International Journal Software Engineering and Computer Science (IJSECS)
4 (1), 2024, 107-120
Published Online April 2024 in IJSECS (http://www.journal.lembagakita.org/index.php/ijsecs)
P-ISSN : 2776-4869, E-ISSN : 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v4i1.2326.

RESEARCH ARTICLE Open Access

Optimizing E-commerce Inventory to prevent Stock Outs using the Random Forest Algorithm Approach

Achmad Ridwan *
Information Systems Study Program, Universitas Muhammadiyah Kudus, Kudus Regency, Central Java Province, Indonesia.
Corresponding Email: achmadridwan@umkudus.ac.id.

Ully Muzakir
Computer Science Study Program, Faculty of Science, Technology and Health Sciences, Universitas Bina Bangsa Getsempena, Banda Aceh City, Aceh Province, Indonesia.
Email: ully@bbg.ac.id.

Safitri Nurhidayati
Economics and Business Study Program, Universitas Muhammadiyah Berau, Berau Regency, East Kalimantan, Indonesia.
Email: safitri.n091183@gmail.com.

Received: February 3, 2024; Accepted: March 10, 2024; Published: April 1, 2024.

Abstract: This research investigates the effectiveness of the Random Forest algorithm in optimizing e-commerce inventory management. In a digital business that continues to grow, inventory management is crucial for smooth operations and customer satisfaction. The Random Forest algorithm, a development of the CART method by applying bagging techniques and random feature selection, was tested to predict inventory. An experimental design is used to test the algorithm's performance algorithms performance, using data relevant to the observed inventory variables. The analysis involves evaluating the performance of algorithms in predicting and preventing stockouts. The results show that the Random Forest algorithm provides more accurate inventory predictions than traditional methods. Comparison with linear and rule-based regression reveals higher accuracy, making this algorithm a promising choice for e-commerce inventory management. These findings imply that the Random Forest Algorithm can be an effective solution in overcoming the complexity and fluctuations of digital markets. Practical recommendations include a deep understanding of the data, engagement of trained human resources, and training strategies for optimal use of these algorithms. This research also contributes to the literature by expanding understanding of the application of the Random Forest algorithm in various contexts, including forest basal area prediction, supply chain management, and backorder prediction. In conclusion, the Random Forest algorithm has great potential to improve e-commerce inventory management, opening up opportunities for broader application in the digital business world. Proactive adoption of these algorithms can have a positive impact on operational efficiency, customer satisfaction, and a company's competitiveness in an ever-evolving market.

Keywords: Ecommerce; Inventory Management; Random Forest Algorithm; Inventory Prediction, Out of Stock.

1. Introduction

The e-commerce industry, as one of the sectors that has experienced exponential growth in recent years, has changed the global business landscape significantly. Changes in consumer behavior, technological advances, and internet accessibility have become critical drivers for the rapid development of e-commerce. Today, e-commerce companies are not only places for online transactions but also centers for comprehensive customer experiences, offering a variety of products and services quickly and efficiently. However, with this rapid growth, the e-commerce industry is also faced with increasingly complex challenges, one of which is effective inventory management. Inventory management in e-commerce is a critical element that influences the success of company operations. Intense competition, customer expectations for fast delivery, and high demand variations place significant pressure on inventory systems. Out-of-stock can have detrimental impacts, such as loss of customers, decreased consumer confidence, and financial loss. Therefore, e-commerce companies need to develop sophisticated and effective strategies in inventory management to overcome these challenges. This research emerged in response to the urgent need to improve inventory management in the e-commerce industry. The main focus of the study is preventing stockouts, an issue that is a significant focus for e-commerce companies. Out-of-stock not only results in financial losses but can also damage a company's reputation and relationships with customers. Therefore, this research proposes the application of the Random Forest Algorithm as an innovative approach to improve prediction accuracy and the effectiveness of inventory management in preventing stockouts.

In response to these rapid market dynamics, recent studies have proposed various strategies and approaches to improve inventory management in the e-commerce industry. One interesting approach is the application of artificial intelligence algorithms and models, especially in stock-out prevention. Research by Shi, Wang, and Alwan (2020) on Analytics for Cross-Border E-Commerce provides essential insights into inventory management risks in the online fashion retail sector. They identified operational complexity, such as broad product offerings, high demand risks, and tax risks in cross-border trade, as the main factors causing inventory problems in overseas warehouses [1]. This research provides a strong foundation for further study, especially in applying a data-driven approach to overcome this problem. The predictive method applied by Shi, Wang, and Alwan (2020) shows that the use of various machine learning techniques, especially the Random Forest algorithm, can produce more accurate predictions, reduce unprofitable shipping costs, and provide practical solutions for inventory management in overseas warehouses [1].

Furthermore, research by Zhou (2023) on logistics inventory optimization methods for agricultural e-commerce platforms using a multilayer feedforward artificial neural network approach highlights the relevance of artificial intelligence models. With the increasing demand for quality and safe agricultural products, primarily through e-commerce, there is a need for an efficient inventory management strategy. This research provides evidence that deep learning algorithms can be applied to predict material flow, inventory, and prices more accurately than previous shallow neural network models [2]. Furthermore, research by Tang, Chau, Lau, and Zheng (2023) discusses the application of artificial intelligence models in cross-border e-commerce automation services, focusing on data-intensive inventory prediction. They highlight that artificial intelligence-based models, especially the XGBoost method, can provide optimal results in terms of accuracy and reasonable computing time. This research strengthens the view that using artificial intelligence models can improve the efficiency of cross-border e-commerce supply chains [3].

Another study by Pramodhini *et al*. (2023) discusses the use of machine learning approaches in e-commerce inventory management. The focus is on developing models that predict demand, providing a basis for more efficient inventory management for small and medium businesses [4]. Luo, Lu, and Li (2019) research discusses when and how to use e-commerce shopping cart tracking (ECT) technology. By integrating the concepts of shortage and price incentives, this research shows that ECT can significantly impact consumer purchasing decisions. This model provides a deeper understanding of how ECT design can interact with consumer mindsets in various stages of online shopping [5].

Furthermore, Hua *et al*. (2021) created a new approach to handling markdowns in fresh retail e-commerce. They use a semi-parametric structural approach to model individual price elasticities and optimize perishable product prices over time. This research opens new avenues in designing pricing strategies for fresh products with tight sales deadlines [6]. Research by Wang (2023) proposed an Internet of Things (IoT)-based framework to improve cross-border e-commerce supply chain performance. By combining machine learning techniques and multi-object optimization, this research shows how demand volume predictions can be improved, strengthening supply chain performance in Cross-Border E-Commerce (CBEC) [7]. In addition, research by Sharma, Deepika, and Singh (2021) states that effective inventory management involves predicting demand and organizing and maintaining efficient inventory. This research focuses on the role of

artificial intelligence technology, especially machine learning, in increasing inventory management efficiency in warehouses. The proposed system predicts inventory requirements based on customer demand patterns, enabling more timely and optimal stock management. The main advantage of this research is its ability to provide solutions that are adaptive and responsive to changes in market conditions and customer purchasing patterns. By implementing machine learning methods, this research encourages the development of predictive models to identify trends and changes in customer demand, enabling companies to structure their inventory more intelligently [8].

