IoT and Multi-Source Data-Driven Intelligent Warehouse Optimization Using KMA-RDTC

Jiping Liu[1, *], Chenxiao Liu [2]
[1.] School of Economics and Management, Jiaozuo University, Jiaozuo, Henan，454000, China
[2.] Shanghai Shuju Information Technology Co., Ltd., Shanghai, 201100, China
E-mail: liujiping-ljp@outlook.com
[*]Corresponding author

Keywords: intelligent warehousing, IoT integration, multi-source data, inventory optimization, operational efficiency, supply chain automation

Received: January 5, 2026

Warehouse operations have developed over the decades through the enhanced Internet of Things (IoT) technology. The traditional operation is perceived as inconsistent and operationally inefficient, which are the main deficiencies in traditional inventory processes. The research fills the gaps by introducing a smart warehouse optimization tool on IoT and multi-source data to optimize and refine inventory accuracy, reduce operational expenses, and optimize the use of resources. Data from 2500 real-time warehouse operations were pre-processed using min-max normalization, and dimensionality reduction was applied via Principal Component Analysis (PCA). The Komodo Mlipir Algorithm-tuned Random Decision Tree Classifier (KMA-RDTC) was proposed to integrate the RDTC with the KMA for optimizing warehouse operations. The RDTC provides robust classification and anomaly detection for inventory data and demand forecasting, while the KMA optimizes RDTC hyperparameters using exploration-exploitation strategies, enabling efficient real-time anomaly detection and forecasting, and enhancing predictive performance, operational accuracy, resource utilization, and decision-making in intelligent warehouse environments. Through experimental testing implemented in Python, it demonstrated superior predictive reliability over traditional Machine Learning (ML) methods by providing accuracy (0.985), prediction accuracy (0.958), sensitivity (0.934), specificity (0.962), Root Mean Squared Error (RMSE) (0.10), and Mean Absolute Error (MAE) (0.07) with Python implementation. These results indicate accurate anomaly detection, efficient resource utilization, and high operational precision. The framework provides a comprehensive design for intelligent warehouse operations, enabling scalable, data-driven, and sustainable logistics aligned with Industry 4.0 principles.

Povzetek: Raziskava predstavi pametno optimizacijo skladiščnih procesov z uporabo IoT, več virov podatkov in modela KMA-RDTC, ki izboljša natančnost zalog, zaznavanje anomalij, napovedovanje povpraševanja ter učinkovito rabo virov v skladu z načeli Industrije 4.0.

1 Introduction

Warehouse effectiveness and efficacy are greatly impacted by the network's structure and the costs associated with receiving, storing, capturing, and exporting. Collection of orders is the most expensive and laborious warehousing operation. This kind of operation necessitates automatic coordination and the evaluation of demands for task execution. The creation of smart devices that allow for a range of services and smart systems has received a lot of attention due to the substantial market for IoT integration [1]. Automated and intelligent warehouse systems are constantly created and utilized by the warehouse industry. The energy utilization of computerized warehouse systems grows with increased warehouse activities [2]. Manufacturing enterprises that operate in competitive marketplaces focus on minimizing expenses and enhancing efficiency. Although expenses are fundamentally associated with profitable activities, including manufacturing and quality control, which are considered essential [3]. Infrastructure, modes of transport, shipment storage spaces, and other multiple sources of information constitute every component of the supply chain during all phases of the purchase, manufacturing, and distribution operation, while the warehouse design is a crucial component of the organization [4]. As the shipping and internet shopping sectors have developed, the benefits of intelligent warehouses are beginning to emerge. The multi-automation system is the fundamental implementation of the intelligent warehouse, which has the potential to enhance productivity and lower expenses. Intelligent warehouses emerged to improve logistics effectiveness, which has grown as one of the most prevalent subjects in technological advances [5-6]. The traditional operation of a warehouse has an imprecise distinction due to social and technological advancements, which results in inadequate storage efficiency. Modern technologies have become increasingly essential to the warehouse system that improves the effectiveness of product transportation, processing, storage, and distribution, consequently resolving the concern of inefficient storage systems. Conventional detection devices lack efficient path planning techniques that are available as a technological assistance due to inadequate processing power and over consumption of resource concerns [7-8]. The development and evolution of the logistics sector is possible through intelligent logistics. The logistics sector is labour-intensive under the conventional manual operation mode, and a manpower shortage has become a prevalent issue in the industry [9]. Storage constraints, inventory management difficulties, ineffective warehouse organization and design, ineffective selection and shipment procedures, a deficiency in automation and technology, inadequate collaboration and interaction, variation in the seasons and fluctuations in demand, along with safety and compliance issues, are several instances of warehouse operation issues [10]. Due to inconsistent decision-making processes and inadequate real-time data integration, the traditional warehouse operation system has complexities with managing demand fluctuations, dynamic inventory circumstances, and operational inefficiencies. An intelligent model that manages data from several sources, forecasts inventory behaviour, and optimizes warehouse operations is required with the development of IoT and various data sources. The research questions are as follows:

RQ1: How can multi-source IoT data be effectively utilized for real-time anomaly detection and demand forecasting in intelligent warehouse operations?

RQ2: Does the integration of the Komodo Mlipir Algorithm with a Random Decision Tree Classifier improve classification performance and reduce prediction errors compared to conventional machine learning models?

RQ3: How does the proposed KMA-RDTC framework enhance operational efficiency, resource utilization, and decision-making in dynamic and large-scale warehouse environments?

1.1 Objective and contributions of this research

The research introduces an intelligent warehouse optimization framework that leverages multi-source IoT data to enhance operational efficiency, improve inventory precision, reduce fulfilment latency, and optimize resource allocation in real time. It proposes the Komodo Mlipir Algorithm-tuned Random Decision Tree Classifier (KMA-RDTC), which integrates robust classification and anomaly detection with hyperparameter optimization to dynamically adapt to changing warehouse conditions. The key contribution of the research is listed as follows.

Proposes a KMA-RDTC-based intelligent warehouse optimization model that enhances inventory accuracy, minimizes operational costs, and improves resource utilization.

Collects multi-source IoT warehouse data and applies preprocessing and feature extraction, including min-max normalization for data quality and PCA for dimensionality reduction while preserving critical variations for predictive modeling.

Integrates RDTC to perform robust classification and anomaly detection, and KMA to fine-tune its hyperparameters using exploration-exploitation strategies.

The research proposed the KMA-RDTC method to efficiently forecast and determine the intelligent warehouse optimization.

Performs better than conventional methods, demonstrating superior predictive accuracy, anomaly detection, and operational efficiency, while supporting Industry 4.0 logistics through improved decision-making, dynamic routing, resource use, and real-time automation.

2 Relevant articles

Automated warehousing and autonomous sorting by utilizing the Convolution Neural Network (CNN) to investigate the use of robotics in intelligent supply-chain and digital logistics, along with achieving an effective operation, energy efficiency, and a decrease in emissions in the field of warehousing and organization, were investigated [11]. Results established the development of a smart supply-chain system by implementing intelligent products in the storage industry. The research's application scope was limited. The consequence of monitoring data deeper than the basic traceability purpose has been examined in the research [12]. Training models for prediction performance that assisted with the design of a warehouse system were determined. The major characteristics of the input data in forecasting performance were determined by the results of the research. The ideal decision was not indicated by the prediction models.

The potential incorporation of computer simulation and reinforcement learning (RL) to develop effective mechanisms that facilitated quick gathering of data from a warehouse in a complicated circumstance was examined [13]. It employed the commercial simulator to test the possibilities. The research limitations included difficulty in generalizing and a lack of significance in practical applications. By enabling the creation of Digital Twin (DT) that integrated IoT technologies, data modeling, and Artificial Intelligence (AI) features, the investigation [14] incorporated the architecture and requirements of the customer. The suggested technique was represented as an effective method for incorporating technological innovations into complicated warehouse environments by providing improved productivity and security.

The repeated assessment was represented as one of the investigation's limitations. The research [15] investigated the IoT-based distribution network design with an integrated forward/reverse distribution system designed to be flexible, resilient, multi-product, and across multiple periods. With an adaptability of 213.528%, the suggested model allows for developing a supply chain network effectively. The intended purpose of the investigation [16] was to determine the real-time information integration processes in IoT-based warehousing necessities. The outcome of the experiments demonstrated that data integration has significance in continuous processing and has a range of experience with IoT-based warehouses in different regions.

An intelligent transportation and logistics warehouse's ideal positioning technique has been suggested in the research [17]. To enable input from IoT devices, it integrated various kinds of complicated processes. The results of the ideal positioning method have superior positioning accuracy and fewer computations compared to traditional techniques. Implementing and assessing a deep learning (DL) technology to improve warehouse inventory management was the focus of the investigation [18]. The CNN was used by the platform to precisely identify and categorize goods in real time, along with highlighting significant advancements, like a 9% rise in inventory accuracy. The reliability of product detection was impacted by variations in the standard of the images obtained.

The research [19] suggested integrating CNNs with Bidirectional LSTM (BiLSTM) models to increase the sustainability and efficiency of supply chains. The suggested hybrid model's specificity was 94.65%. CNNs and BiLSTMs require huge amounts of computing power to train and implement, and have limited computational resources. An investigation connected to Extended Warehouse Management (EWM) that used a variety of ML methods was provided [20]. To determine the optimum metrics, several categorization methods were examined and evaluated in the results. It lacked a thorough investigation to calculate the precise financial, social, and environmental benefits.

The uses of IoT and AI in modern warehouses and transportation were thoroughly examined in the research [21]. The outcomes enabled the shipping and warehousing sector to be more effective, wise, and sustainable. A decreased potential for growth in the warehousing and logistics industries was one of the drawbacks presented in the research. The research [22] developed a system that employed the MongoDataBase (MangoDB) to apply the IoT method in warehousing management. The findings of the research provided the corporations with a practical road map for utilizing IoT to enhance the facilities. The research issues comprised reduced efficiency and higher costs for operation. The purpose of the research [23] was to find the way IoT impacted warehouse management in various enterprises of different sizes while determining whether the benefits and drawbacks of IoT varied under the equivalent conditions. Consequently, experts evaluated the design concept for IoT implementation in warehouse enterprises. Timing constraints for data collection were presented in the research. To optimize customer satisfaction and improve warehouse efficiency, the research [24] presented an Intelligent Storage Location Assignment (ISLA) method that utilized complicated time series clustering algorithms. An efficiency of 61% and 69% was achieved by the ISLA technique. Although the research lacks consistent physical reassignments to adjust to the developments in warehouse operations, it was observed.

