logistics

Article

Enhancing Intermittent Spare Part Demand Forecasting: A Novel Ensemble Approach with Focal Loss and SMOTE

Saskia Puspa Kenaka [1,2,3,*], Andi Cakravastia [1,2,3], Anas Ma'ruf [1,3,4] and Rully Tri Cahyono [1,2,3]

[1] Industrial Engineering Program, Faculty of Industrial Technology, Bandung Institute of Technology, Bandung 40132, Indonesia; andi@itb.ac.id (A.C.); maruf@itb.ac.id (A.M.); r.t.cahyono@itb.ac.id (R.T.C.)
[2] Research Group of Industrial System and Techno-Economics, Bandung Institute of Technology, Bandung 40132, Indonesia
[3] Center for Logistics and Supply Chain Studies, Bandung Institute of Technology, Bandung 40132, Indonesia
[4] Research Group of Manufacturing System, Bandung Institute of Technology, Bandung 40132, Indonesia
* Correspondence: saskia@itb.ac.id

Abstract: *Background*: Accurate inventory management of intermittent spare parts requires precise demand forecasting. The sporadic and irregular nature of demand, characterized by long intervals between occurrences, results in a significant data imbalance, where demand events are vastly outnumbered by zero-demand periods. This challenge has been largely overlooked in forecasting research for intermittent spare parts. *Methods*: The proposed model incorporates the Synthetic Minority Oversampling Technique (SMOTE) to balance the dataset and uses focal loss to enhance the sensitivity of deep learning models to rare demand events. The approach was empirically validated by comparing the model's Mean Squared Error (MSE) performance and Area Under the Curve (AUC). *Results*: The ensemble model achieved a 47% reduction in MSE and a 32% increase in AUC, demonstrating substantial improvements in forecasting accuracy. *Conclusions*: The findings highlight the effectiveness of the proposed method in addressing data imbalance and improving the prediction of intermittent spare part demand, providing a valuable tool for inventory management.

Keywords: intermittent demand; spare part; ensemble learning; SMOTE; loss function

check for updates

Academic Editor: Robert Handfield

Received: 27 December 2024
Revised: 5 February 2025
Accepted: 7 February 2025
Published: 8 February 2025

Citation: Kenaka, S.P.; Cakravastia, A.; Ma'ruf, A.; Cahyono, R.T. Enhancing Intermittent Spare Part Demand Forecasting: A Novel Ensemble Approach with Focal Loss and SMOTE. *Logistics* **2025**, *9*, 25. https://doi.org/10.3390/logistics9010025

1. Introduction

Spare parts are important components in various industrial sectors, ensuring the continuity of machine operation and minimizing costly downtime [1,2]. The periodic replacement of spare parts is unavoidable throughout the life cycle of machines supporting industrial operations. For machines with a service life of up to 30 years, annual spare part expenditure can reach up to 2.5% of the initial cost of the machine, indicating that spare part replacement is unavoidable [3].

Ineffective spare part management can result in overstocking, tying up substantial financial resources, or understocking, leading to unplanned downtime and emergency procurement at premium costs. For instance, in the petrochemicals industry, 54% of inventory comprises high-value spare parts with no recent demand [4], while the US Department of Defense holds $1.5 billion of unused spare parts [5]. Conversely, unplanned downtime in the oil and gas industry can result in annual losses of up to $88 million [3]. Despite their infrequent use, they account for a large portion of the total SKUs and represent significant asset value in inventory [6]. Errors in forecasting demand for these categories amplify operational and financial risks, highlighting the need for effective forecasting techniques [7,8].

Forecasting intermittent and slow-moving spare parts is particularly challenging due to their sporadic demand, which is dominated by long periods of zero demand interspersed with sudden, irregular spikes. This data imbalance complicates forecasting models' ability to accurately identify patterns and predict demand. Moreover, the infrequent yet critical nature of such demand amplifies the consequences of forecasting errors, resulting in overstocking or costly stockouts. These parts were selected as the focus of this study because they represent a significant challenge for forecasting and are critical for minimizing downtime and optimizing inventory management, especially in industries such as oil and gas.

Conventional forecasting methods, such as parametric approaches like Croston's method [9] and its extensions, attempt to handle the sporadic nature of demand by separating the forecasting of demand size and frequency. However, these methods rely on distribution-based assumptions that are often difficult to verify, particularly for datasets with scarce event occurrences, where long intervals of zero demand dominate. Moreover, maximum likelihood estimation, usually employed in these models, becomes less effective in scenarios with limited data, making it even more challenging to accurately predict rare but critical demand occurrences [10].

These limitations have prompted researchers to explore alternative approaches, particularly non-parametric methods that relax distributional assumptions and focus on identifying patterns directly from data. For example, support vector machines (SVMs), Logistic Regression [11], and neural networks [12,13] have been employed to predict demand occurrences. Ensemble methods have also been explored to enhance prediction accuracy by combining multiple models [7,14,15]. However, most machine learning studies focus on forecasting demand occurrence without addressing the interval or size of demand. Additionally, these models often overlook the severe imbalance between periods of demand and non-demand, which remains critical for accurately predicting rare but important demand events.

This study proposes a novel ensemble forecasting approach integrating the Synthetic Minority Oversampling Technique (SMOTE) and focal loss to address the challenges of intermittent demand forecasting. SMOTE enhances the representation of rare demand occurrences by generating synthetic samples for the minority class, enabling models to learn patterns in imbalanced datasets better [16]. Focal loss complements SMOTE by modifying the loss function to prioritize rare events during training, ensuring the model focuses on accurately predicting critical but infrequent demand occurrences [8]. This study combines these techniques to address data-level and model-level challenges in forecasting slow-moving and intermittent spare parts.

While SMOTE and focal loss have been widely applied in domains such as healthcare [17–20], finance [21], and fault detection [22,23], their application in intermittent spare part demand forecasting remains underexplored. This study bridges this gap by leveraging these techniques to improve forecasting accuracy for slow-moving and intermittent spare parts. The proposed model is validated using real-world demand data, offering practical insights to optimize inventory management, reduce downtime risks, and enhance cost efficiency.

The remainder of this paper is structured as follows: Section 2 reviews relevant forecasting techniques, including existing approaches for intermittent demand, and explains SMOTE and focal loss. Section 3 outlines the research methodology. Section 4 proposes the ensemble forecasting model. Section 5 presents the results and discussion based on empirical validation. Finally, Section 6 concludes with the findings, limitations, and directions for future research.

2. Literature Review

Spare part management is a crucial step for inventory control. It forms the foundation for determining policies impacting production processes and cost management. In addition, spare part management will also affect the activities supported, namely the production process. A shortage of spare parts can result in prolonged equipment downtime, resulting in significant financial and operational losses for companies [3]. On the other hand, excessive inventory of spare parts, particularly high-value items, ties up substantial financial resources and increases holding costs, leading to inefficiencies in resource allocation [24]. Achieving a balance between spare part availability and cost efficiency remains a critical challenge, emphasizing the need for accurate demand forecasting.

A better understanding of the characteristics of spare parts is achieved through the classification method. Based on the classification of SBC (Syntetos–Boylan classification), it is known that spare parts tend to have intermittent characteristics, namely, many periods that do not have demand [25]. This characteristic is one of the objects in current spare part forecast research with applications in heavy machinery, aircraft, electronics, maritime, and automotive industries [26,27]. As these industries rely heavily on critical spare parts, inaccuracies in demand forecasting can directly impact operational reliability and financial performance.

This study introduces a novel approach to intermittent demand forecasting by integrating the Synthetic Minority Oversampling Technique (SMOTE) and focal loss to address imbalanced data in demand occurrence prediction. The following sections provide a comprehensive review of existing methods for intermittent demand forecasting, the application of SMOTE in handling imbalanced data, and the use of loss functions to improve model performance.

