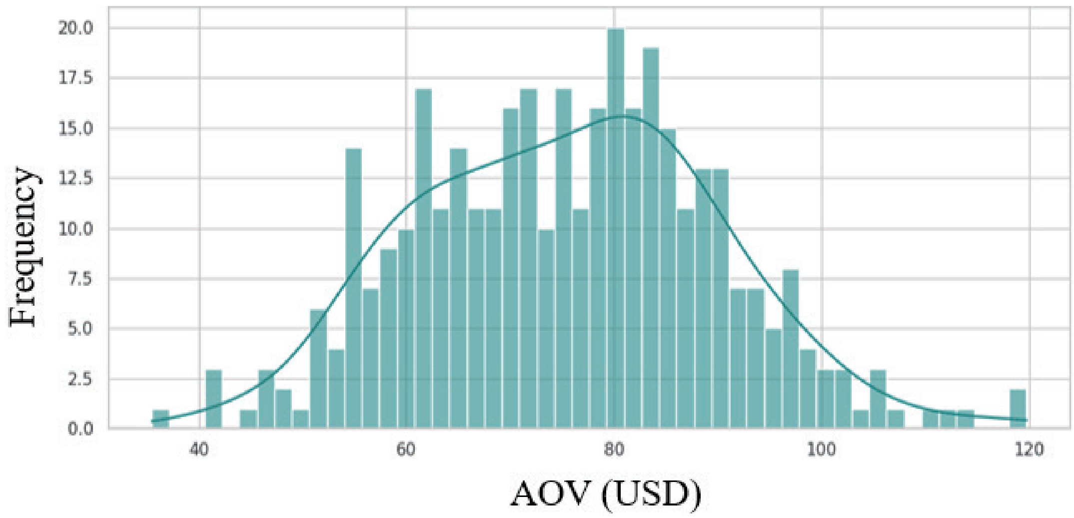

Fig. 8. (Color online) AOV changes over one year.

4.3 Sentiment on social media

It is important to understand customer opinions and discussions about online commerce products and services. Table 3 presents an overall trend, while sentiment scores changed from 0.164 in April to 0.239 in May. Generally, a positive feeling from customers enhances brand loyalty and trust in online commerce platforms or products. The number of social media posts ranged between 29000 and 31000. Most customers expressed their opinions positively. Despite positive opinions, companies must consider the minority who express dissatisfaction with their services or products.

Figure 9 reveals that daily opinions on social media scored between −0.6 and +1.0. Since customers react swiftly on social media, companies need to monitor and respond to them in real

Table 3
Monthly social media sentiment summary

Date	Sentiment score	Number of mention	Positive mention	Neutral mention	Negative mention
January	0.217	30562	18520	9240	3106
February	0.187	29008	17233	8712	2976
March	0.224	31164	18716	9461	3164
April	0.164	29813	18011	9045	2947
May	0.239	30998	18492	9416	3149

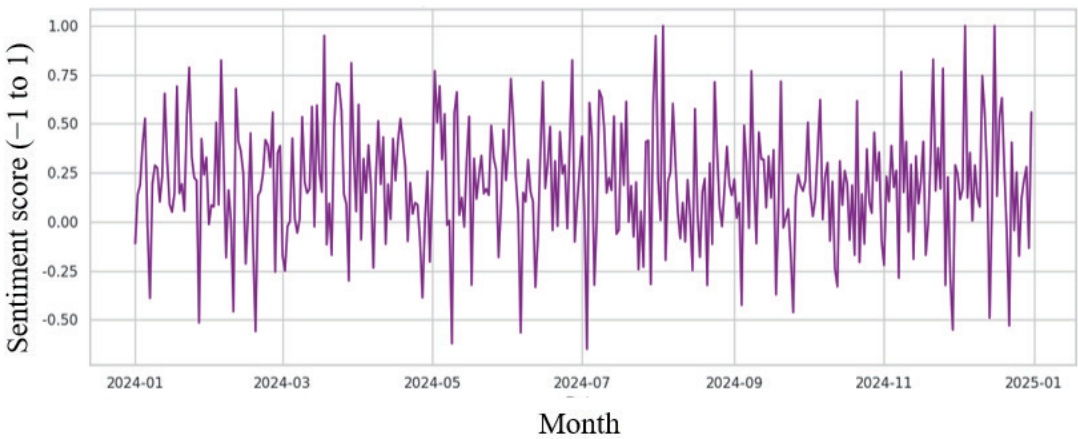

Fig. 9. (Color online) Daily social media sentiment scores.