Enabling the real-time, resource-level carbon tracking and scenario analysis by presenting a bottom-up carbon measurement methodology that is integrated into a warehouse Digital Twin was provided in the research [25]. Results demonstrated that the warehouse DTs extracted the carbon more effectively. The investigation failed to be responsive in making operational decisions. Creating an accurate target function by utilizing AutoML-based regression models to find patterns in the warehouse data was the intent of the investigation [26]. A simulation dataset was used to evaluate ML models and discovered the most effective results in identifying the warehouse data pattern. A limited number of parameters were focused on in the research. The purpose of the research [27] was to complete warehouse inventory operations using an Unmanned Aerial Vehicle (UAV) to identify the goods. Two RL methods were used. The efficacy of the suggested approach has been established by testing conducted in both simulation and actual environments. The ability of a UAV to operate independently in every dimension was limited. The effectiveness of using the UWB method for interior localization, highlighting a vital part in attaining phenomenal precision of distance determination, especially designed for warehouse management, was highlighted in the research [28]. Localization by Optimization (LBO) and ML Enhanced Trilateration (MLET) were shown to be more accurate and robust than the conventional approach during evaluation. The system's overall performance and usefulness were directly impacted by the limited battery lifespan.

To maximize demand forecasting, reduce inventory discrepancies, and save operating expenses, the investigation [29] created a detailed structure involving the utilization of Extreme Gradient Boosting (XGBoost), RF, and artificial data generation. Sustainability measurements showed that the route optimization reduced the fuel usage by 12%, while ML-optimized packaging reduced material waste by 20%. The investigation constrained the model's ability to manage difficult environmental circumstances. The research [30] introduced a smart Indoor Positioning System (IPS) that combines Ultra-Wideband (UWB) detectors with Long Short-Term Memory (LSTM) with Kalman filtering for real-time object tracking, employing a personalized data integration process and optimizing parameters. Localization accuracy can be enhanced by 4% with the Kalman-LSTM model when compared with baseline techniques. It was not suitable for larger or multilevel operations.

Worker safety and operational efficiency in smart warehouses were enhanced by integrating IoT sensors monitoring physiological and motion data with computer vision tracking posture and behavior [31]. ML generated predictive insights for fatigue and injury risk, enabling proactive management. Benefits included improved human-robot collaboration and warehouse performance, while effectiveness depended on sensor placement and data quality.

Optimizing data warehouse performance under dynamic workloads focused on enhancing query speed, resource allocation, and system efficiency using ML techniques [32]. Predictive analytics guided resource allocation, anomaly detection identified performance bottlenecks, and supervised/unsupervised models automated indexing, partitioning, and caching. Simulations demonstrated improved speed, scalability, and cost efficiency, while performance remained sensitive to workload diversity, model accuracy, and adaptation to unforeseen data patterns.

The role of DT Technology (DTT) and IoT-integrated AI in warehouse management was explored for predictive inventory forecasting and real-time decision-making [33]. IoT sensors and WMS combined with ML monitor inventory, product movements, and environmental factors, enabling optimized resource allocation, cost reduction, and improved operational efficiency, while performance may vary with highly dynamic inventory or limited sensor coverage. Table 1 summarizes review of recent approaches in smart warehouse systems, detailing aims, methods, key findings, and identified challenges.

Table 1: Summary of literature on AI- and IoT-driven warehouse optimization

Ref.	Aim	Methods	Results	Limitations
[11]	Investigate robotics in intelligent supply chain and digital logistics	CNN-based automated warehousing and autonomous sorting	Development of a smart supply-chain system; improved operation, energy efficiency, reduced emissions	Limited application scope
[12]	Analyze monitoring data beyond basic traceability	Training prediction models for warehouse design	Identification of major input features affecting forecasting	Prediction models did not indicate ideal decisions
[13]	Develop mechanisms for rapid warehouse data gathering	Computer simulation & Reinforcement Learning	Evaluated possibilities using commercial simulator	Difficult to generalize; limited practical significance
[14]	Integrate IoT, AI, and data modeling in warehouse DTs	Customer requirement-based architecture implementation	Improved productivity and security	Repeated assessment noted as limitation
[15]	Design flexible, resilient IoT-based distribution network	Forward/reverse distribution system modeling	Adaptability of 213.528%; effective supply chain network development	Broader approach for disturbances not addressed
[16]	Determine real-time information integration in IoT warehouses	Experimental evaluation of data integration processes	Data integration significant for continuous processing	Cost-effective IoT logistics not thoroughly investigated
[17]	Suggest ideal warehouse positioning for intelligent transportation	IoT-integrated positioning method	Superior positioning accuracy; fewer computations	High storage costs; scalability management issues
[18]	Improve warehouse inventory management	Deep Learning with CNN for real-time identification	9% increase in inventory accuracy	Detection affected by image quality variations
[19]	Increase supply chain sustainability & efficiency	Hybrid CNN + BiLSTM model	Specificity 94.65%	High computing power required; limited resources
[20]	Optimize Extended Warehouse Management (EWM)	Various ML classification methods	Determined optimum metrics	Lacked financial, social, environmental impact evaluation
[21]	Examine IoT and AI usage in modern warehouses	Literature/experimental analysis	Enhanced efficiency, intelligence, sustainability	Limited growth potential in warehousing/logistics
[22]	Apply IoT in warehouse management	MongoDB-based IoT system	Practical roadmap for facilities enhancement	Reduced operational efficiency; higher costs
[23]	Investigate IoT impact on warehouses of different sizes	Expert evaluation	IoT benefits/drawbacks assessed	Timing constraints for data collection
[24]	Optimize storage location assignment	ISLA method with time series clustering	Efficiency: 61%–69%	Lacked consistent physical reassignments
[25]	Enable real-time carbon tracking in warehouses	Warehouse DT with carbon measurement	Effective carbon extraction	Not responsive for operational decisions
[26]	Discover patterns in warehouse data	AutoML-based regression models	Effective pattern identification	Focused on limited parameters
[27]	Complete inventory using UAVs	UAV with RL methods	Tested in simulation & real environments	Limited UAV autonomy in all dimensions
[28]	Achieve precise indoor localization	UWB with LBO & MLET	More accurate & robust than conventional	Performance affected by battery life
[29]	Maximize demand forecasting & reduce inventory discrepancies	XGBoost, RF, artificial data generation	Fuel reduced by 12%; material waste reduced by 20%	Limited handling of difficult environmental conditions
[30]	Real-time indoor object tracking	IPS: UWB + LSTM + Kalman filtering	4% improved localization accuracy	Not suitable for larger/multilevel operations

[31]	Enhance worker safety & operational efficiency	IoT sensors + computer vision + ML	Improved human-robot collaboration & warehouse performance	Effectiveness depends on sensor placement & data quality
[32]	Optimize data warehouse performance	ML-based predictive analytics, anomaly detection, automated indexing	Improved speed, scalability, cost efficiency	Sensitive to workload diversity, model accuracy
[33]	Predictive inventory forecasting & real-time decision-making	DT + IoT-integrated AI + WMS	Optimized resource allocation, cost reduction, operational efficiency	Performance varies with dynamic inventory; limited sensor coverage

2.1 Research gaps

Despite advancements in warehouse automation, IoT, and AI-driven inventory management, existing state-of-the-art methods face critical limitations. CNN, BiLSTM, and hybrid models achieve high accuracy but demand substantial computing power and cannot operate in real time. Most solutions cannot handle multi-source IoT data, limiting use in complex, dynamic warehouses. DT, UAV, and IoT frameworks improve monitoring, localization, and efficiency but lack adaptability to large-scale, multi-level operations and real-time decision-making. Cost-efficiency, carbon tracking, and consistent storage reassignment are often overlooked, while predictive models focus on limited parameters without delivering end-to-end optimization. These gaps highlight the need for a unified, scalable, and real-time intelligent warehouse framework integrating multi-source IoT data, AI prediction, and resource management to enhance performance, sustainability, and adaptability. To address such gaps, the research proposes the KMA-RDTC framework, which integrates multi-source IoT data to enable real-time anomaly detection and demand forecasting across complex warehouse operations. The RDTC provides robust classification, while the KMA efficiently optimizes hyperparameters, reducing computational overhead. This approach enhances operational accuracy, resource utilization, and adaptive decision-making, offering a scalable, cost-efficient, and end-to-end solution for intelligent warehouse optimization.

3 Research methodology

The research develops an intelligent warehouse optimization system based on IoT and multi-source data that enhances resource efficiency, reduces operating costs, and increases inventory precision. The warehouse IoT multi-source operations data are obtained and preprocessed by the min-max normalization method to provide data quality and consistency. Efficient features are extracted by the PCA method, which reduces the dimensionality of data while maintaining a significant variance for the predictive method. The proposed KMA-RDTC method's applications are explored more extensively. The flow of the research method is shown in Figure 1.

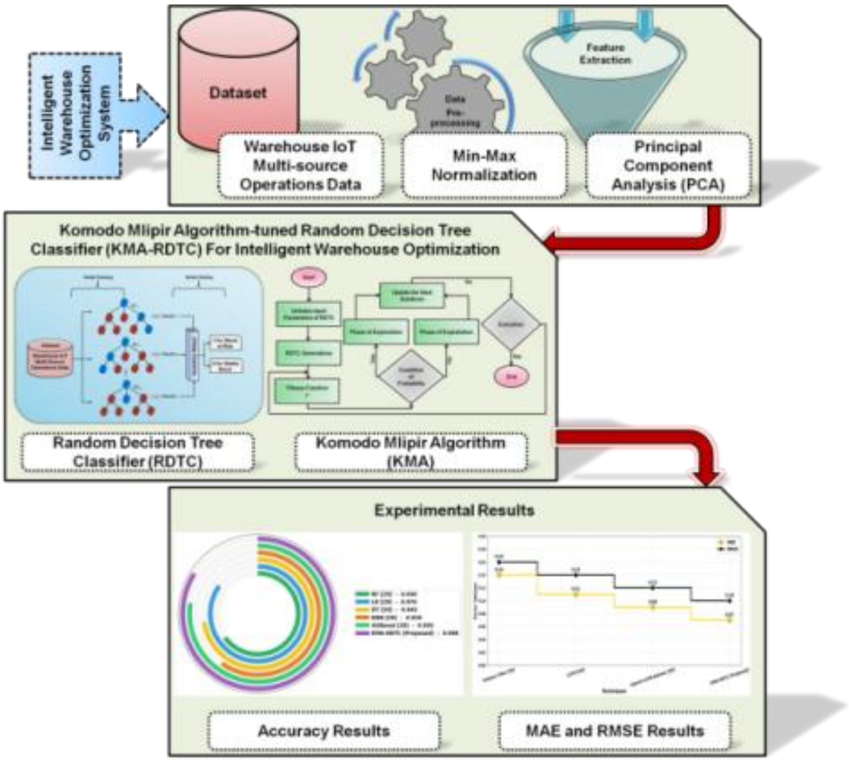

Figure 1: Flow of research method

3.1 Dataset

The warehouse IoT multi-source operations data is obtained from the open source Kaggle (https://www.kaggle.com/datasets/colabsss/warehouse-iot-multi-source-operations-data/data). This dataset includes 2500 real-time warehouse operations obtained from transactional activity logs, RFID-based item tracking, IoT-enabled smart shelves, and environmental sensing devices. It depicts the interactions between inventory levels, environmental factors, item movements, routing routes, and fulfillment operations in multi-location warehouse operations. The information facilitates the investigation of demand trends, order processing, inventory status, operational performance, and stock behaviour under dynamic situations across several warehouse regions. Factors in the dataset are shown in Table 2. The dataset was partitioned using an 80:20 ratio based on chronological order. After sorting the records according to timestamps, the earlier 80% of observations were used for model training, while the most recent 20% were reserved for testing.