2.1. Intermittent Demand Forecasting

Existing research on intermittent demand forecasting has explored diverse approaches to enhance prediction accuracy, ranging from parametric methods to machine learning-based techniques. Table 1 summarizes recent developments in the field, presenting a chronological progression of methods that highlights steady advancements.

The foundational work on the parametric method began with Croston (1972) [9], who introduced a two-component approach to separately forecast the timing of demand occurrences and the demand size using exponential smoothing. This method was later refined by Boylan and Syntetos (2001) [10] to address biases and overestimation issues. Subsequent studies, including those by Jiang et al. (2019) [28], Babai et al. (2021) [29], and Tian and Wang (2021) [30], expanded the application of parametric methods to various domains, such as electric power materials, automotive industries, and e-commerce. However, the reliance on distributional assumptions in parametric models limits their effectiveness for scarce events or small datasets, where validating these assumptions is often challenging [10].

Table 1. Intermittent demand forecasting literature review.

References	Case Study	Method			
		Par	N-Par		
			N-ML	ML	HM
Croston (1972) [9]	Dataset	√			
Boylen and Syntetos (2001) [10]	Dataset	√			
Willemain et al. (2004) [26]	Industrial Companies		√		
Hua & Zhang (2006) [11]	Petrochemical Enterprise			√	

Table 1. *Cont.*

References	Case Study	Method			
		Par	N-Par		
			N-ML	ML	HM
Synder et al. (2012) [31]	Automobile Parts	√			
Kourentzes (2013) [12]	Dataset			√	
Lolli et al. (2017) [13]	Automotive Industry			√	
Jiang et al. (2019) [28]	Electric Power Material	√			
Babai et al. (2021) [29]	Automotive Industry	√			
Yang et al. (2021) [32]	Medicine	√			
Baisariyev et al. (2021) [33]	Spare Part Airlines		√		
Tian & Wang (2021) [30]	E-commerce	√			
Zhuang et al. (2022) [14]	Automotive Aftermarket			√	√
Ye et al. (2022) [34]	Dataset	√			
Rožanec et al. (2022)	Automotive Equipment Manufacturer			√	√
Ahmadov and Helo (2023) [35]	Online Sales			√	
Chien et al. (2023) [15]	After-Market Component Manufacturer				√
Affonso et al. (2024) [1]	Iron Ore Corporation	√	√		
Wang et al. (2024) [36]	Products sold by Walmart	√			

Par—Parametric, N-Par—Non-Parametric, N-ML—Non-Machine Learning, ML—Machine Learning, HM—Hybrid Model.

Non-parametric methods emerged as an alternative to address these limitations. Non-parametric, non-machine learning methods, such as the Bootstrap technique introduced by Willemain et al. (2004) [26] and refined by Baisariyev et al. (2021) [33], provide greater flexibility by eliminating strict distributional assumptions. Despite their adaptability, these methods remain sensitive to data variability, which can affect their accuracy.

Machine learning represents a significant leap in non-parametric approaches, leveraging advanced algorithms to identify patterns in intermittent demand. Hua and Zhang (2006) [11] applied support vector machines and Logistic Regression, while Kourentzes (2013) [12] and Lolli et al. (2017) [13] utilized neural networks to predict demand occurrence. These approaches demonstrate the potential of machine learning in addressing the complexity of intermittent demand forecasting. However, challenges remain, particularly in handling imbalanced data, a distinctive characteristic of intermittent demand.

Hybrid models, which combine machine learning with other techniques, represent the latest advancements in the field. For instance, Zhuang et al. (2022) [14] and Rožanec et al. (2022) [15] utilized ensemble approaches, including transfer learning and LightGBM, to enhance prediction accuracy. While hybrid methods are robust and effective for complex datasets, their implementation often requires careful tuning to avoid overfitting.

These advancements underscore the evolving landscape of intermittent demand forecasting, where parametric, non-parametric, and hybrid methods continue to address unique challenges in prediction accuracy and data imbalance. Despite extensive advancements, forecasting methods have rarely addressed the unique challenges of intermittent spare part demand. The significant data imbalance in slow-moving and intermittent spare parts presents an opportunity to improve prediction accuracy through tailored solutions. Future research will focus on developing innovative methods to resolve imbalanced data issues, tailoring solutions to the specific characteristics of intermittent demand.

2.2. Synthetic Minority Oversampling Technique (SMOTE)

A significant challenge in forecasting intermittent demand is the imbalance between demand and non-demand periods. Addressing this imbalance at the data level is a widely used strategy to improve model performance [21,37]. One such method is the Synthetic Minority Oversampling Technique (SMOTE), which generates synthetic samples for the minority class by interpolating between existing data points. This process results in a more balanced dataset, enhancing the model's ability to learn from rare events [38].

SMOTE has been effectively integrated with various machine learning algorithms, as shown in Table 2, such as support vector machines (SVM) [17,19,21,39–44], Naive Bayes [17,43,45], Decision Trees (DT) [17,42,45], Random Forests (RF) [17,19,42,43,46,47], and gradient boosting (e.g., XGBoost) [48]. It has also been applied in deep learning frameworks, including Convolutional Neural Networks (CNN) [49] and deep neural networks (DNN) [50]. Hybrid models, such as AdaBoost [17,21], stacking [51], and LightGBM [44] further demonstrate SMOTE's adaptability, especially in ensemble learning. These applications have significantly improved classification accuracy across diverse fields, such as geological and mineral resource management [41,48,51–53], cybersecurity and threat detection [45,46,49], financial and asset management [21,40], health diagnosis and risk assessment [17,19,54], engineering diagnostic and fault detection [44,50,55], and structural safety and risk analysis [43,56] to software reliability [42].

Table 2. SMOTE application literature review.

<table>
<tr><th rowspan="2">Application Domain</th><th colspan="3">Applied Method</th></tr>
<tr><th>Machine Learning</th><th>Deep Learning</th><th>Hybrid Model</th></tr>
<tr><td rowspan="4">Geological and Mineral Resource Management</td><td>Prado et al. (2020) [41]</td><td rowspan="4"></td><td rowspan="4">Xiao et al. (2024) [51]</td></tr>
<tr><td>Chen et al. (2022) [52]</td></tr>
<tr><td>Dong et al. (2020) [48]</td></tr>
<tr><td>Ibrahim et al. (2023) [53]</td></tr>
<tr><td rowspan="3">Cybersecurity and Threat Detection</td><td>Qazi and Raza (2012) [45]</td><td rowspan="3">Zhang et al. (2020) [49]</td><td rowspan="3">Karatas et al. (2020) [47]</td></tr>
<tr><td>Ullah and Mahmoud (2019) [46]</td></tr>
<tr><td>Karatas et al. (2020) [47]</td></tr>
<tr><td rowspan="2">Financial and Asset Management</td><td>Sun et al. (2020) [21]</td><td rowspan="2"></td><td rowspan="2">Sun et al. (2020) [21]</td></tr>
<tr><td>Li et al. (2018) [40]</td></tr>
<tr><td rowspan="3">Health Diagnosis and Risk Assessment</td><td>Kumari et al. (2023) [17]</td><td>Bradley and Rajendran (2022) [54]</td><td>Kumari et al. (2023) [17]</td></tr>
<tr><td>Bazarnovi and Mohammadian (2024) [19]</td><td rowspan="2"></td><td rowspan="2">Bradley and Rajendran (2022) [54]</td></tr>
<tr><td>Bradley and Rajendran (2022) [54]</td></tr>
<tr><td rowspan="2">Engineering Diagnostic and Fault Detection</td><td rowspan="2">Zhang et al. (2024) [44]</td><td>Gamel et al. (2024) [50]</td><td rowspan="2">Zhang et al. (2024) [44]</td></tr>
<tr><td>Chen et al. (2024) [55]</td></tr>
<tr><td rowspan="2">Structural Safety and Risk Analysis</td><td>Zhen et al. (2023) [43]</td><td rowspan="2"></td><td rowspan="2"></td></tr>
<tr><td>Chen and Zhang (2022) [56]</td></tr>
<tr><td>Software Reliability</td><td>Feng et al. (2021) [42]</td><td></td><td></td></tr>
</table>

Despite its extensive use in other domains, SMOTE has rarely been applied to forecasting intermittent spare part demand. Considering the significant data imbalance in slow-moving and intermittent spare parts, applying SMOTE in this context offers a promising approach to improve prediction accuracy and address critical data-level challenges.