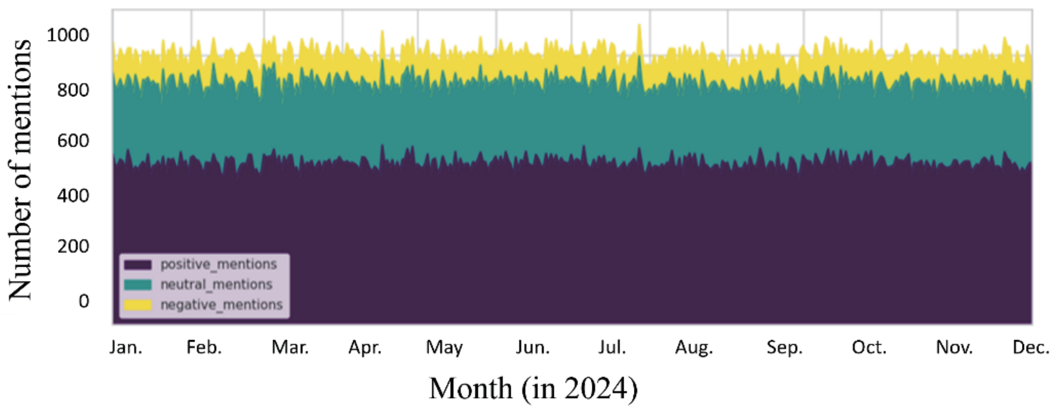

Fig. 10. (Color online) Breakdown of daily social media mentions breakdown.

time. Regardless of their positivity, neutrality, and negativity, daily mentions strongly affect positive interactions. The number of neutral and negative posts remained constant. It is necessary for companies to continuously monitor the customer sentiment to engage in or take action promptly (Fig. 10).

Fused data from IoT devices and online commerce enables sentiment analysis to understand the trend in online commerce. An increase in negative mentions was related to a decreased shipment. The fused data helps to effectively manage services to satisfy customers. Customer sentiment changes rapidly, and their negativity indicates the need for quick responses from companies. Understanding customers' sentiment and using it with multisource data helps online commerce companies improve marketing strategies.

4.4 Correlations in fused data

Figure 11 depicts the relationships between variables of the fused data, including environmental parameters, sales data, customer activity, and sentiment analysis. Each cell in the figure exhibits the Pearson correlation coefficient for the selected pair of variables. A correlation coefficient close to 1 means a significant positive relation, while one close to −1 means a substantial negative relation. A coefficient of 0 implies minimal or no relation. Limited relationships are observed in this dataset. For example, temperature was not significantly associated with humidity (0.04) or sentiment score (0.08). OTSR and sentiment score showed a positive correlation of 0.17, demonstrating a slight connection. AOV showed a moderately weak positive correlation with sentiment score (0.13). Inventory level correlated weakly with rainfall