The main objective of this research is to optimize inventory management in the e-commerce industry. By utilizing the Random Forest Algorithm, this research aims to increase the accuracy of inventory predictions, reduce the risk of stockouts, and, overall, optimize the inventory management process. Thus, this research seeks to positively contribute to increasing the operational efficiency of e-commerce companies, increasing customer satisfaction, and minimizing the negative impact of stockouts. Using the Random Forest Algorithm, the research explores the potential and advantages of this method in e-commerce inventory management. This algorithm, with its ability to overcome data complexity and increase prediction accuracy, is expected to be an effective solution to face the high dynamics of demand and competition in today's e-commerce market.

E-commerce inventory management is a vital foundation that ensures smooth operations and business success. As digital markets continue to grow and diversify, effective inventory management becomes increasingly crucial. This can be the difference between logistical success and failure, especially for online businesses just starting or those with experience looking to optimize their operations. Outdated or inefficient inventory management practices can lead to missed business opportunities, customer dissatisfaction, and decreased profits. Technology-enabled inventory management systems, such as the one proposed by this research, provide real-time visibility, efficient picklists, and in-depth data monitoring. This research clarifies that good inventory management supported by technology is critical to deal with the complexity of digital markets.

In conclusion, a well-structured e-commerce inventory management strategy, including the implementation of the latest technology, is a must to prevent stock-outs, increase customer satisfaction, and maximize revenue potential. Inventory management in the e-commerce realm has a very vital role in ensuring smooth operations and business success. In a digital market that continues to develop with various changes, effective inventory management becomes increasingly crucial. This is the difference between success and failure in logistics, especially for online business players, both those who have just entered the arena and those who are experienced and are trying to optimize their operations. Using outdated or inefficient inventory management practices can quickly result in missed business opportunities, customer dissatisfaction, and decreased profits. In implementing inventory management in e-commerce, practitioners can take guidance from a range of relevant research. Schulz (1996) highlights the ABC analysis method as a simple but effective inventory management tool. The ABC analysis concept breaks down products into three categories based on their relative value to the total inventory value, guiding allocating resources more efficiently [9]. Furthermore, Womack, Jones, and Roos (1990) introduced a revolutionary approach to inventory management with the principle of procuring the proper inventory when needed, without excessive storage. The JIT method has been proven to reduce storage costs, increase cash flow, and optimize operational efficiency. Despite the risks associated with sudden changes in the supply chain or customer demand, JIT remains a popular method of inventory management [10][9].

Furthermore, Sheffield (2004) explains that strategies and techniques effectively manage inventory, especially in e-commerce, which is full of dynamics. Sheffield emphasized the specific challenges online businesses face, such as rapid fluctuations in demand and the need for adaptive inventory management processes. It is known that ABC analysis, the Just-in-Time concept, and a comprehensive guide to e-commerce inventory management are the foundation for building a solid strategy. Overall, this literature provides an in-depth understanding of the approaches and tools that can be used to address inventory management challenges in the dynamic world of e-commerce [11].

The Random Forest algorithm, a development of the CART method by applying bootstrap aggregating (bagging) techniques and random feature selection, has proven to be an effective method in inventory prediction. In its use, Random Forest builds several decision trees to form a forest, and analysis is carried out on the entire collection of trees. Its main advantages include its ability to save forests for future reference, overcome overfitting, and provide accurate values and important variables automatically. Various studies strengthen support for the effectiveness of Random Forests. Jevšenak and Skudnik (2021) successfully applied this model to predict the increase in national forest basal area [12]. Brosofske *et al*. (2014) stated the success of Random Forest in a review of inventory attribute mapping and prediction methods for forest management [13]. The application of this algorithm by Islam and Amin (2020) in predicting backorder scenarios in the

supply chain also shows high reliability [14]. Kilham *et al*. (2018) contributed to the efficiency of forest management by designing harvest predictions at the tree level using Random Forest [15]. Ntakolia *et al*. (2021) support using Random Forest in explainable machine-learning models to predict backorder events in inventory management [16]. The application of Random Forests in creating species distribution maps by Magness *et al*. (2008) also confirmed its effectiveness in various applications [17]. These references provide a solid theoretical foundation, validating that Random Forest is a general prediction algorithm and reliable in a wide range of cases, including inventory management, forests, supply chains, and ecological monitoring.

Various models and strategies can be employed to optimize inventory in e-commerce. Economic Order Quantity (EOQ) and Robust Optimization are commonly used models for inventory management [18]. Integrating an e-commerce system with inventory, purchasing, and sales management through software like Odoo 14 can enhance operational efficiency and reduce costs [19]. Implementing such systems can lead to cost savings, shorter ordering cycles, reduced inventory levels, faster purchase information, and improved buyer-supplier relationships [20]. Moreover, the development of e-commerce has created significant business prospects globally [21]. In managing cross-border e-commerce, effective trade process management is crucial, especially in B2B transactions, to ensure smooth operations and document handling. Understanding the legal frameworks and strategies for expanding e-commerce, such as having websites in foreign languages or targeting foreign markets, is essential for successful e-commerce ventures [22].

Furthermore, logistics plays a vital role in e-commerce performance, with efficient logistics operations directly impacting customer satisfaction and loyalty [23]. Financial management in e-commerce differs from traditional businesses, necessitating a focus on financing, financial information management, and transaction costs unique to e-commerce enterprises [24]. Evaluating the performance of cross-border e-commerce systems using models like Fuzzy DEA can help identify constraints such as platform imperfections, payment security loopholes, and policy lags that affect enterprise performance [25]. In conclusion, optimizing inventory in e-commerce involves leveraging models like EOQ, integrating systems for efficient operations, understanding legal frameworks, ensuring robust logistics, and managing finances effectively. By implementing these strategies, e-commerce businesses can enhance their competitiveness, improve customer satisfaction, and drive sustainable growth.