2.3. Loss Function

Loss functions play a crucial role in forecasting intermittent demand by guiding model training and addressing the challenges of imbalanced datasets. This function quantifies the difference between predicted and actual values, influencing how a model learns patterns in the data. By adjusting the way errors are penalized, loss functions can improve the model's ability to focus on rare but critical events without significantly increasing computational complexity [57].

Loss functions are widely utilized in classification tasks to optimize the predictive accuracy of models, mainly when applied to imbalanced datasets. These functions allocate more importance to rare events, ensuring that critical occurrences are accurately predicted. This capability is particularly relevant for intermittent demand forecasting, where demand periods are sporadic and overshadowed by long periods of zero demand.

Studies have demonstrated the effectiveness of loss functions in addressing data imbalance across various domains. Table 3 summarizes key applications of loss functions, including their use in deep learning models [22,23,57–63], although some also apply to the boosting model [64], for example, tree-based models [65]. Examples include applications in healthcare [59–61,63], finance [64,65], industrial and mechanical [22,23,57], general [58], and environmental applications [62]. In industrial and mechanical applications, loss functions have proven effective for fault diagnosis [22,23], although their application to intermittent demand remains limited. These studies highlight how loss functions enhance model sensitivity to rare events, making them valuable for improving intermittent demand forecasting.

Table 3. Loss function application literature review.

Application Domain	Applied Method		
	Machine Learning	Deep Learning	Hybrid Model
Healthcare		Pasupa et al. (2023) [59]	
		Roy et al. (2022) [60]	
		Gökkan and Kuntalp (2022) [61]	
		Büttner et al. (2024) [63]	
Finance	Hu et al. (2022) [64]		Mushava and Murray (2022) [64]
Industrial and Mechanical		Jiang et al. (2023) [22]	
		Zhao et al. (2022) [23]	
		Nguyen and Thai (2023) [57]	
Environmental Application		You et al. (2023) [62]	
General		Lin et al. (2023) [58]	

The design of a loss function determines how errors are addressed during model optimization, balancing sensitivity to rare events with overall performance. Customizing the loss function to match the unique characteristics of intermittent demand enables models to capture demand patterns better. Table 3 highlights how loss functions have been applied across domains and their relevance to tackling the challenges of imbalanced datasets.

Despite their potential, loss functions remain underexplored in intermittent demand forecasting. Expanding their application in this context presents an opportunity to address the specific challenges of forecasting slow-moving and intermittent spare parts. By effectively leveraging loss functions, this study aims to improve forecasting accuracy and support more efficient inventory management.

3. Research Methodology

The framework used in developing the methodology is the Cross-Industry Standard Process for Data Mining (CRISP-DM) [66]. The CRISP-DM framework was chosen for its flexibility and structured approach to data-driven modeling, making it particularly suitable for handling complex datasets, such as intermittent spare part demand. Below is a complete explanation of each framework phase applied to develop the proposed hybrid ensemble forecasting model, as shown in Figure 1.

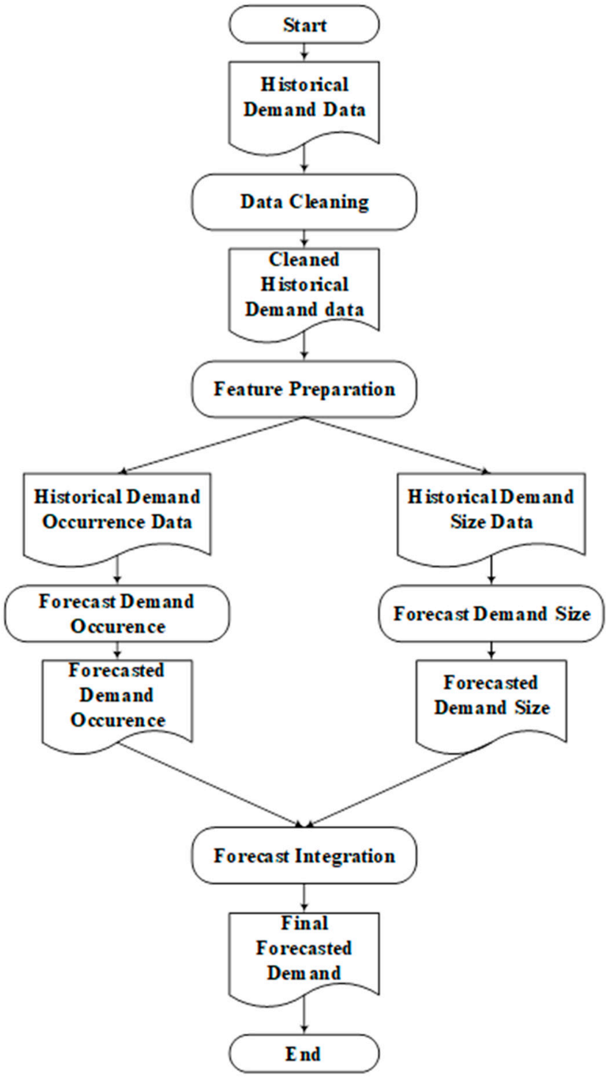

Figure 1. Framework of hybrid ensemble model for advancing intermittent spare part demand forecasting.

1. Business and Data Understanding: This phase understands the operational challenges in the one major oil and gas industry, particularly managing intermittent spare part demand. Spare parts are essential for maintaining high-value equipment, yet their irregular demand patterns result in challenges, such as overstocking costs and stock-out risks. This study addresses these issues by improving forecasting accuracy and addressing the industry's challenges through advanced data-driven methodologies.
2. Data Preparation: The historical spare part dataset underwent a complete preprocessing phase to ensure accuracy and reliability. At first, normalization is performed using the MinMaxScaler to scale the data within a range of [0, 1], which is particularly beneficial for models that are sensitive to the scale of input data [67].

After that, the preparation process involves splitting the dataset into two components: demand occurrence and demand size. This separation allows for the development of classification models for predicting demand events and regression models for forecasting the demand size. A sliding window technique is applied to enhance the performance of sequence-based models like LSTM. This approach generates fixed-length input sequences, effectively capturing temporal dependencies within the data and providing richer information for modeling [68]. The prepared dataset is the foundation for building robust models, ensuring data consistency and structure.

For sequence-based models like LSTM, a sliding window approach was implemented to construct input sequences of fixed lengths, effectively capturing temporal dependencies within the data. Addressing the significant class imbalance present in the dataset, the Synthetic Minority Oversampling Technique (SMOTE) was applied. SMOTE generates synthetic data points for the minority class, improving the model's sensitivity to rare events. This approach follows best practices in handling imbalanced datasets, such as those outlined by [38].