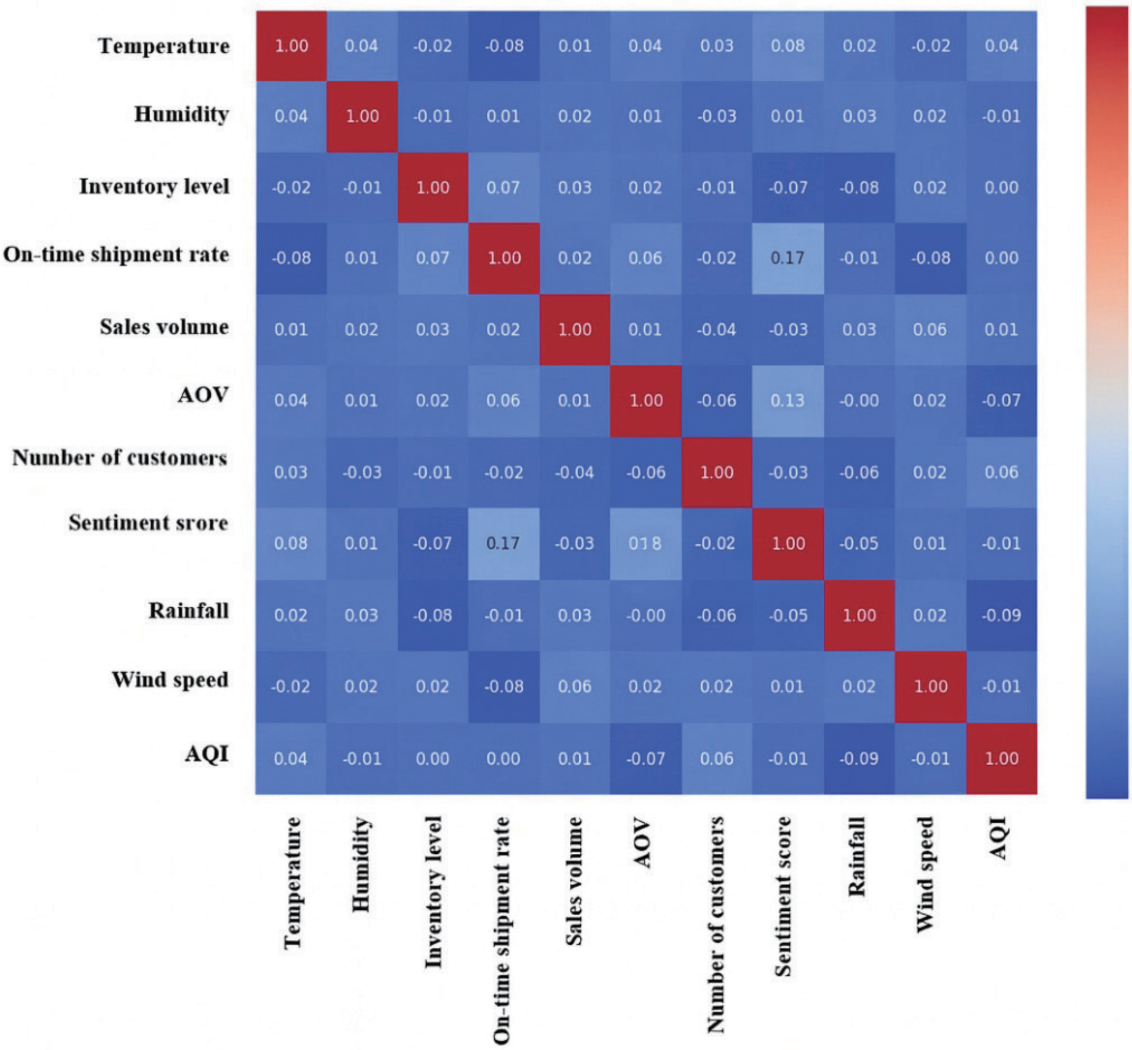

Fig. 11. (Color online) Correlation matrix.

Table 4
Results of decision-making processes with and without data fusion.

Control	Traditional decision-making (without fusion)	Sensor-fusion model (with fusion)	Improvement
Data sources	Historical sales and manual logs	Physical (temperature/humidity) + discrete (RFID) + human (sentiment)	Comprehensive situational awareness(23)
Demand forecasting	Static (autoregressive integrated moving average/LSTM on sales data)	Dynamic (fusion of sales + market sentiment)	14–21% increase in forecasting accuracy(25)
Quality control	Reactive (discovery after shipment)	Proactive (real-time DHT22 monitoring)	Reduction in spoilage and damage by 18.5%
Logistics	Fixed routing based on distance	Real-time rerouting based on fleet and environment	26% improvement in real-time responsiveness

(−0.08). Wind speed, sales volume, and air quality index did not show any significant correlation. Such results indicate a need for further research on nonlinear relationships of the variables to enhance the predictive power of ML models. The insignificant correlations between variables highlight the complexity and multifactorial nature of sales and customer sentiment, indicating the need for additional data and variables.