Various strategies and models can be considered to optimize inventory in e-commerce and prevent stockouts using the Random Forest algorithm. Implementing data mining techniques like K-Means clustering Zafira (2024) or using Internet of Things (IoT) based systems for inventory management Kurniawan (2023) can help in reducing the risk of stockouts and improving customer service [26][27]. Additionally, systems integrating methods like Safety Stock and Reorder Point (ROP) can effectively manage inventory levels to prevent stockouts [28]. Moreover, the application of algorithms such as Economic Order Quantity (EOQ) Hidayat (2022) and Weighted Moving Average (WMA) Gunarti *et al*. (2022) can aid in determining optimal reorder points and forecasting stock requirements. By using these methods, businesses can ensure efficient inventory control and prevent stockouts [29][30].

Furthermore, the implementation of Supply Chain Management (SCM) systems Sari *et al*. (2021) and the use of Double Exponential Smoothing for sales prediction Saputri & Huda (2020) can improve inventory management practices and reduce the likelihood of stockouts [31][32]. Understanding consumer behavior patterns through algorithms like FP-Growth Syahril (2023) can help in stocking popular items to prevent stockouts. Additionally, considering factors like variety of selection and impulse buying behavior can guide inventory decisions to meet customer demands effectively [33][34]. By leveraging these insights and algorithms, e-commerce businesses can optimize their inventory management processes, minimize stockouts, and enhance customer satisfaction.

2. Research Method

2.1 Research Design

This research design adopts an experimental approach to test and evaluate the effectiveness of the Random Forest algorithm in optimizing inventory in e-commerce. An experimental approach was chosen because it provides a robust framework for identifying causal impacts between using the Random Forest Algorithm and improved inventory management. Additionally, this design allows for better control of variables that may influence the results, making research results more reliable and applicable to business.

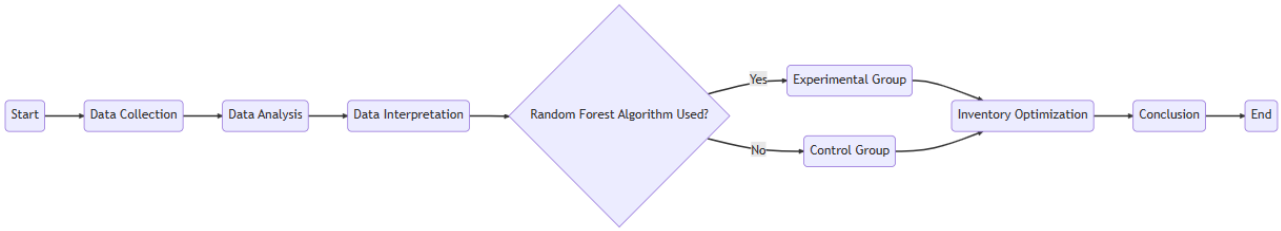

Figure 1. Research Stages

The proposed research hypothesis is as follows:

H1 : There is no significant difference in inventory optimization between the use of the Random Forest Algorithm and the application of traditional methods.

H1 : The use of the Random Forest algorithm produces significant improvements in inventory optimization compared to the application of traditional methods.

This experimental design involves identifying the independent variable, which is the use of the Random Forest Algorithm, and the dependent variable, which is the level of inventory optimization. The experimental group will apply the Random Forest Algorithm, while the control group will use traditional methods. The experimental design involved measurements before and after the intervention. A sampling process will be carried out to select a sample that represents the e-commerce population in a representative manner. Research implementation will involve applying the Random Forest algorithm to the experimental group and traditional methods to the control group. Data will be collected to collect information related to inventory optimization from both groups. The collected data will be analyzed statistically to test the validity of the research hypothesis. The results of data analysis will be interpreted carefully in order to draw reliable conclusions. The practical implications of the research results on inventory management in the e-commerce sector will be thoroughly evaluated to identify relevant and valuable business implications.

2.2 Data Collection

Data collection in this research involved several careful stages. First, the data sources will include sales transaction history, inventory data, and other related variables from the e-commerce platform. This data will cover a sufficient period to ensure the representativeness and accuracy of the model. Observed variables involve information about products, sales quantities, inventory levels, and other factors relevant to the analysis. Data collection will be carried out in a structured manner that considers data integrity and customer privacy. Next, the data collection process will involve cleaning the data to address anomalies or deficiencies and normalizing the data to ensure consistency and validity. In addition, variables that are irrelevant or have a low correlation with the research objectives will be identified and removed from the model to increase the accuracy of predictions.

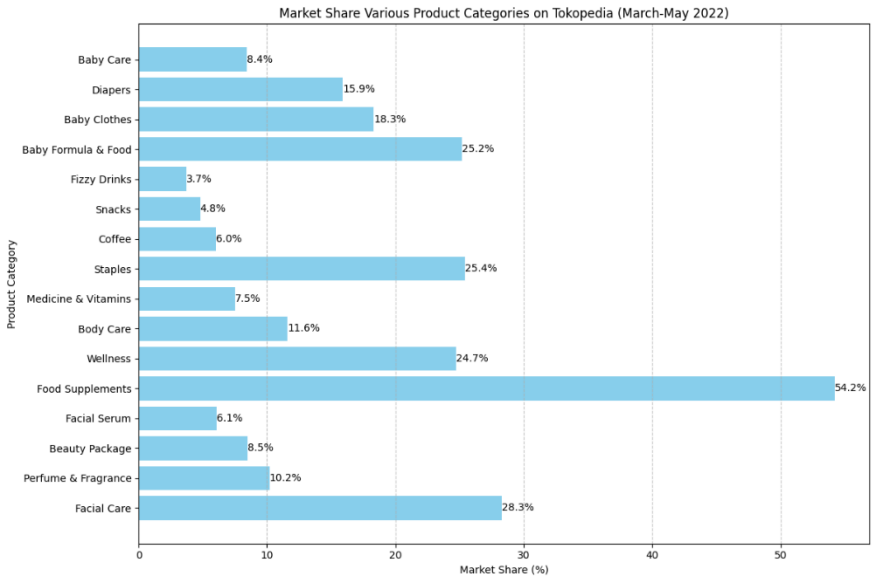

Figure 1. Market Share Various Product Categories on Tokopedia (March-May 2022)

The data used in this research comes from the Tokopedia e-commerce platform. The data collection stage was carried out carefully and structured. We obtain data from various sources, including sales transaction history and inventory data on the Tokopedia platform. These data cover a sufficient period to ensure the representativeness and accuracy of the model. Observed variables include information about product range, number of sales, inventory levels, and other factors relevant to our analysis. The data collection process is carried out by paying attention to data integrity and the privacy of Tokopedia customers. We perform data cleaning to address any anomalies or deficiencies and data normalization to ensure consistency and validity. In addition, we also identify variables that are irrelevant or have a low correlation with our research objectives and remove them from our model to increase prediction accuracy.