Table 2: Factors in dataset

Factors	Descriptions
Timestamp	Records the exact date and time of the warehouse event.
Warehouse_ID	Identifies the warehouse location (e.g., W1, W2, W3).
Section_ID	Indicates the specific section or zone inside the warehouse.
Shelf_ID	Refers to the smart shelf where the product is stored.
Product_ID	Represents the unique code assigned to each product type.
RFID_Tag_ID	A unique identifier for tracking individual product units.
Current_Stock_Level	It shows the quantity of items available on the shelf at that moment.
Stock_Capacity	Displays the maximum number of items the shelf can hold.
Temperature_C	Measures the temperature in degrees Celsius at the corresponding warehouse area.
Humidity_%	Indicates the humidity level in the warehouse at the time of recording.
Order_ID	Links the row to a specific customer or internal order.
Order_Quantity	Represents the number of product units requested in the order.
Order_Fulfillment_Time_sec	Shows the duration taken to complete the order fulfillment process in seconds.
Picking_Route_ID	Identifies the path followed by workers or robots while picking items.
Agent_ID	Represents the worker or autonomous agent handling the task.
Transport_Delay_sec	Indicates any delay encountered during the movement of items within the warehouse.
Actual_Demand	Shows the actual number of units required or consumed during that timeframe.
Target	Flags the inventory status: 0 for stable stock and 1 for stock at risk.

3.2 Data pre-processing by employing min-max normalization

The process of cleaning, organizing and removing irrelevant and anomalous data from the obtained raw and unclean data is known as data pre-processing. This research utilized the min-max normalization method for pre-processing the information. The min-max normalization is mainly used to normalize and scale different operational metrics to a consistent range, commonly [0, 1], which is utilized for data pre-processing in an intelligent warehouse optimization performance. Equation (1) represents the calculation of min-max normalization.

$$x' = \frac{x - \text{mini}\,(b)}{maxi(b) - mini(b)} \qquad (1)$$

The min-max normalization is represented as x', x is the initial value, and the minimum and maximum numbers are represented by $mini(b)$ and $maxi(b)$.

3.3 Feature extraction through principal component analysis (PCA)

An efficient feature extraction PCA method is used for dimensionality reduction. As a consequence, the smallest details are eliminated while the most standard measurements are maintained. It is used to determine the enhanced operational variations, and improved computations efficiency along with removing the redundancy presented in the dataset features. Every instance of the pattern in a d-dimensional domain is transformed into a D-dimensional spatial domain through PCA, which creates new features that are a proportional composition of the initial characteristics and vectors. The maximal variance of each PC excludes variance, and a newly constructed representation of k is referred to as the principal components (PC). Both of the preceding elements are taken into consideration. As a result, the highest variance is maintained by the initial component, while the other component secures a lower variance value. The primary components are represented in Equation (2).

$$PC_j = x_1A_1 + x_2A_2 + \cdots + x_DA_D \quad (2)$$

Variables $x_1, \ldots, x_D$ stands for the number of coefficients, and $A_1, \ldots, A_D$ is the starting function, while j^{th} the component serves as the PC_j.

3.4 Anomaly detection in inventory data utilizing komodo mlipir algorithm-tuned random decision tree classifier (KMA-RDTC)

An intelligent warehouse optimization system based on IoT and multi-source data for improving resource efficiency, operating cost reduction, and increasing inventory precision is the intent of the research. The ML-based KMA-RDTC method is proposed in the research that integrates the Random Decision Tree Classifier (RDTC) for efficient prediction of anomaly detection in inventory data with the Komodo Mlipir Algorithm (KMA) for fine-tuning the RDTC parameters in determining an efficient forecasting performance of intelligent warehouse optimization. Comprehensive explanations are as follows.

3.4.1 Random decision tree classifier (RDTC) for robust demand forecasting in intelligent warehouse optimization

The RDTC combines ensemble learning for robust demand forecasting, and decision trees provide fine-grained anomaly detection in inventory data with Random Forest (RF) and Decision Tree (DT) components. The RDTC is one of the ML classification methods, which automatically obtains an ensemble of predictable patterns from training data that are useful in analyzing information beyond training. Figure 2 displays the RDTC method's architecture.

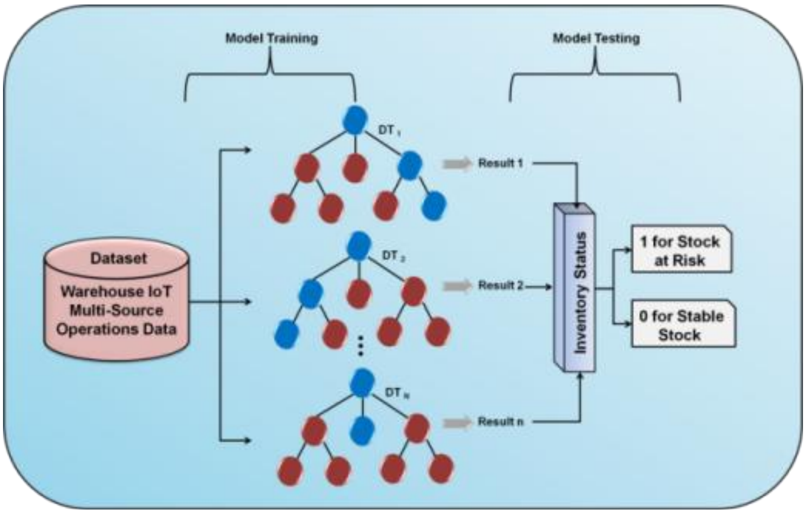

Figure 2: Architecture of the RDTC Method

DTs are the fundamental classifiers that make up RF, which are mixed classifier methods. Root nodes, internal nodes, and end nodes are the three different types of nodes found in a DT, which is a hierarchical structure made up of nodes and directed edges. The complete training dataset serves as the DT's single root node. The $Gini$ indicators are expressed in Equation (3), assuming that the information from specific node s includes data of l categories.

$$Gini(s) = 1 - \sum_{i=1}^{l}[q(i|s)]^2 \quad (3)$$

Variable q is the possibility that the category i attains a node s. When the category field is distributed evenly, $Gini(s)$ is the biggest, and the pertinent data at the time is insufficient. When the minimal $Gini$ is 0, every collection at the selected node belongs to the same category, implying that the most useful information is potentially retrieved. Assuming the set of data is split into k segments, the $Gini$ index of every segment (Equation 4).

$$Gini(S) = \sum_{j=1}^{k}\left(\frac{N_j}{N}\right)Gini(j) \quad (4)$$

Variable N is the data count at the parent node, N_j is the child node data counts with j^{th} node, and k is the child nodes. The preceding tree-development procedure is repeated by Random Forest to create a mixture of several DTs. The RF is robust to over-fitting. It is established that the generalization error's upper bound is less (Equation 5).

$$\frac{\bar{\rho}(1-t^2)}{t^2} \quad (5)$$

Where t is a single tree's classification effectiveness, and $\bar{\rho}$ is the average ratio of correlation between trees. Consequently, it eventually results in a scale of inappropriate points that represents the level of abnormalities of the selection process. Determine the sample N initial anomalous score scale using Equation (6).

$$Rawom(N) = \frac{N}{\bar{q}(N)} \quad (6)$$

A sample in an identical class has a large $Rawom$ value if its $\bar{q}(N)$ value is low. The mapping relationship between an attribute of an object and a particular value or value type of the feature is represented by each branch in the structure. Every leafless node in the RDTC denotes an evaluation condition. Each decision condition is associated with a product feature, and every branch path denotes the attribute value that corresponds to the evaluation condition. The building of decision trees typically involves the subsequent stages of development, such as RDTC generation, which is the process of creating a tree that uses training sample sets and RDTC adjustment, which is the process of verifying, adjusting, and updating the hierarchy of decisions after the creation. RDTC reduction is not an issue when every branch is weak, and there were no over-fitting phenomena. The following Equation (7) is the RDTC generation process's input data in the structure.

$$J = \{(X_{00} \dots, X_{0i} \dots, X_{0N}, S_0) \dots (X_{j0} \dots, X_{ji} \dots, X_{jN}, S_j) \dots\} \quad (7)$$

A binary tree or multi-branch tree is the outcome of decision tree building, where S_0 and S_j are the type marks of $the\ J^{th}$ sample, and X is the value of the i^{th} attribute of $the\ j^{th}$ sample in the collection $(X_{00} \dots, X_{0i} \dots, X_{0N})$ and $(X_{j0} \dots, X_{ji} \dots, X_{jN})$. To choose split attributes, many RDTC sorting algorithms employ various evaluation conditions. The two most crucial evaluation criteria are data gain and data gain percentage. Based on information gain, split attribute selection is used. Assume that the attribute set is q and the trained data ensemble is Q with j and M dimensions (Equation 8).

$$Q = \{q_1, \dots, q_j, \dots q_M\} \quad (8)$$

The percentage of items $(Q(D_i))$ in the total data that fit into i^{th} and j^{th} range is determined in Equation (9), and the test product dataset $Entropy(T, q_j)$ information entropy with N dimensions is indicated by Equation (10).

$$Q(D_i) = \frac{|T_{ji}|}{|T|} \quad (9)$$

$$Entropy(T, q_j) = \sum_{i=1}^{N} -Q(D_i)\, log_2 Q(D_i) \quad (10)$$

Assume that the value range in the sample dataset corresponding to the attribute q_j and u_j, while the subset of samples where attribute q_j acquires the value u is represented by $Gain(T, q_j)$, and the data gain of the data collection T to the attribute q_j is denoted as Equation (11).

$$Gain(T, q_j) = Entropy(T, q_j) - \sum_{u \epsilon u_j} \frac{|T_j(u)|}{|T|} Entropy(T, q_j) \quad (11)$$

$$Splitgain(T, q_j) = \sum_{u \epsilon u_j} \frac{|T_j(u)|}{|T|} log_2 \frac{|T_j(u)|}{|T|} \quad (12)$$

$$Gainratio(T, q_j) = \frac{Gain(T, q_j)}{Splitinfo(T, q_j)} \quad (13)$$

The proportion of information of the attribute $Splitinfo(T, q_j)$ is computed as follows, while the data gain of the testing set is T_j, and the test dataset T information gain rate ($Gainratio$) in relation to the attribute q_j was determined in Equations (12-13). When robust prediction and anomaly detection are needed across vast, heterogeneous datasets, the RDTC performs highly effectively. Demand forecasting, fraud detection, quality inspection, defect diagnosis, and dynamic inventory monitoring are all possible applications. It is appropriate for real-time transportation, industrial automation, and sensor-based monitoring systems due to the ensemble structure, which improves accuracy in noisy conditions.