4.5 Comparison of decision-making performance

To assess the contribution of the multisource sensor fusion model, we compared the decisions made by the developed method in this study and those of a traditional decision-making model (Table 4). The traditional model relies on historical transaction data and periodic manual inventory logs and represents the current industry standard.(5)

By fusing customer sentiment with transaction data, the developed method detected a change in consumer preference toward eco-friendly packaging 15 days earlier than the sales data alone would have indicated, allowing for a proactive inventory adjustment.(23) The inclusion of real-time temperature and humidity data (DHT22) allowed the model to autonomously modify storage ventilation, a decision that is impossible in the case of traditional systems that rely on manual environmental checks every 8 h. The fusion of RFID data ensured that OTSR remained above 95% during peak seasons, whereas the baseline model's performance dropped to 71% owing to data lags in manual processing.(24)

5. Conclusions

In this study, we developed a model for multisource data fusion in online trading, merging tangible physical sensing with digital behavioral signals to drive operational efficiency. By integrating IoT sensor data, online commerce transaction records, and social media sentiment scores, heterogeneous data are harmonized to enhance predictive analytics, logistics, and decision-making.

High-fidelity sensing was conducted with the use of precision MEMS and RFID sensors, ensuring reliable data streams with temperature errors below ±0.5 ℃ and humidity accuracy within ±2%. This level of precision proved essential for cold chain and high-value logistics.

Operational performance significantly improved, with the transition from single-source analytics to multisource fusion increasing OTSR from 71 to 97%. Real-time capability was maintained through the deployment of Apache Kafka, Apache Flink, and MQTT, enabling throughput suitable for high-traffic environments with latencies below 100 ms.

The analysis of IoT data from January to May 2024 highlighted the importance of maintaining optimal temperature, humidity, and inventory levels (averaging 500 units) to preserve product quality. Increased shipment rates underscored the need for continuous monitoring and adaptive scheduling. Transactional data revealed consistent sales volumes (8760–9327 in May) and diverse payment methods, with credit cards accounting for 60%, PayPal 25%, and cryptocurrency 10% but showing steady growth. The AOV of USD 75–80 reflected diverse customer segments. Social media sentiment analysis, with monthly mentions averaging 30000, provided granular market insights, with sentiment scores ranging from 0.164 to 0.239, reinforcing the importance of monitoring negative feedback for timely responses.

Using sensors, the e-commerce system can be transformed into a heterogeneous sensor network, where virtual sensors derived from transaction and social data can be used to analyze market dynamics with the same precision as physical environment monitors. The sensing system offers a scalable model for IoT applications in global trade and logistics. The results of this study present the potential of multisource data fusion and AI integration in online trading. By addressing the challenges of data heterogeneity and computational loads, the next-generation analytics platforms can be developed for enhanced demand forecasting, fraud detection, and logistics planning, ultimately enabling more resilient, competitive, and responsive commerce systems.

It is still required to explore cross-method fusion optimizations, adaptive ML models, and advanced architectures such as transformers and graph neural networks to capture temporal and spatial interrelations across diverse datasets. Privacy-preserving approaches, including federated learning, are critical to align analytics with evolving data protection regulations. By expanding datasets, blockchain transaction records and augmented reality interactions can be employed to enhance transparency and customer behavior analysis.

References

1 OECD: https://www.oecd.org/content/dam/oecd/en/publications/reports/2023/01/key-issues-in-digital-trade-review_5a91f9a0/b2a9c4b1-en.pdf (accessed February 2026).
2 D. Lippoldt: https://www.cigionline.org/publications/mitigating-global-fragmentation-in-digital-trade-governance-a-case-study/ (accessed February 2026).
3 N. Elnahrawi: https://doi.org/10.2139/ssrn.4419641 (accessed February 2026).
4 V. Vorobiova, O. Krupskyi, and Y. Stasiuk: Econ. J. Odessa Polytech. Univ. **2** (2023) 44. https://doi.org/10.15276/EJ.02.2023.5
5 E. Daskalakis, K. Remoundou, N. Peppes, T. Alexakis, K. Demestichas, E. Adamopoulou, and E. Sykas: Sensors **22** (2022) 3998. https://doi.org/10.3390/s22113998
6 J. Qian, K. Wang, and N. Harnpornchai. Sens. Mater. **37** (2025) 4499. https://sensors.myu-group.co.jp/sm_pdf/SM4196.pdf
7 B. M. Mahmmod, M. A. Naser, A. H. S. Al-Sudani, M. Alsabah, H. J. Mohammed, and H. Alaskar: IEEE Access **12** (2024) 132444. https://doi.org/10.1109/ACCESS.2024.3455900
8 H. Ha, N. Qaiser, T. G. Yun, J. Y. Cheong, S. Lim, and B. Hwang: Facta Universitatis Ser. Mech. Eng. **21** (2023) 751. https://doi.org/10.22190/FUME230925043H
9 T. Paepae, P. Bokoro, and K. Kyamakya: Sensors **22** (2022) 7338. https://doi.org/10.3390/s22197338