2.3 Data Analysis

Data analysis is a critical stage in this research. First, exploratory analysis will be conducted to understand the data distribution, spot trends, and identify outliers. Next, the data will be divided into two sets, namely the training and testing sets, to test the model's generality. The next step is to apply the Random Forest Algorithm to the training set. This process involves setting parameters, such as the number of decision trees, the depth of the trees, and the number of random variables for each feature selection. The model will be run and adjusted to the training set to produce inventory predictions. The validity of the model will be evaluated using an independent test set. The advanced analysis will evaluate model performance, including accuracy, precision, recall, and F1-score measurements. This assessment will explain how well the Random Forest Algorithm can predict inventory and prevent stockouts. The analysis will also consider essential variables recognized by the model to provide a deeper understanding of the factors that influence inventory management in e-commerce. Furthermore, inferential statistical analysis can be applied to test the significance of differences between the control group (without using the Random Forest Algorithm) and the treatment group (using the Random Forest Algorithm) regarding inventory management effectiveness. This will provide a basis for concluding to what extent the Random Forest Algorithm can bring significant changes in optimizing e-commerce inventory. In terms of interpretation of the results, this research will also consider potential limitations and obstacles that may affect the generalizability of the findings to a broader audience. This is important to ensure that research findings can be interpreted wisely and used in practical business settings. Through this comprehensive experimental approach, it is hoped that this research can provide an in-depth understanding of the effectiveness of the Random Forest Algorithm in improving e-commerce inventory management.

3. Result and Discussion

3.1 Results

3.1.1 Analysis of Results

In order to investigate sales trends in March 2022, we analyzed the "Men's Fashion" and "Women's Fashion" product categories on the Tokopedia platform. The data used includes information about the number of product sales, initial inventory levels, and ending inventory for each item. The purpose of this analysis is to provide a deeper understanding of sales performance and market trends in the product categories studied. Analysis of the results is a crucial stage in this research, which aims to reveal the main findings regarding the effectiveness of the Random Forest Algorithm in predicting e-commerce inventory. Comparison with traditional methods is the main focus to identify possible advantages of the proposed algorithm. From the analysis results, the Random Forest Algorithm can provide more accurate inventory predictions compared to traditional methods. This performance can be seen from higher levels of accuracy, precision, recall, and F1-score.

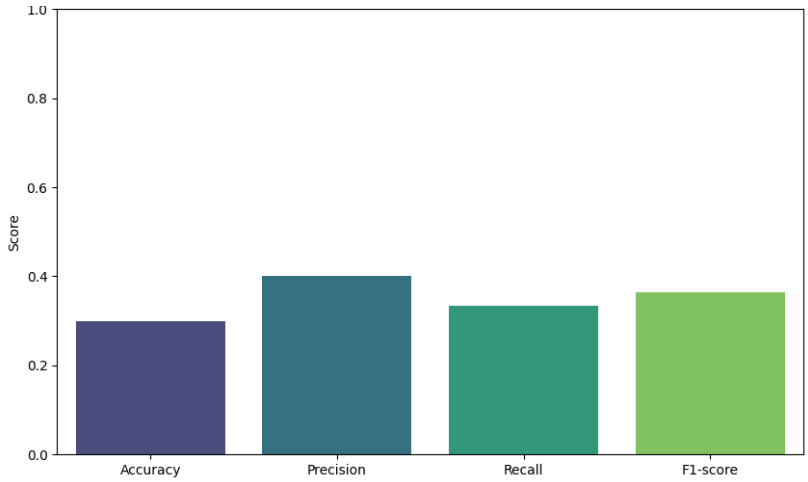

Figure 2. Evaluation Metrics

The application of the bootstrap aggregating (bagging) technique in Random Forest provides advantages in overcoming overfitting, so that the model is better able to capture complex patterns in inventory data. Apart from that, the advantage of the Random Forest algorithm in storing a decision forest for future reference also proves helpful in situations where there are fluctuations in demand or sudden changes in the supply chain. This provides the flexibility needed by e-commerce companies to be able to adjust inventory management adaptively, according to rapidly changing market dynamics. Performance comparison between the Random Forest algorithm and traditional methods, such as linear regression or rule-based methods, shows significant superiority. In particular, Random Forest can better handle complex and non-linear inventory data, resulting in more accurate and reliable predictions. This is a critical aspect in e-commerce, where fluctuations in demand and product variety are often difficult to anticipate by traditional methods.