3.4.2 Komodo mlipir algorithm (KMA) for fine-tuning the RDTC's hyper parameter performances

The Komodo dragon is a member of the Varanidae family and a survivor of the dinosaur era. To optimize the hyperparameters of the RDTC method, the KMA is utilized. The KMA is a population-based metaheuristic inspired by the social hierarchy, foraging strategies, and reproductive mechanisms of Komodo dragons [34]. Males of Komodo dragons are capable of searching for food from several kilometres distant and frequently feed on weaker offspring and larvae. Frequently, Komodo dragon females are parthenogenetic, indicating females do not require males to produce newborns. Figure 3 represents the algorithmic flow of KMA optimization.

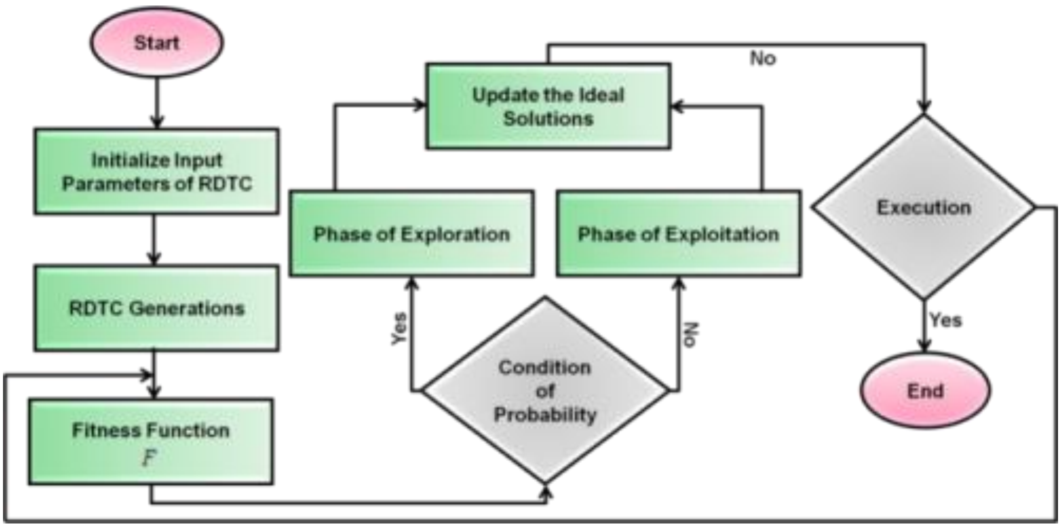

Figure 3: Algorithmic Flowchart of KMA Optimization

The KMA simulates the Komodo dragon's social order, foraging, and breeding. Based on their fitness (qualities), Komodo dragon individuals (N) were divided into three categories such as huge males, females, and small males. Large and smaller males in the KMA are optimized by movement, while females are maximized by reproducing. The KMA introduces the idea of Mlipir. In addition to hunting each other, small groups of male Komodo dragons frequently search for food particles from larger male dragon individuals. Females avoid the possibility of being consumed by massive males by engaging in the behaviour. Individuals refer to the following behaviour as the mobility of small males known as Mlipir. The Mlipir movement mechanism enables small males to explore promising regions identified by high-fitness individuals while simultaneously avoiding domination, thereby balancing exploration and exploitation in the search process. High-quality large men tend to respond to low-quality large males, while large males with less quality have a 0.5 capturing chance or excluding high-quality large males. KMA is selected for RDTC hyperparameter optimization because the search space is high-dimensional and non-linear, where traditional grid search becomes computationally expensive and random search lacks adaptive guidance. Compared with standard tuning algorithms, KMA maintains population diversity while dynamically refining high-quality solutions, leading to efficient convergence toward near-optimal hyperparameter configurations.

The computational model states (U_{ji} and l_j') that enormous men exhibit the following behaviours expressed in equations (14–15).

$$U_{ji} = \begin{matrix} O_1(l_i - l_j)\ if\ F(l_i) < F(l_j)\ or\ O_2 < 0.5 \\ O_1(l_j - l_i)\ Other \end{matrix}$$
(14)

$l_j' = l_j + \sum_i^p U_{ji}\ ,\ \ Where, i \neq j$(15)

The fitness (or quality) of i and j large men is represented by $F(l_i)$ and $F(l_j)$. The positions (l_j') of the large men are represented by l_i and l_j, the large male quantity is represented as p, and O_1 and O_2 were the pair of arbitrary counts with the distribution. To produce two offspring, a female and a huge male of the highest quality are exploited through the phase of exploration. Following is a mathematical description of the exploration process between two individuals (Equations 16-17).

$l_{jk}' = O_k \cdot l_{jk} + (1 - O_k) \cdot l_{ik}$(16)

$l_{ik}' = O_k \cdot l_{ik} + (1 - O_k) \cdot l_{jk}$(17)

Where the random amount in the interval of the k^{th} dimension normal distribution is O_k, the large males with high quality are denoted by l_{jk} and l_{ik} and female in the k^{th} dimension is indicated by two offspring (l_{jk}' and l_{ik}') with the process of exploration with the highest quality of large male and female . To complete the exploitation process, each female Komodo dragon's distinct dimension receives a low value. Symmetric average distribution is used to generate a small value at random (l_{ji}'), with the following Equations (18-19).

$(l_{j1}, l_{j2}, \dots, l_{jn}) \to (l_{j1}', l_{j2}', \dots, l_{jn}')$(18)

$l_{ji}' = l_{ji} + (2O - 1)\alpha|upbnd_i - lowbnd_i|$(19)

Where $l_{j1}, l_{j2}, \dots, l_{jn} \in lowbnd_i, upbnd_i$ represent the vector of position in the Komodo dragon individuals ($l_{j1}', l_{j2}', \dots, l_{jn}'$) with n dimension, while the lower and higher boundaries are denoted by $lowbnd_i$ and $upbnd_i$ with i dimension, the normal distribution (l_{ji}) of random number is determined by O , and α is an exploitation distance measure at 0.1. Equations (20-21) are an explanation of the short path travel, while percentages are chosen at random for interaction with the large male.

$$U_{ji} = \begin{matrix} \sum_{k=1}^{N} O_1(l_{ik} - l_{jk})\ if\ O_2 < c \\ 0 \quad Other \end{matrix}$$
(20)

$l_j' = l_j + \sum_p^{l=1} U_{ji}\ , Where\ i \neq j$(21)

Where O_1 and O_2 are the selections at random in the duration $[0, 1]$ with the typical distribution, l_{ik} and l_{jk} are small males (j), and the large male (i) in the k dimension, mlipir probability is denoted by c, the random dimensional selection (l_j) using the normal distribution is I, and p is the total amount of massive males. An adaptive model tuning and effective hyperparameter optimization are performed by the KMA. It is frequently used in real-time decision systems, resource-efficient optimization, predictive model selection, and data-intensive IoT applications. The KMA optimization is useful for automated operational intelligence systems, smart logistics, and predictive maintenance, while it enhances the model performance in forecasting, identifying, and detecting anomalies during the process.

Thus, KMA adaptively optimizes RDTC hyperparameters by balancing global exploration and local exploitation. The attraction-repulsion behavior of large males strengthens refinement, while the Mlipir movement of small males enhances diversity and prevents premature convergence. Parthenogenetic reproduction further intensifies local search around high-quality solutions. Through iterative fitness evaluation, KMA selects optimal tree depth, number of estimators, split criteria, and anomaly thresholds, improving classification accuracy, generalization, overfitting control, and convergence stability in mango defect detection.

3.4.3 Integrated KMA-RDTC applications in intelligent warehouse optimization

The KMA-RDTC is an advanced ML framework designed for robust anomaly detection and demand forecasting in intelligent warehouse systems. It integrates RDTs within an ensemble structure, enabling precise pattern recognition across large, heterogeneous inventory datasets while reducing overfitting through RF principles. Hyperparameter optimization is achieved using the KMA, a population-based metaheuristic inspired by Komodo dragon social hierarchy, foraging, and reproductive strategies. The Mlipir mechanism allows smaller male agents to explore promising regions identified by high-fitness individuals, balancing exploration and exploitation for efficient convergence. By combining RDTC's ensemble learning with KMA's adaptive tuning, the

KMA-RDTC framework enhances predictive accuracy, facilitates real-time anomaly detection, and supports dynamic inventory management, enabling optimized demand forecasting, automated operational decisions, and scalable smart warehouse performance in Industry 4.0 environments. Hyperparameters of KMA-RDTC are provided in Table 3. Algorithm 1 shows the KMA-RDTC method.