3. Modeling: The modeling phase is divided into two stages. Namely, a classification model is used to forecast demand occurrence, and a regression model is employed to forecast the demand size [14]. The classification stage predicts whether a demand event will occur within a given time frame. A range of base models are used, including Logistic Regression, Decision Trees, Random Forests, LightGBM, SVM, and multi-layer perceptron (MLP) until deep learning models such as LSTM and DNN. In this phase, the Synthetic Minority Oversampling Technique (SMOTE) is applied to handle the significant class imbalance in the demand occurrence dataset, improving the model's ability to detect rare demand events. Additionally, deep learning models such as LSTM and DNN are employed to capture nonlinear relationships and temporal dependencies. Focal loss is incorporated into the training process of deep learning models to enhance their sensitivity to minority-class predictions. The outputs of these base models are combined using a stacking ensemble approach, where a linear regression meta-learner aggregates predictions to produce the final forecast. Genetic Algorithms (GA) optimize the ensemble weights and decision thresholds. The GA iteratively refines the weights assigned to each base model by minimizing the ensemble's overall error [69]. The final stage integrates model outputs into the forecasting system. Forecasted demand occurrence from classification models is combined with forecasted demand size from regression models to produce the final forecast.
4. Evaluation and Deployment: The model evaluation uses k-fold cross-validation for time series data, ensuring robust and unbiased performance assessment. The evaluation metrics include the Area Under the Curve (AUC) for forecast demand occurrence and the Mean Squared Error (MSE) for forecast demand size. The k-fold cross-validation approach ensures that the models are validated comprehensively and are capable of generalizing to unseen data [68].

The AUC evaluates the model's ability to distinguish between demand and no-demand periods by measuring the area under the Receiver Operating Characteristic (ROC) curve, representing the trade-off between true positive and false positive rates across various thresholds. Values closer to 1.0 indicate superior discriminatory performance, particularly in handling imbalanced datasets [70].

The Mean Squared Error (MSE) quantifies the average squared difference between predicted and actual demand sizes for regression tasks. Lower MSE values indicate higher accuracy and minimal deviation in the prediction. This metric is widely adopted in analysis for its straightforward interpretation and effectiveness in capturing prediction errors.

This methodology offers a systematic approach to addressing the challenges of forecasting intermittent spare part demand. The proposed framework ensures accurate and reliable forecasts by combining structured data preparation, advanced modeling techniques, and rigorous evaluation, supporting better inventory decision-making.

Following evaluation, the model is deployed within the case study company's spare part inventory management system. Automated pipelines are set up to process incoming data, generate real-time predictions, and integrate the outputs into operational workflows. Continuous monitoring of AUC and MSE metrics ensures the model maintains its predictive reliability over time.

4. Proposed Hybrid Ensemble Model for Advancing Intermittent Spare Part Demand Forecasting

The proposed model addresses the challenges of intermittent and slow-moving spare part demand through a structured two-stage approach: forecasting demand occurrence (classification) and estimating demand size (regression). While both stages share a similar workflow—including data preprocessing, application of various models, and result aggregation using ensemble techniques—the models used in each stage are distinct, as shown in the flowchart (Figure 2). The detailed steps for implementing this model can be seen in Algorithm 1, which provides a comprehensive outline of the procedures for preprocessing, model training, and optimization.

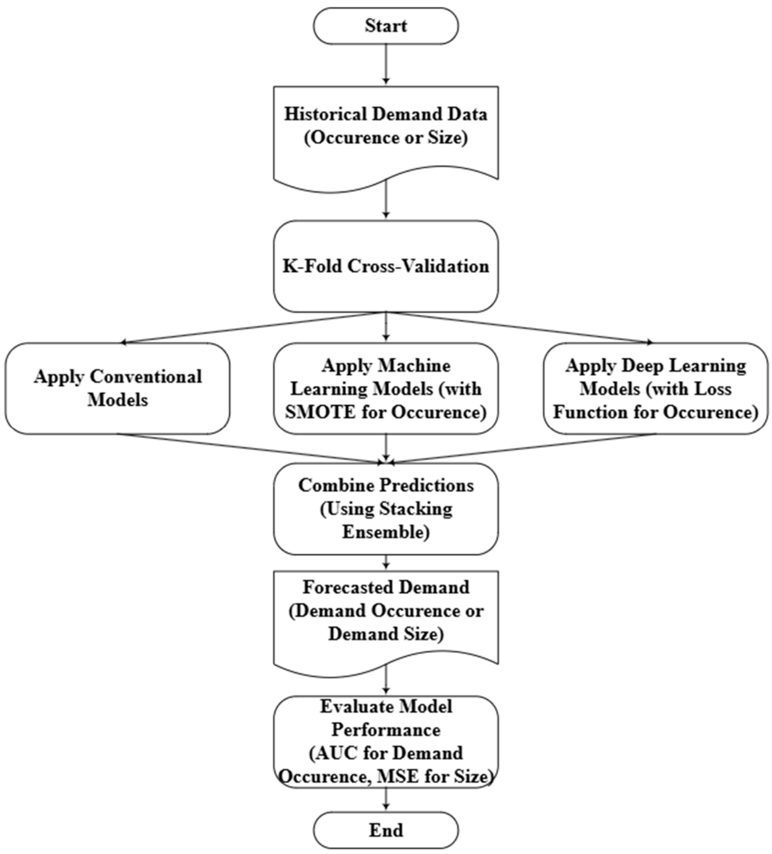

Figure 2. Detailed framework of forecast stage in hybrid ensemble model for advancing intermittent spare part demand forecasting.

In the demand occurrence (classification) stage, the model predicts whether demand will occur during a given period. This stage employs conventional models, including Logistic Regression and Decision Trees. These straightforward and interpretable methods serve as baseline predictors and have been validated extensively in forecasting tasks, particularly for intermittent demand [11,65,71]. Machine learning models such as Random Forest, support vector machine (SVM), and LightGBM are also implemented to handle complex relationships within the data [14,15,21,39]. To address the class imbalance characteristic of intermittent demand, the Synthetic Minority Oversampling Technique (SMOTE) is applied during data preprocessing, improving the representation of minority classes [37,38,42,52]. Deep learning models, such as long short-term memory networks (LSTM) and deep neural networks (DNN), are utilized for their ability to capture sequential dependencies and nonlinear patterns [7,50,72,73]. Focal loss and advanced loss functions focus on difficult-to-classify instances, further enhancing classification accuracy [58]. The results of these models are combined through a stacking ensemble approach, with linear regression used as the meta-learner. This combination has improved predictive accuracy, as demonstrated in previous studies [2,7,52,74,75]. The development of the two-stage model is based on the model [14], while the ensemble model is based on the model [7], both designed for intermittent data. The Hyperparameters' search range used in the models, listed in Table 4, align with best practices from prior studies [76,77].

Table 4. Search range hyperparameter.

Model	Hyperparameter	Search Range
Logistic Regression	Regularization Strength	[0.1, 1, 10, 100]
Decision Tree	Maximum Depth of the Tree	[3, 5, 7, 10]
Random Forest	Number of Trees in the Forest	[50, 100, 200]
	Maximum Depth of Each Tree	[3, 5, 7, 10]
Neural Network	Number of Neurons in Hidden Layer	[(50,), (100,), (50,50)]
	Ridge Regression	[0.0001, 0.001, 0.01]
LightGBM	Maximum Number of Leaves	[5, 10, 20, 30]
	Maximum Depth of the Tree	[3, 5, 7, 10]
	Learning Rate	[0.01, 0.05, 0.1, 0.2]
	Feature Fraction	[0.6, 0.8, 1.0]
	Minimum Data in Each Leaf	[5, 10, 20]
	Minimum Gain to Split	[0.01, 0.1, 0.2]
Ensemble (GA)	Population Size	[50, 100, 200]
	Crossover Rate	[0.5, 0.7, 0.9]
	Mutation Rate	[0.1, 0.2, 0.3]
	Tournament Size	[3, 5, 7]
	Number of Generations	[20, 40, 60]
	Threshold (for Classification)	[0–1]
Moving Average	Window Size	[1–(number of data − 1)]
ARIMA	Order of the AutoRegressive Part	[0, 1, 2, 3, 4, 5]
	Order of Differencing	[0, 1, 2]
Neural Network	Order of the Moving Average Part	[0, 1, 2, 3, 4, 5]
LSTM	Number of Neurons	[5, 10, 15, 20]
KNN	Number of Neighbors	[1, 2, 3, 4, 5, 6, 7]
DNN	Number of Layers	[1, 2, 3, 4, 5]
	Number of Units in Each Layer	[32, 64, 96, 128, 160, 192, 224, 256, 288, 320, 352, 384, 416, 448, 480, 512]
	Learning Rate	[0.01, 0.001, 0.0001]

The demand size (regression) stage estimates the magnitude of demand if the classification stage predicts its occurrence. While the workflow mirrors the classification stage, the models are explicitly designed for regression tasks. Conventional models such as Naive, Moving Average, and ARIMA are included to capture linear trends and smooth short-term fluctuations in demand data [2,7]. Machine learning models, including KNN and LightGBM, are implemented for their ability to model nonlinear patterns and provide robust predictions for irregular data [2,7,14]. For deep learning, LSTM and DNN capture long-term dependencies and intricate nonlinear patterns in the data. The ensemble of these models is optimized using Genetic Algorithms (GA), which iteratively refine weights and thresholds through selection, crossover, and mutation processes, ensuring the final model adapts effectively to intermittent demand size forecasting [69]. Algorithm 1 details the steps for this optimization process.