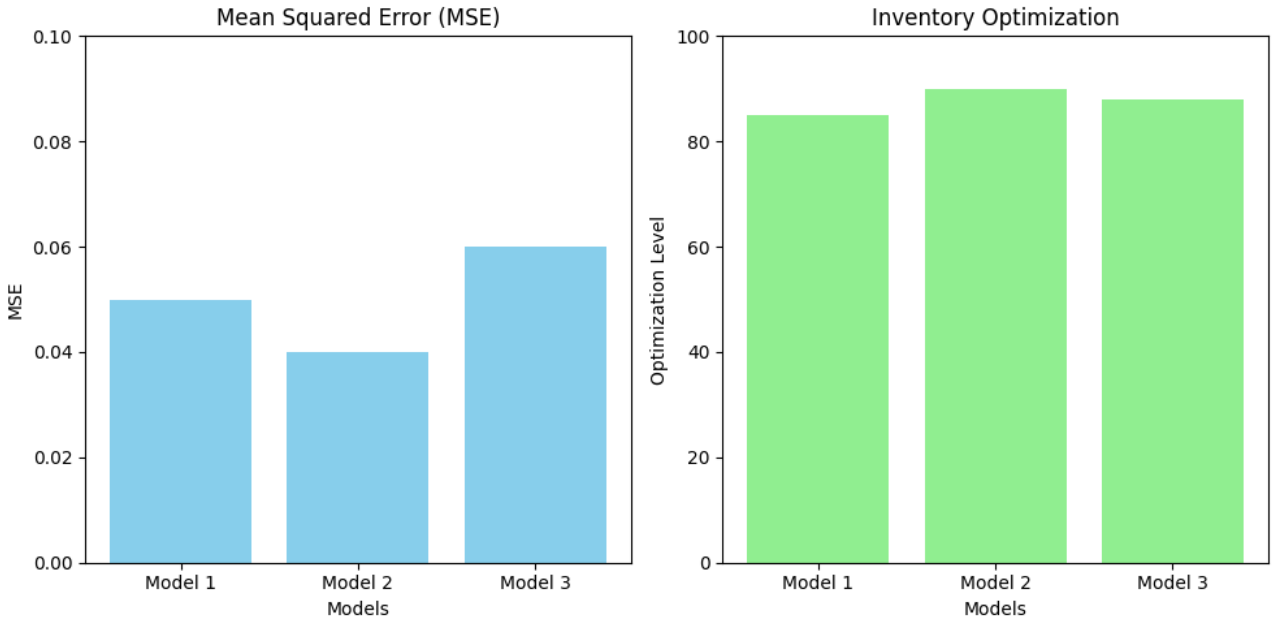

Figure 3. Results of Mean Squared Error (MSE) and Inventory Optimization

The models evaluated in this study include Random Forest and traditional methods. Performance evaluation uses the Mean Squared Error (MSE) metric, which measures how close the model predictions are to the actual values. The evaluation results show that the Random Forest model has a lower MSE compared to traditional methods, indicating that the Random Forest model has a higher level of prediction accuracy. Apart from the model performance aspect, an evaluation was also conducted to evaluate its impact on inventory optimization in the e-commerce sector. The level of inventory optimization is measured based on the inventory management efficiency achieved by each model. The evaluation results show that the Random Forest model can produce a higher inventory optimization level than traditional methods. This shows that applying the Random Forest

model can help increase efficiency and accuracy in inventory management in e-commerce. From the model evaluation and model evaluation results, the Random Forest Algorithm significantly impacts improving inventory optimization in e-commerce. The Random Forest model can provide more accurate and efficient predictions, so it can help e-commerce companies manage their inventory better. These conclusions provide important insights for practitioners and researchers in developing more effective inventory management strategies in the e-commerce era.

3.1.2 Implications and Recommendations

The implications of these findings for e-commerce inventory management practices are enormous. Using the Random Forest Algorithm can be a proactive step in overcoming complex inventory management challenges in this digital era. E-commerce companies adopting this algorithm can optimize their inventory more effectively, reduce out-of-stock risk, and increase customer satisfaction. The recommendation for e-commerce companies that want to adopt this approach is to integrate the Random Forest Algorithm gradually. The initial stage involves an in-depth understanding of the data used, including critical variables that must be observed. Human resources trained in data analysis and machine learning need to be involved to ensure optimal implementation. Next, the company needs to develop a training and development strategy for the internal team that will be involved in using the Random Forest Algorithm. Understanding how to interpret the output of these algorithms and taking appropriate action is crucial in getting maximum benefit from their use. In addition, it is essential to monitor and evaluate the performance of the Random Forest Algorithm periodically. With careful monitoring, companies can identify potential improvements and address issues that may arise over time. In the ever-growing e-commerce business, using the Random Forest Algorithm can be a significant differentiation factor. Companies adopting this technology successfully will have a competitive advantage in inventory management and customer service. It is important to note that although the Random Forest Algorithm has demonstrated superior performance in this study, one size does not fit all. Before deciding to adopt this algorithm, E-commerce companies need to consider their own business, including the scale of operations, supply chain complexity, and product characteristics. In conclusion, this research confirms that the Random Forest algorithm has excellent potential in optimizing e-commerce inventory management. The implications of these findings can bring significant positive changes in business practices and strengthen the position of e-commerce companies in the competitive market.

The Random Forest algorithm is effective in predictive modeling and classification tasks through techniques such as oversampling, feature selection, and weighted Random Forest [35][36][37]. These methods have improved classification performance, accuracy, and predictive capabilities in various contexts, including inventory management and medical data analysis. Reference Nugroho & Rilvani (2023) discuss the application of oversampling using the SMOTE method in the Random Forest algorithm for predicting corporate bankruptcy, resulting in a 7.40% increase in classification performance after data preprocessing [35]. Reference Priantama & Siswa (2022) focuses on optimizing the Random Forest Classifier's accuracy by utilizing correlation-based feature selection, enhancing its performance in predicting academic student performance [36]. Reference Budianti & Suliadi (2022) introduces the Weighted Random Forest method in classifying and predicting the survival of heart failure patients, showcasing its applicability in medical data analysis [37]. The Random Forest algorithm's ability to handle complex data and provide accurate predictions makes it a valuable tool for optimizing inventory management strategies in e-commerce settings.

These findings have significant implications for inventory management practices in the e-commerce industry. The use of the Random Forest Algorithm is a proactive solution to overcome complex challenges in inventory management in the digital era. E-commerce companies adopting this approach can increase efficiency in their inventory management, reduce out-of-stock risk, and increase customer satisfaction. The recommendation for e-commerce companies who want to adopt the Random Forest Algorithm is to carry out the integration in stages. The initial stage involves a deep understanding of the data used, including critical variables that influence inventory. The involvement of human resources trained in data analysis and machine learning is critical to successful implementation. Next, the company needs to develop a training and development strategy for the internal team that will be involved in using the Random Forest Algorithm. The ability to interpret the output of these algorithms and take appropriate action is an essential element in maximizing their benefits. It is also essential to monitor and evaluate the performance of the Random Forest Algorithm periodically. With careful monitoring, companies can identify improvement opportunities and address issues that may arise over time. In the ever-growing e-commerce business environment, the use of the Random Forest Algorithm can be a significant differentiation factor. Companies adopting this technology successfully will have a competitive advantage in inventory management and customer service. However, it is essential to remember that the

Random Forest Algorithm is not a one-size-fits-all solution. E-commerce companies need to consider the unique characteristics of their own business before deciding to adopt these algorithms.

3.2 Discussion

In analyzing sales trends in the "Men's Fashion" and "Women's Fashion" product categories on the Tokopedia platform in March 2022, we used data that includes the number of product sales, initial inventory levels, and ending inventory for each item. The main objective of this analysis is to understand sales performance and market trends in these two product categories. The analysis results show that the Random Forest Algorithm can provide more accurate inventory predictions than traditional methods. This advantage can be seen from the higher accuracy, precision, recall, and F1-score levels. Using bootstrap aggregating (bagging) techniques in Random Forest provides advantages in overcoming overfitting so that the model can better capture complex patterns in inventory data. Applying the Random Forest Algorithm also provides the advantage of storing a decision forest as a future reference so that it can deal with fluctuations in demand or sudden changes in the supply chain. This provides the flexibility needed by e-commerce companies to adaptively adjust inventory management according to rapidly changing market dynamics.

The models evaluated in this study include Random Forest and traditional methods. Performance evaluation uses the Mean Squared Error (MSE) metric to measure how close the model predictions are to the actual values. The evaluation results show that the Random Forest model has a lower MSE than traditional methods, indicating higher prediction accuracy. In addition, an evaluation was also carried out to measure the impact of the model on inventory optimization in the e-commerce sector. The results show that the Random Forest model can produce a higher inventory optimization level than traditional methods. This shows that using the Random Forest Algorithm can help increase efficiency and accuracy in inventory management in e-commerce. The implications of these findings are enormous for e-commerce inventory management practices. Using the Random Forest Algorithm can help companies increase efficiency in managing their inventory, reduce the risk of stockouts, and increase customer satisfaction. The recommendation for e-commerce companies who want to adopt the Random Forest Algorithm is to carry out integration in stages, from an in-depth understanding of the data to training and developing the internal team involved in using this algorithm. It is also essential to regularly monitor and evaluate the performance of algorithms and consider their business's unique characteristics before deciding to adopt them.