Table 3: Hyperparameters of the KMA-RDTC Method

Category	Hyperparameter	Search Range
RDTC	Number of Trees	100 – 300
	Maximum Tree Depth	10 – 50
	Minimum Samples Split	2 – 10
	Minimum Samples Leaf	1 – 5
	Feature Selection Strategy	{sqrt, log2, auto}
	Anomaly Score Threshold	0.5 – 0.9
KMA	Population Size	20 – 50
	Maximum Iterations	50 – 150
	Mlipir Probability	0.3 – 0.7
	Exploitation Distance	0.1 (fixed)
	Random Factors	[0,1]
Computational Parameters	Training Time	412 seconds
	Inference Time	3.8 ms/sample
	Memory Usage	2.4 GB
	Parameter Count	6
	Convergence Time	78 iterations (~325 seconds)

Algorithm 1: *KMA-RDTC*

Input:

Dataset $D = \{(X_1, S_1), (X_2, S_2), \ldots, (X_n, S_n)\}$
Hyperparameter search space H
Population size N
Maximum iterations T_{max}
Mlipir probability c
Exploitation factor α

Output:

Optimized RDTC model $M*$

Step 1: Initialize Komodo Population

1. Generate N candidate solutions $L = \{l_1, l_2, \ldots, l_N\}$ *where each* $l_j \in H$ *(random hyperparameter vectors)*
2. Evaluate fitness $F(l_j)$ *for each* l_j*:*
- *Train RDTC with hyperparameters* l_j *on Dtrain*
- *Compute validation accuracy / anomaly score*
- *Assign fitness value*

3. Sort population based on fitness
4. Divide population into:
- *Large males (best fitness group)*
- *Females (next best group)*
- *Small males (remaining group)*

Step 2: Main Optimization Loop

For $t = 1$ *to* T_{max} *:*
A. Large Male Movement (Dominance Update)
For each large male l_j*:*
For each other large male l_i $i \neq j$*:*
If $F(l_i) < F(l_j)$ *or* $O_2 < 0.5$*:*
$U_{ji} = O_1(l_j - l_i)$
Else:
$U_{ji} = O_1^*(l_j - l_i)$
Update position:
$l_j' = l_j + \sum_i^p U_{ji}$
B. Female Reproduction (Exploration)
For each female l_j*:*
Select best large male l_i
For each dimension k*:*
$O_k = random(0,1)$
$l_{jk}' = O_k \cdot l_{jk} + (1 - O_k) \cdot l_{ik}$
$l_{ik}' = O_k \cdot l_{ik} + (1 - O_k) \cdot l_{jk}$
C. Female Exploitation (Local Search)
For each female l_j*:*
For each dimension i*:*
$O = random(0,1)$
$l_{ji}' = l_{ji} + (2O - 1)\alpha|upbnd_i - lowbnd_i|$
D. Mlipir Movement (Small Male Following)
For each small male l_j*:*
If $O_2 < c$*:*
Select large male l_i *randomly*
For each dimension k*:*
*Uji += O1 * (lik - ljk)*
$U_{ji} += O_1^*(l_j - l_i)$
Update:
$l_j' = l_j + \sum_i^p U_{ji}$
E. Boundary Control
Ensure all updated l_j' *remain within H bounds*
F. Fitness Evaluation
For each updated l_j'*:*
Train RDTC using l_j'
Compute fitness $F(l_j')$
G. Population Update
Combine old and new individuals
Sort by fitness

Select top N individuals
Reassign roles (large males, females, small males)
End For
Step 3: Final Model Training
Select best hyperparameter vector l_{best}
Train RDTC on full dataset D using l_{best}
Return optimized model $M*$

4 Experimental results

The research was intended to provide an IoT-enabled, ML-based KMA-RDTC method for efficient optimization of intelligent warehouse inventory data anomalous detection.

Implementation performance is executed by employing several system specifications that include Python, 30 GB RAM, and an Intel Core i7 processor in Windows 11. The proposed KMA-RDTC method shows an enhanced performance on various parameters with the obtained data, which are as follows.

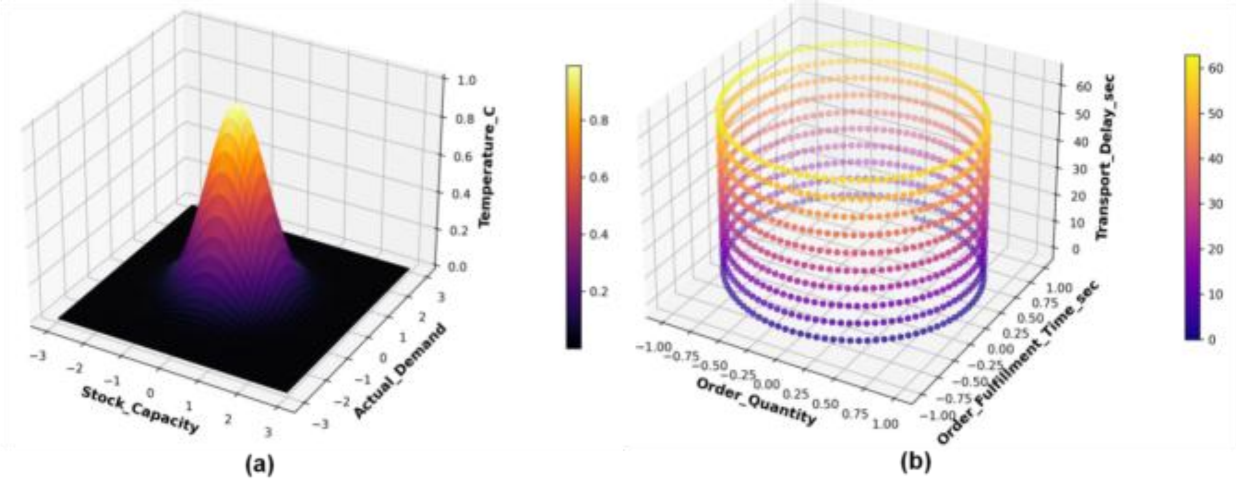

Figure 4: Results of (a) Stock Capacity and Temperature-based Fluctuations and (b) Dynamic Transformation

Figure 4(a) shows the relationship between stock capacity and temperature-based fluctuations observed by the IoT sensors. Figure 4(b) provides an attention to dynamic transformation in time-based fulfilment parameters, stock level, and demand. It supports the predictive KMA-RDTC optimization method by determining a complex association identified in real-time warehouse systems.

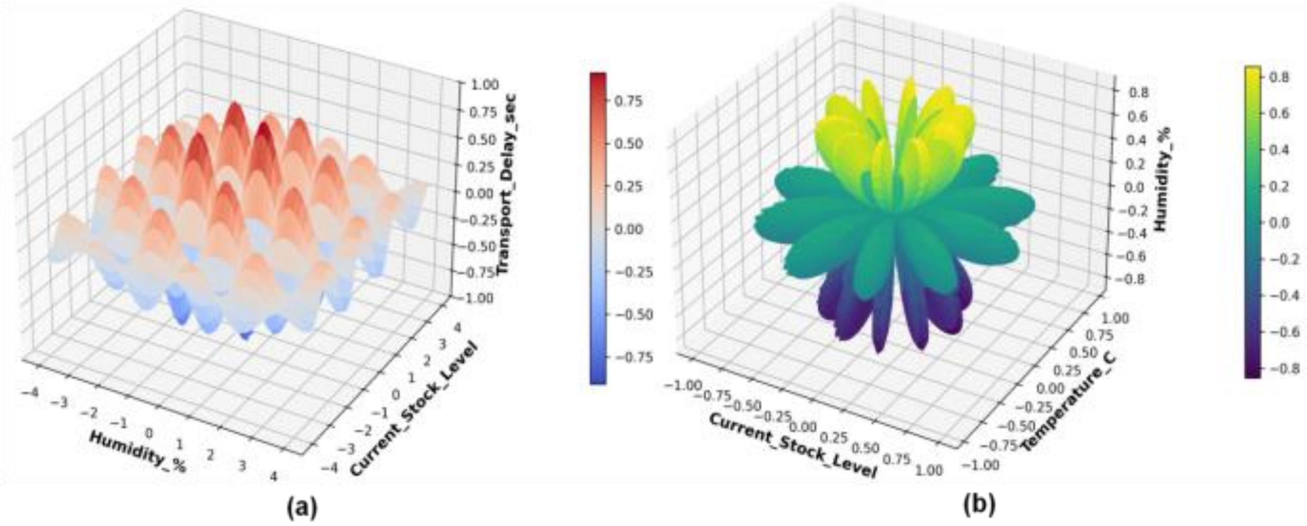

Figure 5: Visual Depictions of (a) Humidity Variations and (b) Multi-Sensor Interactions

Humidity variations impact shipment delay by illustrating the way environmental conditions influence warehouse productivity in Figure 5(a). Temperature, stock level, and humidity are connected in Figure 5(b), which shows multi-sensor interactions. An intelligent system manages the environmental uncertainty to make reliable forecasts and implement adaptive inventory control.

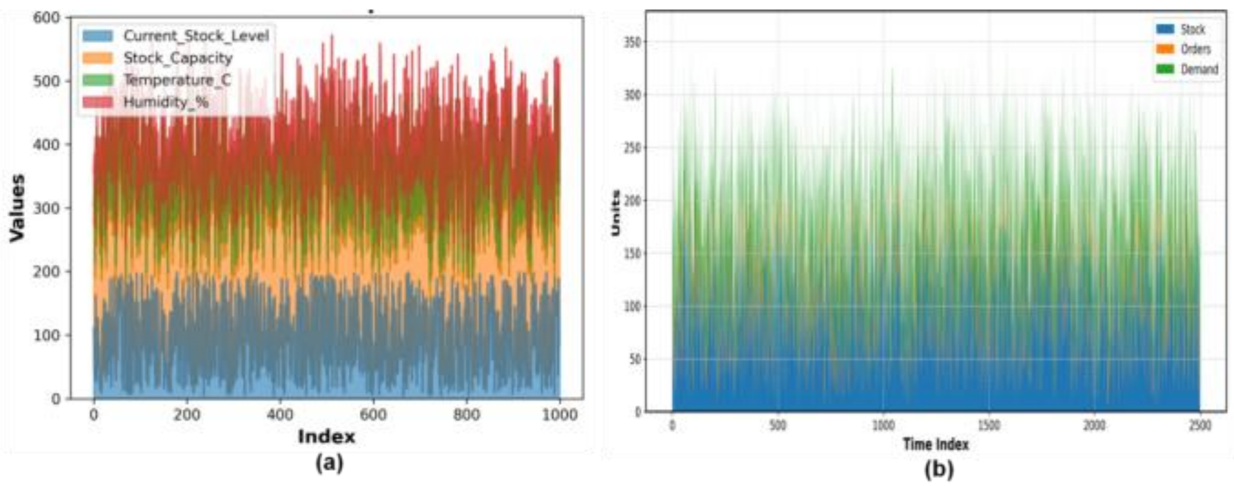

Figure 6: Results of (a) Synchronized Time-Series and (b) Stock and Order Patterns

Synchronized time-series show continuous data collection from the IoT devices that record capacity, stock levels, and environmental factors in Figure 6(a). Operational variability in Figure 6(b) highlighted the patterns in stock and order patterns. These findings show essential real-time data integration in the suggested model's enhanced fulfilment forecasting and dynamic stock allocation.