The flowchart in Figure 2 illustrates the shared steps for data preprocessing, model training, and ensemble aggregation in both stages while highlighting the distinct models used for classification and regression. This systematic design ensures the model is robust and reliable for demand occurrence and size forecasts. Evaluation metrics, such as Area Under the Curve (AUC) for classification and Mean Squared Error (MSE) for regression, are used to assess the model's performance. Combining these advanced techniques, the proposed model effectively addresses the complexities of forecasting intermittent spare part demand [13,20]. The results of the classification and regression stages are integrated to form the final demand forecast. If the classification model predicts demand occurrence, the regression stage estimates its magnitude. Conversely, if no demand occurrence is predicted, the demand for that period is set to zero. This hierarchical approach ensures accurate and realistic demand predictions.

Algorithm 1 Ensemble Model for Advancing Intermittent Spare Part Demand Forecasting

Input:

- Training samples X_{train}, Y_{train}; validation samples X_{val}, Y_{val}; test samples X_{test}, Y_{test}
- Classifiers

$$L_c = \left\{ \begin{array}{c} Logistic\ Regression,\ Decision\ Tree, \\ Random\ Forest,\ SVM,\ MLP, \\ LightGBM, LSTM, DNN \end{array} \right\}$$

- Regressor

$$L_r = \left\{ \begin{array}{c} Naive,\ ARIMA, \\ Moving\ Average,\ LSTM, KNN, DNN \end{array} \right\}$$

Step 1: Data Processing

1. Handle Missing Data
   - Impute or remove missing values in $X_{train}, X_{val}, X_{test}$ to ensure consistency across the dataset.
2. Handle Imbalanced Data
   - Use SMOTE to generate synthetic samples for the minority class in Y_{train}.

$$x_{new} = x_i + \lambda(x_{nn} - x_i)$$

Algorithm 1 *Cont.*

Step 2: Grid Search to Find the Best Hyperparameter for the Base Model

3. Define hyperparameter search space for each base model $L_b = \{L_{c1}, L_{c2}, \ldots, L_{cm}, L_{r1}, L_{r2}, \ldots, L_{rn}\}$.
   - Let H_b represent the hyperparameter search space for each base model L_b, where

     $$\mathrm{H}_b = \{h_1, h_2, \ldots, h_q\}$$

     where h_q are the different hyperparameter configurations for model L_b.
4. For each hyperparameter configuration $h_1 \epsilon \mathrm{H}_b$ of base model L_b:
   - Train model $L_b(h_i)$ using the training dataset X_{train}, Y_{train} with oversampling using SMOTE (for classifier).
   - Apply time series cross-validation and evaluate model performance for each split using X_{val}, Y_{val}.
   - Use focal loss as the loss function during training:

     $$\mathrm{FL}(y, \hat{y}) = -\alpha(1 - \hat{y})^{\gamma}.\log(\hat{y})$$

   - Compute prediction scores on the validation dataset X_{val}, Y_{val} and evaluate the model using a metric M_b, which can be the following:

     $$M_b = \begin{cases} AUC(L_b(h_i)),\ if\ L_b\ is\ a\ classification\ model \\ MSE(L_b(h_i)),\ if\ L_b\ is\ a\ regression\ model \end{cases}$$

5. Select the best hyperparameter configuration h for each base model L_b:
   - For classification models:

     $$h_b = \arg\max_{h_i \epsilon \mathrm{H}_b} AUC(L_b(h_i))$$

   - For regression models:

     $$h_b = \arg\min_{h_i \epsilon \mathrm{H}_b} MSE(L_b(h_i))$$

6. Train the base model L_b with the best hyperparameters h_b:
   - Train model $Lb(\hbar b)$ on the entire training dataset *Xtrain*, *Ytrain*.
   - Store the final trained model $Lb(\hbar b)$ for further use in ensemble learning.

Step 3: Genetic Algorithm for Ensemble Weight and Threshold Optimization
GA Parameter Optimization:

1. Initialize a search space for GA parameters: population size, crossover rate, mutation rate, and number of generations.
2. For each combination of GA parameters $\{p, c, m, g\}$, run GA for both the classification and regression stages:
   - For classification, evaluate the performance using AUC.
   - For regression, evaluate the performance using MSE.
3. Select the best combination of GA parameters $\{p^*, c^*, m^*, g^*\}$:
   - For classification models:

     $$(p^*, c^*, m^*, g^*) = \arg\max_{p,c,m,g} AUC_{GA}$$

   - For regression models:

     $$(p^*, c^*, m^*, g^*) = \arg\min_{p,c,m,g} MSE_{GA}$$

Algorithm 1 *Cont.*

Classification Stage:

1. Initialize GA with a population of weights for combining the classification models $w = \{w_1, w_2, \ldots, w_m\}$ and the threshold α to be optimized.
2. Define fitness function for classification:
   - For each individual $[w_1, w_2, \ldots, w_m; \alpha]$:
     - Compute the weighted ensemble prediction:
       $$d_i = \sum_{j=1}^{m} w_j.L_{cj}(x_i)$$
     - Apply the threshold α to convert the result to binary: v
       $$v_i = \begin{cases} 1, & if\ d_i > \alpha \\ 0, & if\ d_i \leq \alpha \end{cases}$$
     - Calculate the AUC of the ensemble and define the fitness function as:
       $$F_{class}\left(w_j, \alpha\right) = 1 - AUC(v_i)$$
3. Optimize ensemble weights and threshold using GA:
   - Apply GA operations (selection, crossover, mutation) to optimize the weights w_j and threshold α.
4. Select the best solution w^*, α^* that maximizes AUC.

Regression Stage:

1. Initialize GA with a population of weights for combining the classification models $w = \{w_1, w_2, \ldots, w_m\}$ and the threshold α to be optimized.
2. Define fitness function for regression:
   - For each individual $[w_1, w_2, \ldots, w_m; \alpha]$:
     - Compute the weighted ensemble prediction:
       $$f_i = \sum_{k=1}^{n} w_k.L_{rk}(x_i)$$
     - Define the fitness function to minimize the error:
       $$F_{reg}(w_k) = MSE(f_i)$$
3. Optimize ensemble weights and threshold using GA:
   - Apply GA operations (selection, crossover, mutation) to optimize the weights w_k.
4. Select the best solution w^* that minimizes the error.