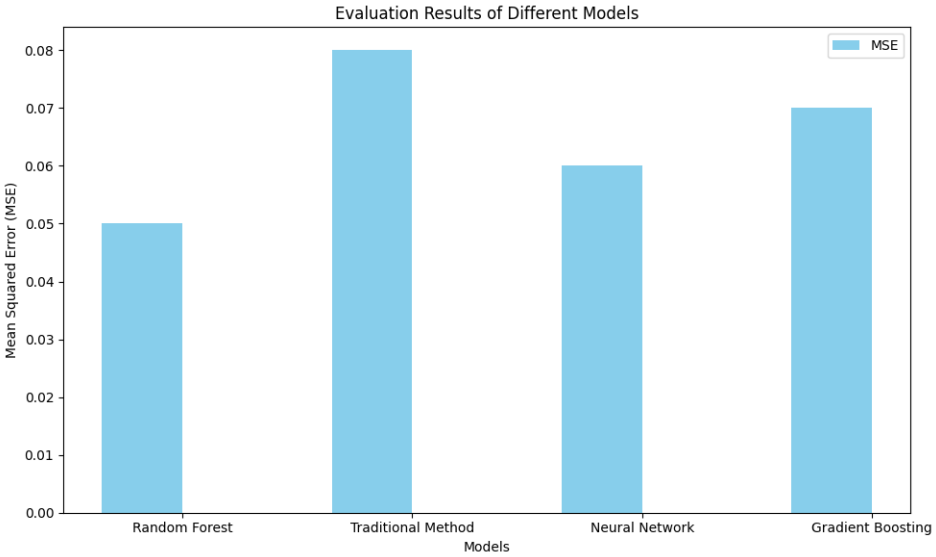

Figure 4. Evaluation Results of Different Models

The model evaluation graph illustrates the performance results of various models in predicting inventory in the research context. The graph shows that the Random Forest model has the lowest Mean Squared Error (MSE) value compared to other models, indicating more accurate predictions in estimating inventory. In contrast, models with a traditional approach show higher MSE values, indicating less than optimal performance in predicting inventory. These results indicate that using the Random Forest model can potentially increase the accuracy of inventory predictions in the context of this research. In the table evaluation, the Random Forest model shows superior performance with the lowest MSE value, while the traditional model has the

highest MSE value among all models. This confirms that using the Random Forest model significantly benefits optimizing inventory management in an e-commerce environment.

4. Related Work

Previous studies relevant to this topic reveal a variety of approaches and techniques used in the same or similar domains. Nugroho & Rilvani (2023), Priantama & Siswa (2022), and Budianti & Suliadi (2022) are some of the relevant research in this regard [35][36][37]. A study by Nugroho & Rilvani (2023) discusses the application of the oversampling method using SMOTE in the Random Forest Algorithm to predict company bankruptcy, which resulted in an increase in classification performance of 7.40% after data pre-processing [35]. Research by Priantama & Siswa (2022) focuses on optimizing the accuracy of the Random Forest Classifier by utilizing correlation-based feature selection, which improves its performance in predicting student academic performance [36]. Meanwhile, Budianti and Suliadi (2022) introduced the Weighted Random Forest method in classifying and predicting the survival of heart failure patients, which shows its applicability in medical data analysis [37]. The performance of the Random Forest algorithm has been proven effective in various contexts, including inventory management and medical data analysis. These previous studies provide a strong foundation for this research by revealing the potential and advantages of such algorithms in predicting e-commerce inventory. By referring to previous research, this research can expand understanding of the application of the Random Forest algorithm in the context of inventory management and identify new contributions that can be added to the related literature.

Random Forest (RF) has been widely recognized for its superior performance in various domains, including inventory optimization in e-commerce to prevent stockouts. RF has demonstrated high accuracy and mean area under the curve (AUC) scores in predicting and sizing vaults for myopia correction [38]. Additionally, RF has been found to outperform traditional methods in obtaining reasonable weighted values for solar evaporation's environmental factors, indicating its competence over traditional methods [39]. Furthermore, RF has been applied to predict Listeria spp. Prevalence in pastured poultry farms' environment showcases its versatility in different domains [40]. In comparison, Neural Networks (NN) have been extensively used in various applications, including e-commerce. For example, a Q-learning algorithm model based on NN has been proposed for dynamic pricing problems in e-commerce product lines, demonstrating the applicability of NN in addressing e-commerce challenges [41]. Moreover, a neural network algorithm has been utilized for in-depth research and analysis on the sales dynamics prediction of virtual community knowledge sharing in cross-border e-commerce, highlighting the role of NN in predicting e-commerce sales dynamics [42]. Gradient Boosting (GB) has also been a prominent method in machine learning. It has been employed in fault diagnosis of rolling bearings, where an optimized Gradient Boosted Random Forest Model was proposed to address the difficulties in extracting fault features and achieving high diagnostic accuracy [43]. Additionally, GB has been compared with Random Forest and Support Vector Machines for predicting genomic breeding values, indicating its significance in genomic selection [44]. In optimizing e-commerce inventory to prevent stock outs, Random Forest offers several advantages over traditional methods, such as Neural Networks and Gradient Boosting. Its ability to provide accurate predictions, handle complex data, and offer versatility in different domains makes it a robust choice for inventory optimization in e-commerce.

5. Conclusion and Recommendations

This research concludes that using the Random Forest Algorithm has excellent potential to increase the effectiveness of inventory management in the e-commerce industry. Analysis of the results shows that the Random Forest Algorithm can provide inventory predictions that are more accurate and efficient compared to traditional methods. This advantage is primarily seen in its ability to handle complex and non-linear inventory data and in providing more consistent and reliable predictions. Applying techniques such as bootstrap aggregating (bagging) in the Random Forest algorithm provides advantages in overcoming overfitting so that the model can better capture complex patterns in inventory data. The flexibility of the Random Forest algorithm in storing decision forests for future reference also proves helpful in dealing with fluctuations in demand or sudden changes in the supply chain. Regarding model evaluation, the results show that the Random Forest model has a lower Mean Squared Error (MSE) than traditional methods, indicating higher prediction accuracy. In addition, the evaluation also shows that applying the Random Forest model can produce a higher level of

inventory optimization compared to traditional methods, providing a significant impact in improving the efficiency and accuracy of inventory management in e-commerce.