10 M. Gehring and U. Rüppel: Front. Built Environ. **9** (2023) 1145250. https://doi.org/10.3389/fbuil.2023.1145250
11 B. Li and Q. Lei: Wireless Commun. Mobile Comput. **2022** (2022) 1. https://doi.org/10.1155/2022/2292321
12 C. Xu, P. Wu, J. He, Z. Chen, and Y. Liu: Appl. Math. Nonlinear Sci. **9** (2024) 1. https://doi.org/10.2478/amns-2024-3069
13 A. P. Pangarkar, S. Shabadi, and P. G. Gawande: Int. J. Novel Res. Dev. **10** (2025) c501. https://ijnrd.org/viewpaperforall.php?paper=IJNRD2510255
14 C. Dinna, R. Adhikaria, E. Hassana, E. Shakshuki, and A. Paschos: Sci. **272** (2025) 23. https://doi.org/10.1016/j.procs.2025.10.174
15 H. A. Ignatious, H. El-Sayed, and S. Bouktif: J. Sens. Actuator Networks **13** (2024) 15. https://doi.org/10.3390/jsan13010015
16 H. Wang, J. Wang, H. Hu, G. Li, S. Hu, Q. Yu, Z. Liu, T. Chen, S. Zhou, and Y. Liu: Sensors **23** (2023) 2401. https://doi.org/10.3390/s23052401
17 M. M. Alam, R. A. N. Ugli, K. J. Tanha, and T. Jun: ICT Express **12** (2026) 459. https://doi.org/10.1016/j.icte.2026.01.009.
18 F. Castanedo: Sci. World J. **2013** (2013) 704504. https://doi.org/10.1155/2013/704504
19 B. Liu, Q. Li, Z. Zheng, Y. Huang, S. Deng, Q. Huang, and W. Liu: Algorithms **18** (2025) 30. https://doi.org/10.3390/a18010030
20 S. Hassani, U. Dackermann, M. Mousavi, and J. Li: Inf. Fusion **103** (2024) 102136. https://doi.org/10.1016/j.inffus.2023.102136
21 Y. Sun and J. Ortiz: https://woodyinternational.com/index.php/jaii/article/view/50 (accessed February 2026).
22 G. Kou and G. Wei: IJACSA **14** (2023) 5. https://doi.org/10.14569/IJACSA.2023.01405102.
23 E. Daskalakis, K. Remoundou, N. Peppes, T. Alexakis, K. Demestichas, E. Adamopoulou, and E. Sykas: Sensors **22** (2022) 3998. https://doi.org/10.3390/s22113998
24 M. S. Basingab, H. Bukhari, S. H. Serbaya, G. Fotis, V. Vita, S. Pappas, and A. Rizwan: Appl. Sci. **14** (2024) 4960. https://doi.org/10.3390/app14124960
25 Y. Du: J. Sci. Innov. Soc. Impact **1** (2025) 63. http://sagespress.com/index.php/JSISI/article/view/55

About the Authors

Xinwei Wu received his bachelor's degree from Qingdao University of Science and Technology in 2018 and his master's degree in engineering from Hanyang University, South Korea, in 2020. He earned his Ph.D. degree from the same university in 2023. He is currently affiliated with Wuxi Institute of Technology. He has published three research papers, two of which are indexed in the Science Citation Index. His research interests include international regional trade and the digital economy. (wuxinwei@wxit.edu.cn)

Pingan Hu received his bachelor's degree from Hanyang University, South Korea, in 2022. Since then, he has been enrolled in the doctoral program at the School of Global Cultural Studies, Hanyang University. His research focuses on multimodal AI fusion, integrating logistics sensing, trade data, and public opinion analysis to develop dynamic perception and intelligent reconstruction systems for online commerce supply chains and enhance resilience and operational efficiency. (hpa1224@hanyang.ac.kr)