Step 4: Combine Classification and Regression Predictions
Classification Prediction:

1. For each test sample $x_i \epsilon X_{test}$, compute the weighted sum of classification predictions:
   $$d_i = \sum_{j=1}^{m} w_j.L_{cj}(x_i)$$
2. Apply the optimized threshold α^* and set:
   $$v_i = \begin{cases} 1, & if\ d_i > \alpha \\ 0, & if\ d_i \leq \alpha \end{cases}$$

Algorithm 1 *Cont.*

Regression Prediction (for: $v_i = 1$):

1. For each test sample $x_i \epsilon X_{test}$ where $v_i = 1$, compute the weighted sum of regression predictions:
$$f_i = \sum_{k=1}^{n} w_k.L_{rk}(x_i)$$
2. Store the predicted demand quantity f_i.
3. For samples where $v_i = 0$, set $f_i = 0$.

Output:

Return the final classification predictions V and regression predictions F.

5. Experiment and Results

The historical data collected span from January 2012 to December 2023, representing 624 weeks of recorded spare part usage. From 716 spare parts, this study focuses specifically on those exhibiting intermittent demand patterns. These patterns are characterized by large inter-arrival times and sparse occurrences, highlighting the irregularity and low frequency of demand events. Figure 3 illustrates the demand patterns for a representative spare part over the entire period, segmented into pre-COVID-19 (2012–2019), during COVID-19 (2020–2021), and post-COVID-19 (2022–2023). The data reveal a highly intermittent pattern with long intervals between demand occurrences and occasional spikes, typical of slow-moving spare parts predominantly used during scheduled maintenance.

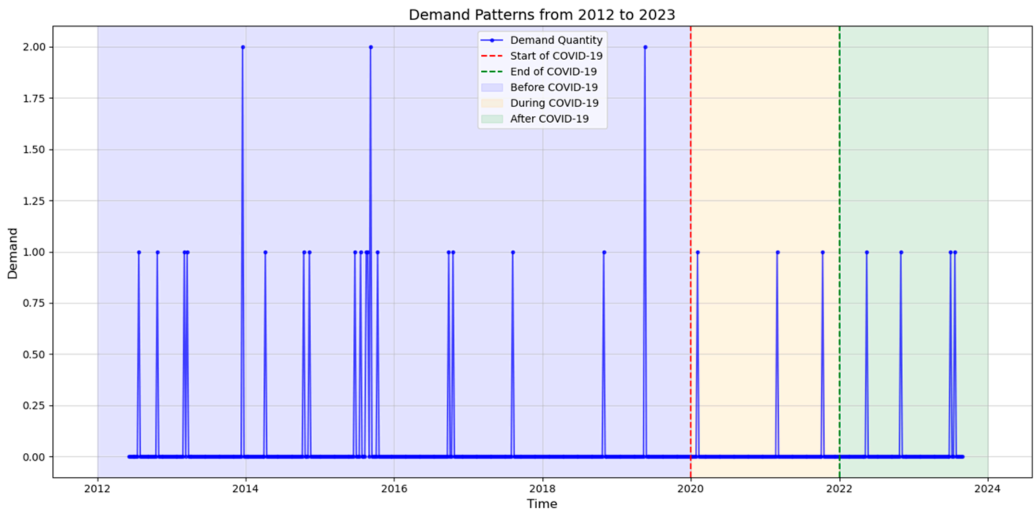

Figure 3. Demand pattern of spare parts from 2012 to 2024.

During the pre-COVID-19 period, the demand pattern remained stable, characterized by consistent intervals and spikes. Similarly, during the COVID-19 period, no significant deviations from these patterns were observed, with the intervals between demand events and the frequency of spikes mirroring those of the pre-pandemic period. This stability can be attributed to the continuous operation of industrial machinery, which necessitated ongoing maintenance regardless of external disruptions.

In the post-COVID-19 period, the patterns persisted without notable changes, further reinforcing the conclusion that spare parts demand is primarily driven by internal factors, such as maintenance schedules, rather than external disruptions like the pandemic. These findings underscore the resilience of spare parts demand, even amidst global crises, and highlight the critical role of maintenance activities in sustaining industrial operations. However, it is important to note that, while demand remained stable, global disruptions like COVID-19 had a more significant impact on the supply side, where vendors faced challenges in meeting demand due to logistical constraints, labor shortages, and supply chain disruptions [78]. These issues, although external to the demand itself, underscore the importance of robust vendor management and supply chain resilience to mitigate potential risks in the availability of spare parts.

The datasets used in this study reflect varying characteristics of intermittent demand, as summarized in Table 5. The first dataset has an ADI of 203.7 and CV^2 of 0.317, representing highly intermittent and variable demand. Datasets 2 to 5 exhibit distinct combinations of ADI and CV^2, capturing the spectrum of intermittent demand patterns observed in actual operations. The high ADI values across all datasets underscore the demand's sparse and irregular nature, with values far exceeding typical thresholds for classifying intermittent demand. The variability in CV^2 reflects the unpredictability of demand magnitudes, indicating that, while demand occurrences are sparse, their intensity can vary significantly. Moreover, the substantial proportion of periods without demand highlights traditional forecasting methods' challenges when dealing with data, as conventional models often struggle to balance sparsity and variability.

Table 5. Characteristics of datasets.

Dataset	ADI	CV^2
Dataset 1	203.7	0.317
Dataset 2	209.7	0.146
Dataset 3	204.35	0.252
Dataset 4	5.9	0.132
Dataset 5	4	0.254

The selection of five datasets with varying characteristics was driven by the need to evaluate forecasting model performance across various real-world scenarios. The first dataset represents a baseline scenario with a combination of sparsity and moderate variability, while Datasets 2 and 3 emphasize extreme sparsity with very high ADI values. Conversely, Datasets 4 and 5 feature more frequent demand occurrences with lower ADI values, providing a contrast that ensures comprehensive model testing. This diverse dataset design ensures that forecasting models are robust and generalizable across different intermittent demand patterns, a crucial requirement for industrial applications. These findings demonstrate the importance of advanced forecasting models that can effectively handle irregular demand patterns while addressing the variability and sparsity inherent in the data. Even during significant global disruptions, the stability of demand patterns further emphasizes the necessity of reliable forecasting systems to optimize inventory management and minimize downtime in industrial settings.

The datasets reveal several critical insights about the characteristics of intermittent demand. High ADI values reflect long intervals between demand occurrences, which aligns with the definition of intermittent demand. The variability in demand, as captured by CV^2, indicates that while demand events are sparse, their magnitudes are unpredictable. Additionally, the large proportion of periods without demand underscores the sparse and irregular nature of the data, posing significant challenges for traditional forecasting models. The observed consistency in the maximum and minimum demand values suggests that, despite irregular occurrences, the magnitude of demand events remains relatively stable within datasets.

After analyzing dataset characteristics and the inherent challenges in forecasting intermittent demand, the next step evaluates the performance of various models in addressing these challenges. Hyperparameter tuning was conducted using Genetic Algorithms (GA) to identify the optimal configurations for each model, with the best GA parameters determined through grid search. Once the optimal hyperparameters were identified, the models were tested on all datasets, with validation performed using k-fold cross-validation to ensure reliability and robustness. The results of these evaluations are presented in the AUC and MSE in Table 6.

Table 6. Optimized hyperparameter.

Model	Hyperparameter	Value
Logistic Regression	Regularization Strength	0.1
Decision Tree	Maximum Depth of the Tree	3
Random Forest	Number of Trees in the Forest	3
	Maximum Depth of Each Tree	50
Neural Network	Number of Neurons in Hidden Layer	(50,)
	Ridge Regression	0.0001
LightGBM	Maximum Number of Leaves	20
	Maximum Depth of the Tree	10
	Learning Rate	0.1
	Feature Fraction	0.8
	Minimum Data in Each Leaf	5
	Minimum Gain to Split	0.01
Ensemble (GA)	Population Size	100
	Crossover Rate	0.7
	Mutation Rate	0.3
	Tournament Size	3
	Number of Generations	60
	Threshold (for classification)	0.444
Moving Average	Window Size	1
ARIMA	Order of the AutoRegressive Part	(0, 0, 0)
	Order of Differencing	1
Neural Network	Order of the Moving Average Part	3
LSTM	Number of Neurons	20
KNN	Number of Neighbors	3
DNN	Number of Layers	1
	Number of Units in Each Layer	0.01
	Learning Rate	288

The overall performance, as shown in Table 7 and Figure 4, demonstrates that the proposed model outperformed all other models, achieving an AUC of 0.735, significantly higher than traditional machine learning models such as Logistic Regression (0.548), Decision Tree (0.563), and Random Forest (0.523). Even advanced models, including LSTM and DNN, with overall AUC scores of 0.569 and 0.462, respectively, fell short of the proposed approach. While LightGBM performed comparatively better, with an AUC of 0.718, the proposed model demonstrated superior capability in addressing the unique characteristics of intermittent demand. These results highlight the effectiveness of the proposed ensemble approach, leveraging focal loss and SMOTE techniques to handle data imbalance and sparsity.