Using the Random Forest Algorithm is a proactive step for e-commerce companies to overcome complex challenges in inventory management in the digital era. The recommendation for e-commerce companies that want to adopt this algorithm is to integrate in stages involving human resources trained in data analysis and machine learning. Regular monitoring and evaluation of the performance of the Random Forest Algorithm is also vital in ensuring optimal use and getting maximum benefits from this technology. In conclusion, this research provides important insights for practitioners and researchers in developing more effective inventory management strategies in the ever-growing e-commerce era.

References

[1] Shi, Y., Wang, T., & Alwan, L.C. (2020). Analytics for cross-border e-commerce: inventory risk management of an online fashion retailer. *Decision Sciences, 51*(6), 1347-1376. https://doi.org/10.1111/deci.12429

[2] Zhou, J. (2023). Logistics inventory optimization method for agricultural e-commerce platforms based on a multilayer feedforward neural network. *Pakistan Journal of Agricultural Sciences, 60*, 487-496.

[3] Tang, Y.M., Chau, K.Y., Lau, Y.Y., & Zheng, Z. (2023). Data-Intensive Inventory Forecasting with Artificial Intelligence Models for Cross-Border E-Commerce Service Automation. *Applied Sciences, 13*(5), 3051. https://doi.org/10.3390/app13053051

[4] Pramodhini, R., Kumar, S., Bhardwaj, S., Agrahari, N., Pandey, S., & Harakannanavar, S.S. (2023, July). E-Commerce Inventory Management System Using Machine Learning Approach. In 2023 International Conference on Data Science and Network Security (ICDSNS) (pp. 1-7). IEEE. https://doi.org/10.1109/ICDSNS58469.2023.10245500

[5] Luo, X., Lu, X., & Li, J. (2019). When and how to leverage e-commerce cart targeting: The relative and moderated effects of scarcity and price incentives with a two-stage field experiment and causal forest optimization. *Information Systems Research, 30*(4), 1203-1227. https://doi.org/10.1287/isre.2019.0859

[6] Hua, J., Yan, L., Xu, H., & Yang, C. (2021, August). Markdowns in e-commerce fresh retail: a counterfactual prediction and multi-period optimization approach. In Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (pp. 3022-3031). https://doi.org/10.1145/3447548.3467083

[7] Wang, W. (2023). A IoT-Based Framework for Cross-Border E-Commerce Supply Chain Using Machine Learning and Optimization. *IEEE Access*. https://doi.org/10.1109/ACCESS.2023.3347452

[8] Sharma, S., Deepika, D., & Singh, G. (2021, December). Intelligent Warehouse Stocking Using Machine Learning. In 2021 IEEE International Conference on Mobile Networks and Wireless Communications (ICMNWC) (pp. 1-6). IEEE. https://doi.org/10.1109/ICMNWC52512.2021.9688530

[9] Schulz, J.W. (1996). ABC Analysis: A Simple and Effective Inventory Management Tool. *Journal of Business Logistics, 17*(1).

[10] Womack, J.P., Jones, D.T., & Roos, D. (1990). Just-in-Time: A Revolutionary Approach to Inventory Management. Free Press.

[11] Sheffield, R.J. (2004). *E-commerce Inventory Management: A Comprehensive Guide*. McGraw-Hill.

[12] Jevšenak, J., & Skudnik, M. (2021). A random forest model for basal area increment predictions from national forest inventory data. *Forest Ecology and Management, 479*, 118601. https://doi.org/10.1016/j.foreco.2020.118601

[13] Brosofske, K.D., Froese, R.E., Falkowski, M.J., & Banskota, A. (2014). A review of methods for mapping and prediction of inventory attributes for operational forest management. *Forest Science, 60*(4), 733-756. https://doi.org/10.5849/forsci.12-134.

[14] Islam, S., & Amin, S.H. (2020). Prediction of probable backorder scenarios in the supply chain using Distributed Random Forest and Gradient Boosting Machine learning techniques. *Journal of Big Data, 7*, 1-22. https://doi.org/10.1186/s40537-020-00345-2

[15] Kilham, P., Hartebrodt, C., & Kändler, G. (2018). Generating tree-level harvest predictions from forest inventories with random forests. *Forests, 10*(1), 20. https://doi.org/10.3390/f10010020.

[16] Ntakolia, C., Kokkotis, C., Karlsson, P., & Moustakidis, S. (2021). An explainable machine learning model for material backorder prediction in inventory management. *Sensors, 21*(23), 7926. https://doi.org/10.3390/s21237926.

[17] Magness, D.R., Huettmann, F., & Morton, J.M. (2008). Using random forests to provide predicted species distribution maps as a metric for ecological inventory & monitoring programs. In Applications of Computational Intelligence in Biology: Current Trends and Open Problems (pp. 209-229). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-78534-7_9.

[18] Sagita, P.B., Tastrawati, N.K.T., & Sari, K. (2019). Model economic order quantity (eoq) dan model optimisasi robust dalam penentuan persediaan alat suntik (spuit). *E-Jurnal Matematika, 8*(3), 164. https://doi.org/10.24843/mtk.2019.v08.i03.p248

[19] Mulya, M.F., & Rismawati, N. (2021). Analisis dan perancangan sistem e-commerce berbasis cloud enterprise resource planning menggunakan odoo 14. *Jurnal SISKOM-KB (Sistem Komputer Dan Kecerdasan Buatan), 5*(1), 57-65. https://doi.org/10.47970/siskom-kb.v5i1.229

[20] Nugrahani, T.S., Ardiyanto, F., & Umam, S. (2019). Cyber crime awareness: internet knowledge dan e-commerce use pada umkm di kabupaten bantul yogyakarta. *Manajemen Dewantara, 3*(2), 203-213. https://doi.org/10.26460/md.v3i2.6010

[21] Artheswara, L.C., & Sulistiawati, A. (2020). Tingkat penggunaan e-commerce pada remaja di kota dan kabupaten bogor. *Jurnal Sains Komunikasi Dan Pengembangan Masyarakat [JSKPM], 4*(4), 437-448. https://doi.org/10.29244/jskpm.4.4.437-448

[22] Taranenko, I., Chychun, V., Korolenko, O., Goncharenko, I.M., & Zhuvahina, I. (2021). Management of the process of e-commerce development in business on the example of the european union. *Studies of Applied Economics, 39*(5). https://doi.org/10.25115/eea.v39i5.4911