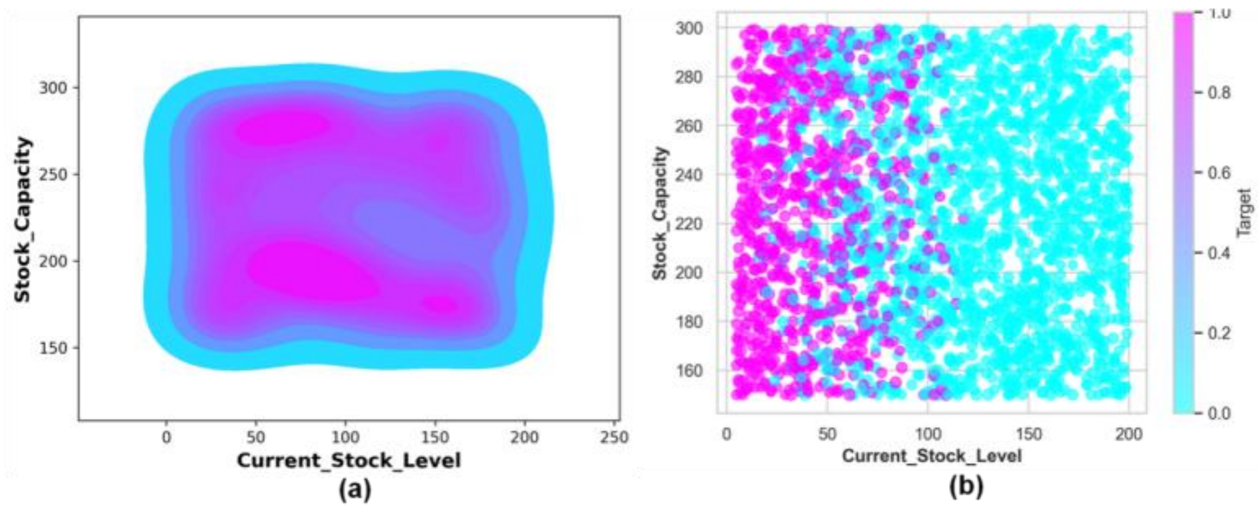

Figure 7: Evaluation of Stock Capacity vs Current Stock Level in (a) Non-Linear Relationship and (b) Operational State Associations

A non-linear relationship between inventory levels and stock capacity is shown in Figure 7(a). Target-based results in Figure 7(b) explore operational states associated with anomalous or demand circumstances.

.

Results highlighted the multi-level feature of warehouse variability and the necessity of feature extraction along with the ensemble learning in the proposed method.

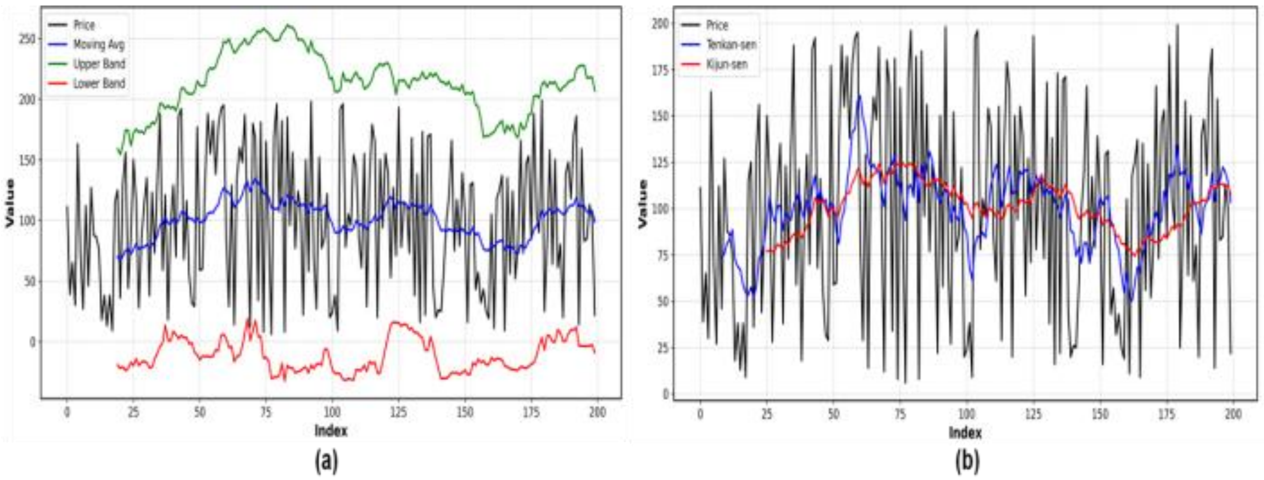

Figure 8: Results of (a) varying inventory levels and operational unpredictability and (b) patterns in warehouse behaviour

In Figure 8(a), varying inventory levels and operational unpredictability are demonstrated, while in Figure 8(b), both immediate and long-term patterns in warehouse behaviour are highlighted. The proposed method addresses the temporal issues by combining real-time IoT data with enhanced forecasting for increased stability.

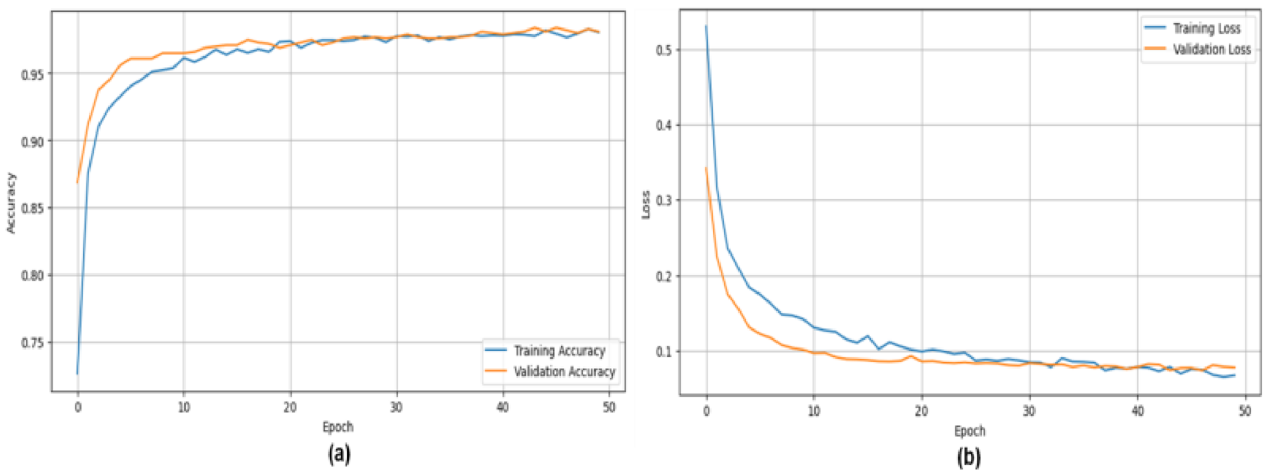

Figure 9: Results of accuracy and loss validation

The number of accurate predictions made is known as the accuracy score (Figure 9a). The numbers that show the deviation from the intended target state are called loss values (Figure 9b). The performance of a forecasting model is evaluated by the accuracy, which presents the forecasting amounts when the anticipated value equals the actual value is known as accuracy. It is frequently represented and evaluated during the training phase. A loss function considers the probability or uncertainty of a forecast based on the degree that the prediction fluctuates from the actual value in the performance of intelligent warehouse optimization. The errors created for every sample in train or validation sets are added collectively.

The performance parameters demonstrate the efficiency and robustness of the proposed optimization framework. Demand Variability Handling reflects adaptability to fluctuating order volumes without performance decline. Cost Reduction indicates improved resource allocation, scheduling, and inventory control. Energy Consumption of the Model represents computational power usage during training and execution, showing energy efficiency. Inventory Precision measures improved stock accuracy and reduced discrepancies. Operational Time denotes total execution duration, highlighting computational efficiency. Error Detection Accuracy reflects reliable anomaly and defect identification. Fulfilment Latency Reduction indicates faster order processing and improved logistics coordination. These parameters confirm enhanced adaptability, efficiency, accuracy, and operational stability.

Table 4: Proposed KMA-RDTC method performance

Parameter	Values
Demand Variability Handling	38%
Cost Reduction	22%
Energy Consumption of the Model	145Watts
Inventory Precision	30%
Operational Time	18min
Error Detection Accuracy	92%
Fulfilment Latency Reduction	25%

Table 4 provides the outcomes of the proposed method's significant performance by exploring fulfillment latency reduction (25%), error detection accuracy (92%), operational time (18min), inventory precision (30%), energy consumption (145Watts), cost reduction (22%), and demand variability handling (38%). Table 5 examines the ablation results of the proposed KMA-RDTC method improvements during integration.

Table 5: Ablation Results of the KMA-RDTC Method

Techniques	Accuracy	MAE	RMSE
RDTC	0.932	0.11	0.18
KMA	0.945	0.09	0.15
KMA-RDTC	0.968	0.07	0.10

According to the ablation results, the hybrid KMA-RDTC method provides essential outcomes with accuracy (0.968), MAE (0.07), and RMSE (0.10), which shows that an integrated method represents more improvements.

The classification effectiveness of the proposed KMA-RDTC model for identifying "stock at risk" (1) and "stable stock" (0) conditions is given in Table 6. Precision showed how accurately predicted stock conditions were classified without excessive false alarms, while recall reflected the model's ability to correctly detect actual risky or stable instances. The F1-score combined precision and recall into a single balanced measure, ensuring both false positives and false negatives were minimized.

Table 6: Proposed KMA-RDTC model classification performance for binary stock classification

Metrics (%)	1-Stock at Risk	0-Stable Stock
F1-Score	0.946	0.978
Precision	0.958	0.976
Recall	0.934	0.962

The proposed KMA-RDTC model delivered balanced and reliable classification for both "stock at risk" (1) and "stable stock" (0) conditions. High precision, recall, and F1-scores confirmed accurate risk detection and stable stock recognition with minimal misclassification, while an AUC of 0.973 indicated strong separability between the two classes. The results validated robust anomaly detection and dependable real-time decision support in intelligent warehouse operations.