Table 7. AUC comparison across model validation.

Dataset	Model								
	LR	DT	RF	SVM	NN	LightGBM	LSTM	DNN	Proposed
Dataset 1	0.559	0.596	0.597	0.510	0.440	0.705	0.450	0.439	0.723
Dataset 2	0.630	0.592	0.463	0.488	0.377	0.856	0.376	0.479	0.885
Dataset 3	0.526	0.657	0.607	0.507	0.333	0.873	0.330	0.344	0.868
Dataset 4	0.521	0.472	0.450	0.489	0.538	0.555	0.549	0.543	0.604
Dataset 5	0.504	0.501	0.498	0.490	0.518	0.602	0.504	0.504	0.593
Overall	0.548	0.563	0.523	0.497	0.441	0.718	0.442	0.462	0.735

LR—Logistic Regression, DT—Decision Tree, RF—Random Forest, NN—Neural Network.

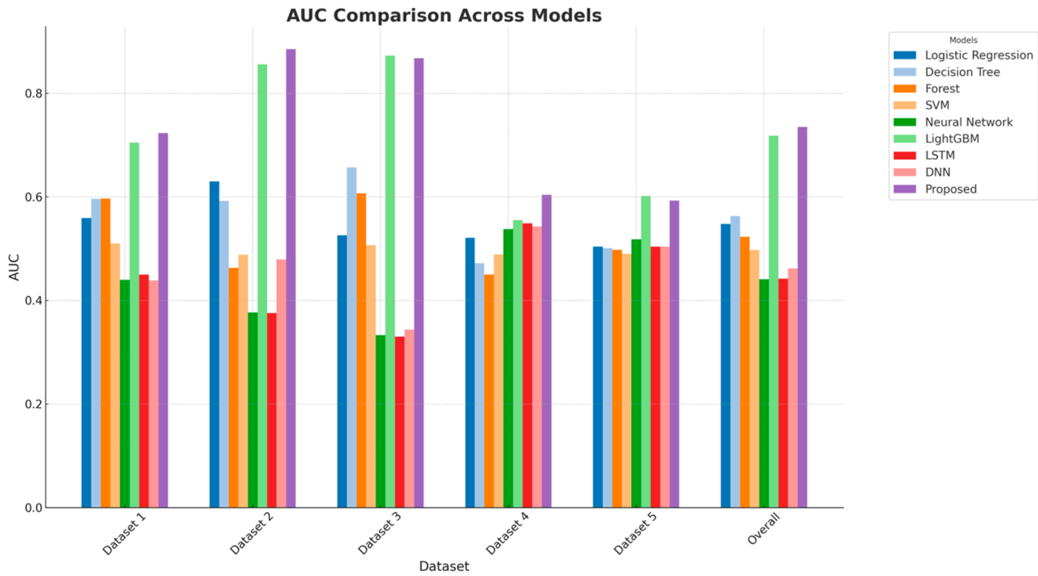

Figure 4. AUC comparison across model validation.

The results for the MSE, as shown in Table 8 and Figure 5, further confirm the superiority of the proposed model, which achieved the lowest overall MSE of 3.47. This represents a significant improvement over baseline methods such as Naive (8.51), ARIMA (8.04), and traditional machine learning approaches like Moving Average (8.31) and KNN (5.68). Even advanced methods like LSTM (5.69) and LightGBM (4.47) could not match the accuracy of the proposed model, demonstrating its ability to handle the sparse and highly variable nature of intermittent demand data.

Table 8. MSE comparison across model validation.

Dataset	Model							
	Naive	Moving Average	ARIMA	LSTM	KNN	DNN	LightGBM	Proposed
Dataset 1	3.83	3.83	4.59	4.64	3.83	5.81	3.83	1.44
Dataset 2	8.75	8.75	26.46	3.93	4.36	5.15	4.00	2.24
Dataset 3	20.92	20.92	21.42	9.55	11.81	16.04	9.06	4.57
Dataset 4	8.14	8.14	30.57	4.47	4.40	4.88	4.42	4.04
Dataset 5	8.02	8.02	15.85	5.99	6.19	6.55	6.02	5.04
Overall	8.51	8.31	8.04	5.69	5.68	5.54	4.47	3.47

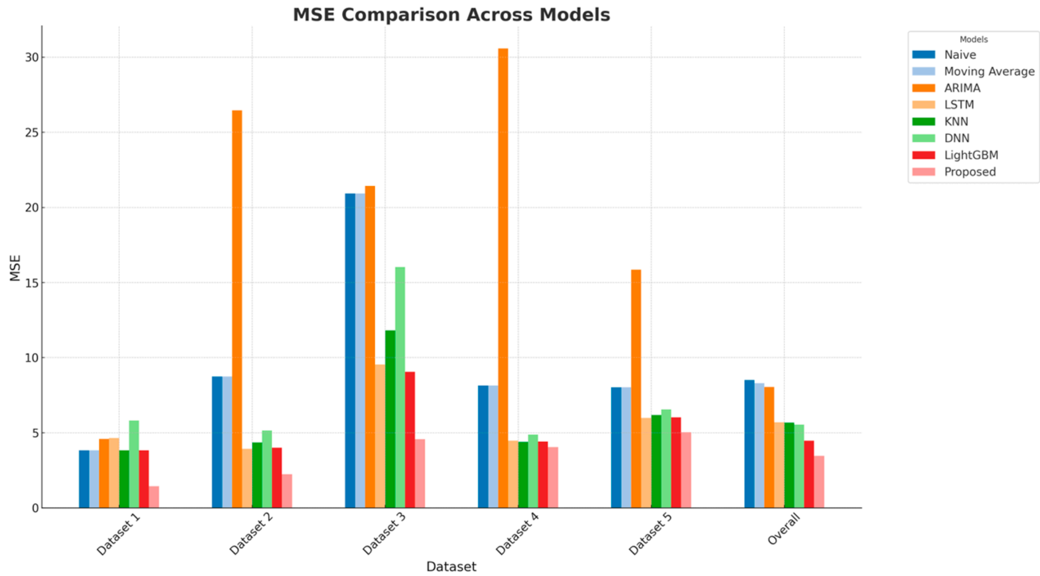

Figure 5. MSE comparison across model validation.

Dataset-specific results provide further insights into the robustness of the proposed model. On Dataset 1, the proposed model achieved an AUC of 0.723 and the lowest MSE of 1.44, outperforming models such as Logistic Regression (AUC 0.559) and LightGBM (MSE 3.83). For Dataset 2, which exhibits extreme sparsity, the proposed model achieved the highest AUC of 0.885 and the lowest MSE of 2.24, demonstrating its strength in capturing highly intermittent patterns. Similarly, for Dataset 3, characterized by high variability in demand magnitudes, the proposed model maintained superior performance with an AUC of 0.868 and an MSE of 4.57.

The proposed model delivered robust results for Datasets 4 and 5, where demand is more frequent but intermittent; it achieved AUC scores of 0.604 and 0.593 and MSE values of 4.04 and 5.04, respectively. These findings underline the versatility of the proposed approach in addressing diverse patterns of intermittent demand.