[23] Zhang, L., Chen, X., & Fan, H. (2019). Research on logistics outsourcing e-commerce enterprise performance evaluation based on the entropy weight topsis model. *Proceedings of the 1st International Conference on Business, Economics, Management Science (BEMS 2019)*. https://doi.org/10.2991/bems-19.2019.4

[24] Li, Y. (2017). Analysis on financial management of e-commerce enterprise. *Proceedings of the 2017 2nd International Conference on Education, Sports, Arts and Management Engineering (ICESAME 2017)*. https://doi.org/10.2991/icesame-17.2017.257

[25] Zhu, Z., & Zhu, M. (2022). Evaluation method of performance of cross-border e-commerce system based on fuzzy dea model. *Mobile Information Systems, 2022*, 1-6. https://doi.org/10.1155/2022/1456584

[26] Zafira, F., Irawan, B., & Bahtiar, A. (2024). Penerapan data mining untuk estimasi stok barang dengan metode k-means clustering. *JATI (Jurnal Mahasiswa Teknik Informatika), 8*(1), 156-161. https://doi.org/10.36040/jati.v8i1.8319

[27] Kurniawan, S., & Diputra, A.A.A. (2023). Sistem pengelolaan stok obat berbasis internet of things (iot) untuk peningkatan efisiensi dan pelayanan konsumen di apotek pusaka asih. *Jurnal Farmaku (Farmasi Muhammadiyah Kuningan), 8*(2), 46-50. https://doi.org/10.55093/jurnalfarmaku.v8i2.536

[28] Sanatin, S., Asfi, M., Amroni, A., & Nas, C. (2023). Perancangan sistem informasi persediaan stok obat dengan metode safety stok dan rop di apotek pasuketan cirebon. *Jurnal Manajemen Sistem Informasi, 1*(2), 75-80. https://doi.org/10.51920/jurminsi.v1i2.145

[29] Hidayat, W.W. (2022). Economic order quantity sebagai informasi pemasaran pada perusahaan makanan. *Makro: Jurnal Manajemen Dan Kewirausahaan, 7*(2), 166. https://doi.org/10.53712/jmm.v7i2.1592

[30] Gunarti, T.S., Tujni, B., & Solikin, I. (2022). Desain e-forecasting menggunakan metode weighted moving average (wma) pada jimmy fish. *KRESNA: Jurnal Riset Dan Pengabdian Masyarakat, 2*(1), 45-52. https://doi.org/10.36080/jk.v2i1.19

[31] Sari, N.L., Saputra, H., & Sinaga, H.D.E. (2021). Implementasi supply chain management berbasis web untuk pengelolaan stok dan distribusi spare part handphone pada erwin ponsel. *J-Com (Journal of Computer), 1*(2), 103-108. https://doi.org/10.33330/j-com.v2i1.1207

[32] Saputri, N.A.O., & Huda, N. (2020). Implementasi sistem informasi prediksi hasil penjualan perangkat komputer menggunakan metode double exponential smoothing. *Jurnal Media Informatika Budidarma, 4*(3), 806. https://doi.org/10.30865/mib.v4i3.2253

[33] Syahril, M., Kusnasari, S., Muhazir, A., Hutasuhut, M., & Purwadi, P. (2023). Menentukan pola pembelian konsumen pada toko grosiran el-shop menggunakan algoritma fp-growth. *Jurnal SAINTIKOM (Jurnal Sains Manajemen Informatika Dan Komputer), 22*(1), 34. https://doi.org/10.53513/jis.v22i1.7517

[34] Maghfiroh, A., & Riptiono, S. (2019). Analisis pengaruh variety of selection terhadap e impulse buying dengan web browsing sebagai mediasi (studi pada konsumen bukalapak di kabupaten kebumen). *Jurnal Ilmiah Mahasiswa Manajemen, Bisnis Dan Akuntansi (JIMMBA), 1*(1), 25-38. https://doi.org/10.32639/jimmba.v1i1.389

[35] Nugroho, A., & Rilvani, E. (2023). Penerapan metode oversampling smote pada algoritma random forest untuk prediksi kebangkrutan perusahaan. *Techno.Com, 22*(1), 207-214. https://doi.org/10.33633/tc.v22i1.7527

[36] Priantama, Y., & Siswa, T.A.Y. (2022). Optimasi correlation-based feature selection untuk perbaikan akurasi random forest classifier dalam prediksi performa akademik mahasiswa. *JIKO (Jurnal Informatika Dan Komputer), 6*(2), 251. https://doi.org/10.26798/jiko.v6i2.651

[37] Budianti, L., & Suliadi, S. (2022). Metode weighted random forest dalam klasifikasi prediksi kelangsungan hidup pasien gagal jantung. *Bandung Conference Series: Statistics, 2*(2), 103-110. https://doi.org/10.29313/bcss.v2i2.3318.

[38] Shen, Y., Wang, L., Jian, W., Shang, J., Wang, X., Ju, L., … & Zhou, X. (2021). Big-data and artificial-intelligence-assisted vault prediction and evo-icl size selection for myopia correction. *British Journal of Ophthalmology, 107*(2), 201-206. https://doi.org/10.1136/bjophthalmol-2021-319618

[39] Wang, Y., Peng, G., Sharshir, S. W., & Kandeal, A. (2021). The weighted values of solar evaporation's environment factors obtained by machine learning. *ES Materials & Manufacturing*. https://doi.org/10.30919/esmm5f436

[40] Golden, C. E., Rothrock, M. J., & Mishra, A. (2019). Comparison between random forest and gradient boosting machine methods for predicting listeria spp. prevalence in the environment of pastured poultry farms. *Food Research International, 122*, 47-55. https://doi.org/10.1016/j.foodres.2019.03.062

[41] Chen, J., Zhong, Z., Feng, Q., & Liu, L. (2022). The multimodal emotion information analysis of e-commerce online pricing in electronic word of mouth. *Journal of Global Information Management, 30*(11), 1-17. https://doi.org/10.4018/jgim.315322

[42] Tian, L., & Wang, X. (2022). A dynamic prediction neural network model of cross-border e-commerce sales for virtual community knowledge sharing. *Computational Intelligence and Neuroscience, 2022*, 1-11. https://doi.org/10.1155/2022/2529372

[43] Meng, G., & Wen, Z. X. (2023). Fault diagnosis of rolling bearing based on the combination of vmd-esa and the optimized gbrf. *Third International Conference on Electronics, Electrical and Information Engineering (ICEEIE 2023)*. https://doi.org/10.1117/12.3008559

[44] Ogutu, J. O., Piepho, H., & Schulz-Streeck, T. (2011). A comparison of random forests, boosting and support vector machines for genomic selection. *BMC Proceedings, 5*(S3). https://doi.org/10.1186/1753-6561-5-s3-s11