4.1 Phase of comparison performance

Developing an innovative IoT-enabled and ML-based KMA-RDTC method for providing an efficient optimization performance of intelligent warehouse operations is the main concern of the research. This section shows the comparison between the proposed and various existing methods, which include RF [29], Logistic Regression (LR) [29], DT [29], K-Nearest Neighbour (KNN) [29], XGBoost [29], Kalman Filter [30], LSTM [30], Hybrid LSTM-Kalman [30], Decision Tree (DT) [31], Logistic Regression (LR) [31], Support Vector Machine (SVM) [31], and Random Forest (RF) [31]. It employs accuracy as a predictive metric with Prediction Accuracy, Sensitivity, and Specificity as a classification metrics and MAE and RMSE as an error metric.

- Accuracy: It shows the percentage of the total forecasts in both real and false cases. Accuracy calculation in mathematical format is expressed in Equation (22).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (22)$$

Note: TP represents the true positives, TN indicates the true negatives, FP denotes the false positives, and FN denotes the false negatives.

Table 7: Numerical Results of the KMA-RDTC and Conventional Methods with Accuracy

Techniques	Accuracy
RF [29]	0.93
LR [29]	0.97
DT [29]	0.945
KNN [29]	0.92
XGBoost [29]	0.955
KMA-RDTC [Proposed]	0.985

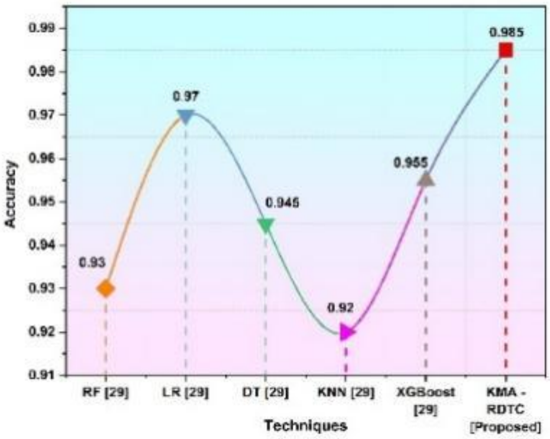

Figure 10: Outcomes of Proposed and Existing Methods with AccuracyIn the above, Table 7 and Figure 10 represent the accuracy results during comparison. The proposed KMA-RDTC method indicates more accuracy (0.985), which outperforms the traditional ML methods in optimizing the intelligent warehouse operations.

- MAE: The mean variations between the actual and anticipated values show the average amount of errors without allowing directions into the factor (Equation 23).

$$MAE = \frac{1}{N}\sum_{j=1}^{N}\left|b_j - \hat{b}_j\right| \quad (23)$$

- RMSE: The standard deviation of the remaining data (errors) is measured by taking the square root of the Mean Squared Error (MSE). Through squaring, it significantly impacts greater errors (Equation 24).

$$RMSE = \sqrt{\frac{1}{N}\sum_{j=1}^{N}\left(b_j - \hat{b}_j\right)^2} \quad (24)$$

Note: Variable N is the data point quantities, b_j indicates an actual value, $\hat{b}_j$ is the forecasted value, $\left|b_j - \hat{b}_j\right|$ determines an absolute result, and $\left(b_j - \hat{b}_j\right)^2$ denotes the squared variations.

Table 8: Comparison Results of MAE and RMSE

Techniques	MAE	RMSE
Kalman Filter [30]	0.14	0.16
LSTM [30]	0.11	0.14
Hybrid LSTM-Kalman [30]	0.09	0.12
KMA-RDTC [Proposed]	0.07	0.10

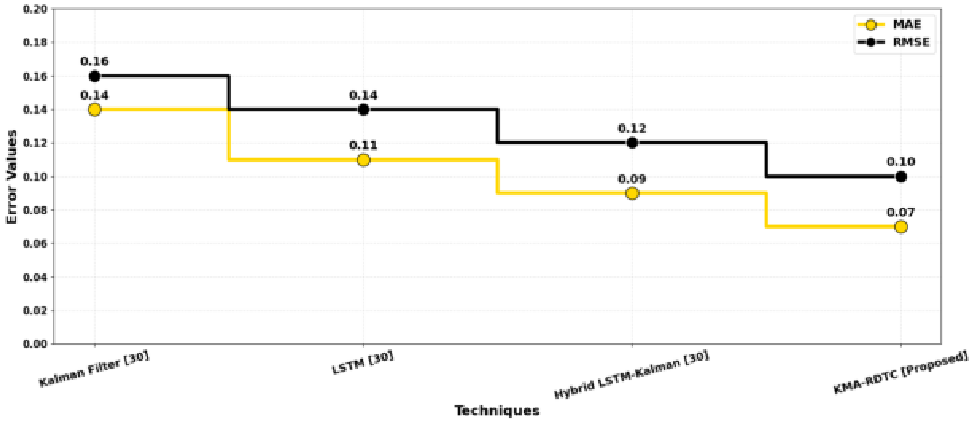

Figure 11: Result of MAE and RMSE

Error metrics such as MAE and RMSE results of the proposed and existing methods are shown in Table 8 and Figure 11. Whereas, the MAE of the KMA-RDTC method is 0.07 and the RMSE is 0.10, which indicates significant results of the proposed method in an intelligent warehouse optimization system compared to the conventional methods.

Prediction Accuracy denotes the proportion of correctly classified instances out of the total evaluated samples, reflecting the model's overall capability to accurately distinguish between normal and anomalous warehouse operational conditions. Sensitivity also referred to as the true positive rate, measures the capability to correctly detect actual positive cases, such as operational anomalies, demand fluctuations, or system irregularities. Specificity, or the true negative rate, evaluates the capability to correctly recognize normal operational states without false labelling as anomalies.

Table 9: Evaluation of classification performance across different warehouse anomaly detection models

Methods	**Prediction Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**
DT [31]	85.3	81.0	87.5
LR [31]	87.8	85.5	90.2
SVM [31]	90.5	88.0	92.0
RF [31]	93.2	90.8	94.5
KMA-RDTC [Proposed]	95.8	93.4	96.2

Table 9 shows that the proposed KMA-RDTC achieved a Prediction Accuracy of 95.8%, Sensitivity of 93.4%, and Specificity of 96.2%, outperforming conventional DT, LR, SVM, and RF methods. These results indicate more accurate anomaly detection, effective identification of operational irregularities, and enhanced reliability for real-time warehouse monitoring and optimization.

4.2 Discussions

The investigation in [29] demonstrated that integrating XGBoost, Random Forest, and artificial data generation optimized demand forecasting, reduced inventory discrepancies, and improved sustainability by lowering fuel use and material waste. However, the model struggled under complex environmental conditions, showing predictive frameworks needed greater robustness for dynamic warehouses. The IPS in [30], combining UWB sensors with LSTM and Kalman filtering, improved localization accuracy by 4% and enabled real-time object tracking. Scalability remained a concern, as it was less effective for larger or multi-level operations. In [31], IoT sensors and computer vision monitored physiological and motion data, generating predictive insights for fatigue and injury risk. This improved human-robot collaboration and overall warehouse performance, though effectiveness depended on sensor placement and data quality, indicating a need for standardized sensing strategies.

To overcome the limitations, the research proposed the KMA-RDTC model, which combined robust classification and anomaly detection with intelligent hyperparameter optimization. The RDTC handles complex, multi-source IoT data for accurate inventory prediction and anomaly detection, while the KMA fine-tunes RDTC parameters through exploration and exploitation strategies inspired by natural behavior. This joint mechanism allows the model to adapt dynamically to changing warehouse conditions, avoid suboptimal solutions, and converge efficiently, enabling real-time decision-making.

The findings demonstrated that the proposed KMA-RDTC model outperformed conventional methods in intelligent warehouse optimization. The model achieved high accuracy (0.985) and prediction accuracy (0.958) reflected the model's ability to correctly classify normal and anomalous inventory conditions. Sensitivity (0.934) indicated effective detection of actual anomalies and demand fluctuations, while specificity (0.962) confirmed that normal operations were rarely misclassified. Low RMSE (0.10) and MAE (0.07) showed precise alignment between predicted and actual values, highlighting the robustness of demand forecasting. These results indicated that the model effectively handled multi-source IoT data in real time, providing accurate demand forecasting and robust anomaly detection under dynamic warehouse conditions. Unlike conventional ML or hybrid models, the integrated KMA-RDTC framework improved operational decision-making, optimized resource utilization, and enhanced the reliability of warehouse management systems, demonstrating clear superiority over traditional approaches.

4.3 Practical implications

The proposed method shows applicability across commercial and industrial environments. The IoT-based multi-source structure assists storage facilities in identifying environmental anomalies, estimating demand, and monitoring operational conditions. Organizations can reduce risks, allocate resources efficiently, and detect issues early using the KMA-RDTC method. The scalable framework supports multi-source operations while improving productivity, planning accuracy, operational costs, and service responsiveness in Industry 4.0 warehouse systems.

The analytical framework can support operational monitoring and decision support in modern warehousing environments by integrating anomaly detection, demand analysis, and inventory monitoring within a unified system. Such integration assists operational planning and resource coordination across different parts of a warehouse environment and improves the understanding of inventory conditions and operational variations. The framework can also assist settings where automated technologies and human operators work together by providing predictive insights that support scheduling decisions, inventory planning, and operational adjustments when unusual patterns appear in stock levels or order processing activities. From an operational perspective, warehouse activities can be viewed across multiple functional levels such as storage locations, operational sections, and overall warehouse management, where analytical outputs from the proposed model can support monitoring and coordinated responses to demand variations, stock conditions, and operational changes. By linking predictive analytics with operational insights, the approach contributes to practical decision support in data-driven warehouse management systems that rely on continuous information from interconnected sensing and tracking technologies.

5 Conclusions

The research developed an IoT-enabled, ML-based method for effectively evaluating and optimizing the intelligent warehouse operations. The warehouse IoT multi-source operations data, with 2500 real-time warehouse operations, were obtained and pre-processed by employing the min-max normalization for providing data quality and consistency. The PCA method was used to reduce the dimensionality while providing variations in the predictive model. The KMA-RDTC method was proposed in the research, where the RDTC method determined the effective forecasting and optimized performance in intelligent warehouse operations, while the KMA optimization significantly fine-tuned the RDTC method by enhancing the performance. Comparisons of the proposed and existing methods were performed with a Python implementation. Through crucial results in terms of accuracy (0.968), RMSE (0.10), and MAE (0.07), the proposed KMA-RDTC method demonstrated more significance in intelligent warehouse optimization operations. It provided a comprehensive design for smart warehouse operations, emphasizing the potential of multi-source, IoT-driven data systems to achieve high-performance, sustainable logistics within Industry 4.0 circumstances.

5.1 Limitations and future directions

The potential latency in massive data sources and reliance on constant sensor connectivity are several drawbacks presented in the proposed method. In low-signal or unstructured circumstances, it varied with the performances. The model was validated on a specific dataset and setting, limiting generalizability across heterogeneous warehouses, and cybersecurity risks in IoT data transmission and centralized processing were not extensively addressed, potentially affecting secure industrial deployment.To create fully automated, adaptive warehouse environments in relation to Industry 4.0, future research should focus on adopting DT simulations for proactive decision-making, integrating more neural network methods for greater pattern detection, improving scalability through edge computing, and integrating autonomous robotics. It also extends validation across diverse warehouse environments and integrates robust cybersecurity measures, secure data transmission, and decentralized architectures to ensure scalable and secure industrial deployment.

Data availibility statement

All data generated or analysed during this study are included in this article.

Funding

This work was supported by Henan Provincial Department of Science and Technology soft science research project "Research on the construction and implementation strategy of Henan government affairs data sharing system based on the perspective of regional digital government construction" (No.232400411191).

Author contributions

Conceptualization, Jiping Liu, Chenxiao Liu; writing—original draft preparation, Jiping Liu, Chenxiao Liu; data curation, Jiping Liu, Chenxiao Liu; investigation, Jiping Liu, Chenxiao Liu; writing—review and editing,Jiping Liu, Chenxiao Liu.

References

[1] Likhouzova, T., & Demianova, Y. (2022). Robot path optimization in a warehouse management system. Evolutionary Intelligence, 15(4), 2589-2595. https://doi.org/10.1007/s12065-021-00614-w

[2] Derpich, I., Sepúlveda, J. M., Barraza, R., & Castro, F. (2022). Warehouse Optimization: Energy Efficient Layout and Design. Mathematics, 10(10), 1705. https://doi.org/10.3390/math10101705

[3] Stanisławski, R., & Szymonik, A. (2021). Impact of selected intelligent systems in logistics on the creation of a sustainable market position of manufacturing companies in Poland in the context of Industry 4.0. Sustainability, 13(7), 3996. https://doi.org/10.3390/su13073996

[4] 4.Li, H., Wang, S., Zhen, L., & Wang, X. (2024). Data-driven optimization for automated warehouse operations decarbonization. Annals of Operations Research, 343(3), 1129-1156. https://doi.org/10.1007/s10479-022-04972-1

[5] 5.Gao, J., Li, Y., Xu, Y., & Lv, S. (2022). A two-objective ILP model of OP-MATSP for the multi-robot task assignment in an intelligent warehouse. Applied Sciences, 12(10), 4843. https://doi.org/10.3390/app12104843

[6] 6.Yu, C., Liao, W., & Zu, L. (2023). Dynamic scheduling optimization method for multi-AGV-based intelligent warehouse considering bidirectional channel. Systems, 12(1), 9. https://doi.org/10.3390/systems12010009

[7] 7.Chen, Y., Wu, J., He, C., & Zhang, S. (2023). Intelligent warehouse robot path planning based on improved ant colony algorithm. IEEE Access, 11, 12360-12367. https://doi.org/10.1109/ACCESS.2023.3241960

[8] 8.Liu, J., & Liu, H. (2023). Research on Path Optimization Method for Warehouse Inspection Robot. Applied Artificial Intelligence, 37(1), 2254048. https://doi.org/10.1080/08839514.2023.2254048

[9] 9.Zhao, Q., Zhang, X., & Wang, P. (2024). Multi-type equipment selection and quantity decision optimization in intelligent warehouse. IEEE Access, 12, 63515-63527. https://doi.org/10.1109/ACCESS.2024.3395288

[10] 10.Dave, R., & Sarkar, B. (2023). The Rise of Intelligent Warehouses: A New Era of Efficiency and Sustainability. Int. J. Eng. Trends Technol, 71(7), 357-366. https://doi.org/10.14445/22315381/IJETT-V71I7P234

[11] 11.Liu, H., Zhou, L., Zhao, J., Wang, F., Yang, J., Liang, K., & Li, Z. (2022). Deep-learning-based accurate identification of warehouse goods for robot picking operations. Sustainability, 14(13), 7781. https://doi.org/10.3390/su14137781

[12] 12.Tufano, A., Accorsi, R., & Manzini, R. (2022). A machine learning approach for predictive warehouse design. The International Journal of Advanced Manufacturing Technology, 119(3), 2369-2392. https://doi.org/10.1007/s00170-021-08035-w

[13] 13.Leon, J. F., Li, Y., Martin, X. A., Calvet, L., Panadero, J., & Juan, A. A. (2023). A hybrid simulation and reinforcement learning algorithm for enhancing efficiency in warehouse operations. Algorithms, 16(9), 408. https://doi.org/10.3390/a16090408

[14] 14.Pracucci, A. (2024). Designing Digital Twin with IoT and AI in warehouse to support optimization and safety in Engineer-to-Order manufacturing process for prefabricated building products. Applied Sciences, 14(15), 6835. https://doi.org/10.3390/app14156835

[15] 15.Goli, A., Babaee Tirkolaee, E., Golmohammadi, A. M., Atan, Z., Weber, G. W., & Ali, S. S. (2023). A robust optimization model to design an IoT-based sustainable supply chain network with flexibility. Central European Journal of Operations Research, 33, 1-22. https://doi.org/10.1007/s10100-023-00870-4

[16] 16.Sahara, C. R., & Aamer, A. M. (2022). Real-time data integration of an internet-of-things-based smart warehouse: a case study. International Journal of Pervasive Computing and Communications, 18(5), 622-644. http://dx.doi.org/10.1108/IJPCC-08-2020-0113

[17] 17.Chen, J., Xu, S., Liu, K., Yao, S., Luo, X., & Wu, H. (2022). Intelligent transportation logistics optimal warehouse location method based on Internet of Things and blockchain technology. Sensors, 22(4), 1544. https://doi.org/10.3390/s22041544

[18] 18.Villegas-Ch, W., Navarro, A. M., & Sanchez-Viteri, S. (2024). Optimization of inventory management through computer vision and machine learning technologies. Intelligent Systems with Applications, 24, 200438. https://doi.org/10.1016/j.iswa.2024.200438

[19] 19.Dalal, S., Lilhore, U. K., Simaiya, S., Radulescu, M., & Belascu, L. (2024). Improving efficiency and sustainability via supply chain optimization through CNNs and BiLSTM. Technological Forecasting and Social Change, 209, 123841. https://doi.org/10.1016/j.techfore.2024.123841

[20] 20.Erdos, F., & Farhat, R. (2023). Sustainability Approach of SAP Application Management Service Solutions in the Field of Warehouse Management. Chemical Engineering Transactions, 107, 259-264. https://doi.org/10.3303/CET23107044

[21] 21.Liu, T., & Ju, H. (2024). Research on the Application of IOT and AI in Modern Logistics and Warehousing. J. Ind. Eng. Appl. Sci, 2, 1-4. https://doi.org/10.5281/zenodo.10755279

[22] 22.Hamdy, W., Al-Awamry, A., & Mostafa, N. (2022). Warehousing 4.0: A proposed system of using Node-Red for applying Internet of Things in warehousing. Sustainable Futures, 4, 100069. https://doi.org/10.1016/j.sftr.2022.100069

[23] 23.Jarašūnienė, A., Čižiūnienė, K., & Čereška, A. (2023). Research on impact of IoT on warehouse management. Sensors, 23(4), 2213. https://doi.org/10.3390/s23042213

[24] 24.Kalkha, H., Khiat, A., Bahnasse, A., & Ouajji, H. (2024). Enhancing warehouse efficiency with time series clustering: A hybrid storage location assignment strategy. IEEE Access, 12, 52110-52126. https://doi.org/10.1109/ACCESS.2024.3386887

[25] 25.Elbouzidi, A. D., Frédéric, R., Pellerin, R., Lamouri, S., & Ait El Cadi, A. (2025). Leveraging digital twins for enhanced sustainable warehouse management. Cleaner Logistics and Supply Chain, 17, 100287. https://doi.org/10.1016/j.clscn.2025.100287

[26] 27.Francuz, Á., & Bányai, T. (2025). Intelligent Control Approaches for Warehouse Performance Optimisation in Industry 4.0 Using Machine Learning. Future Internet, 17(10), 468. https://doi.org/10.3390/fi17100468

[27] 27.Lin, H. Y., Chang, K. L., & Huang, H. Y. (2024). Development of unmanned aerial vehicle navigation and warehouse inventory system based on reinforcement learning. Drones, 8(6), 220. https://doi.org/10.3390/drones8060220

[28] 28.Gnaś, D., Majerek, D., Styła, M., Adamkiewicz, P., Skowron, S., Sak-Skowron, M., ... & Pietrzyk, R. (2024). Enhanced indoor positioning system using ultra-wideband technology and machine learning algorithms for energy-efficient warehouse management. Energies, 17(16), 4125. https://doi.org/10.3390/en17164125

[29] 29.Arvind, V. R., Shrinidhi, R. M., Deepa, T., & Maheedhar, M. (2025). Intelligent Warehousing: A

Machine Learning and IoT Framework for Precision Inventory Optimization. IEEE Access, 13, 169381-169414. https://doi.org/10.1109/ACCESS.2025.3614679

[30] 30.Masoudi, B., Razi, N., & Rezazadeh, J. (2025). IoT-Enabled Indoor Real-Time Tracking Using UWB for Smart Warehouse Management. Computers, 14(12), 510. https://doi.org/10.3390/computers14120510

[31] 31.Ping, Gang. (2025). OPTIMIZING HUMAN FACTORS AND SAFETY MANAGEMENT IN SMART WAREHOUSES USING A FUSION OF INTERNET OF THINGS (IoT) AND COMPUTER VISION. Journal of Tianjin University Science and Technology. 58 (11). https://doi.org/10.5281/zenodo.17548327

[32] 32.Nookala, G. (2021). Automated Data Warehouse Optimization Using Machine Learning Algorithms. Journal of Computational Innovation, 1(1).

[33] 33.Devi, D. P., Allur, N. S., Dondapati, K., Chetlapalli, H., Kodadi, S., & Perumal, T. (2023). Digital twin technology and IoT-enabled AI using real-time analytics for smart warehouse management and predictive inventory optimization. International Journal of Marketing Management, 11(4), 78-98.

[34] 34.Li, L., & Zhao, M. (2022). A novel komodo mlipir algorithm and its application in pm2. 5 detections. Atmosphere, 13(12), 2051. https://doi.org/10.3390/atmos13122051