In conclusion, the validation phase results demonstrate the proposed model's superior performance compared to traditional and benchmark intermittent demand forecasting methods. These findings provide a strong foundation for the subsequent testing phase, where the model's performance will be further evaluated under test conditions to confirm its generalizability and robustness.

The models used for performance comparison in the test phase include the Syntetos–Boylan Approximation (SBA) model [10,27], the Bootstrapping model [26], the stacking ensemble model [7], and the best threshold intermittent demand combination forecasting (BTIDCF) model [14]. The SBA and Bootstrapping models were selected for their frequent application in intermittent demand scenarios. In contrast, the stacking ensemble and BTIDCF models were chosen as benchmarks for advanced ensemble and two-stage methods in handling intermittent data. The comparison in the test phase aims to assess the generalizability and robustness of the proposed ensemble model against these established methods. All results are based on the best hyperparameters optimized for each base model during the validation phase.

The results of the testing phase with AUC performance can be seen in Figure 6 and Table 9, while the MSE performance can be seen in Figure 7 and Table 10. The test phase results show that the proposed model outperforms the AUC and MSE metrics comparison models. The overall AUC for the proposed model is 0.68, significantly higher than SBA (0.50), Bootstrapping (0.50), the stacking ensemble model [7] (0.64), and the BTIDCF model [14] (0.50). This superior performance demonstrates the model's ability to generalize well across datasets with varying characteristics. Similarly, the proposed model achieves the lowest overall MSE of 2.07, far surpassing SBA (3.08), Bootstrapping (14.25), the stacking ensemble model [7] (4.99), and the BTIDCF model [14] (3.69), further validating its capacity to minimize forecast errors.

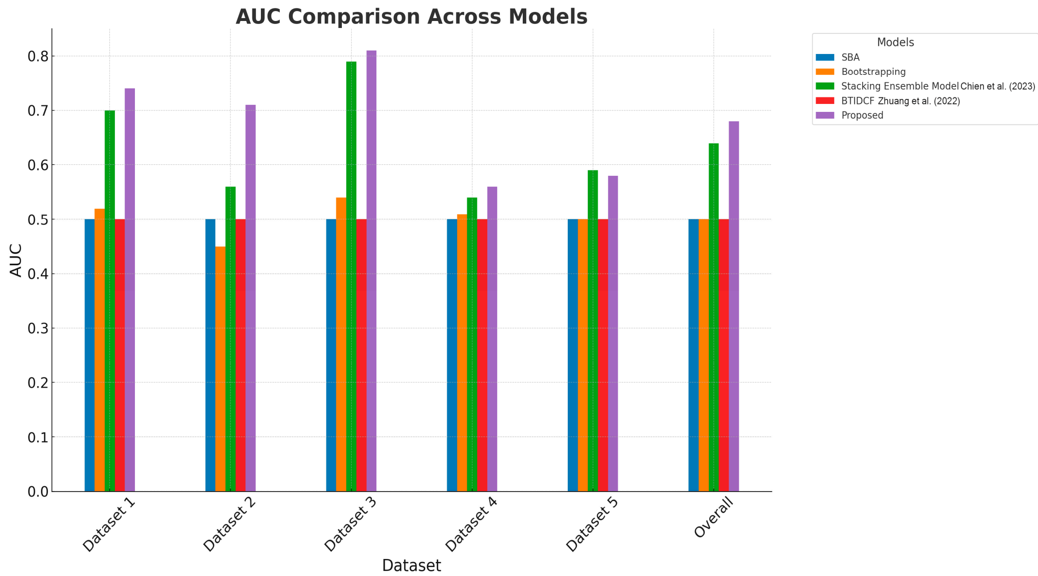

Figure 6. AUC comparison across model test [7,14].

Table 9. AUC comparison across model test.

Dataset	Model				
	SBA	**Bootstrapping**	**Stacking Ensemble Model [7]**	**BTIDCF [14]**	**Proposed**
Dataset 1	0.50	0.52	0.70	0.50	0.74
Dataset 2	0.50	0.45	0.56	0.50	0.71
Dataset 3	0.50	0.54	0.79	0.50	0.81
Dataset 4	0.50	0.51	0.54	0.50	0.56
Dataset 5	0.50	0.50	0.59	0.50	0.58
Overall	0.50	0.50	0.64	0.50	0.68

Table 10. MSE comparison across model test.

Dataset	Model				
	SBA	**Bootstrapping**	**Stacking Ensemble Model [7]**	**BTIDCF [14]**	**Proposed**
Dataset 1	1.63	11.66	3.83	1.63	1.44
Dataset 2	1.05	12.36	2.84	1.05	3.71
Dataset 3	1.34	5.41	11.44	1.34	2.01
Dataset 4	4.24	23.04	2.85	4.89	1.51
Dataset 5	7.17	18.77	3.98	9.52	1.70
Overall	3.08	14.25	4.99	3.69	2.07

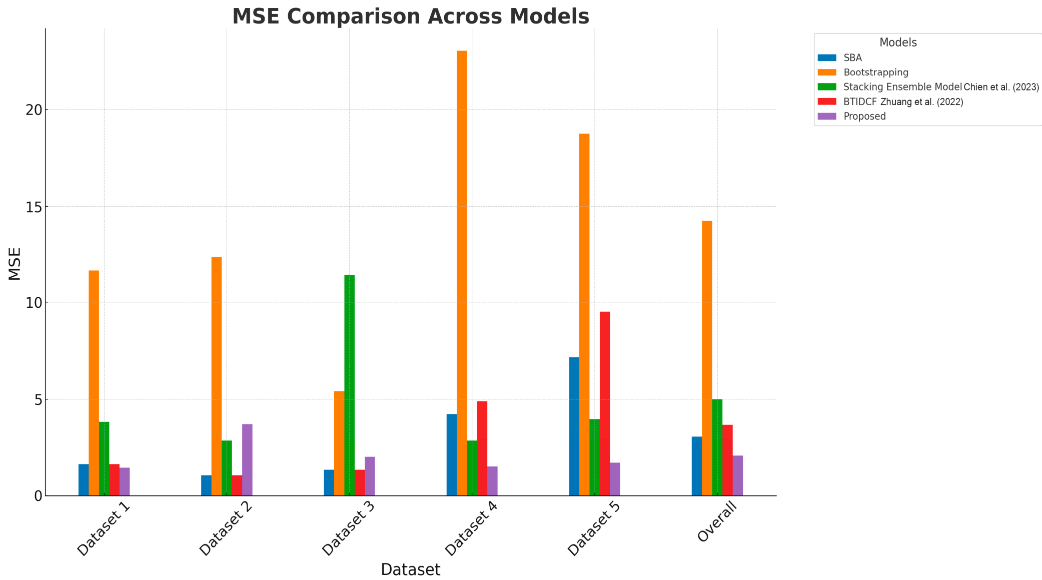

Figure 7. MSE comparison across model test [7,14].

For Dataset 1, the proposed model achieves the best performance, with an AUC of 0.74 and an MSE of 1.44, outperforming SBA (AUC 0.50 and MSE 1.63) and the stacking ensemble model [7] (MSE 3.83). This indicates its reliability in capturing intermittent patterns with moderate sparsity. For Dataset 2, characterized by high sparsity, the proposed model attains an AUC of 0.71 and an MSE of 3.71, effectively managing sparse demand patterns. In contrast, other models, such as Bootstrapping, exhibit much higher MSE (12.36). This underscores the adaptability of the proposed approach in sparse scenarios.

Dataset 3, which reflects significant variability in demand magnitudes, highlights the proposed model's ability to handle unpredictable patterns, achieving the highest AUC of 0.81 and a much lower MSE of 2.01 compared to SBA (AUC 0.50 and MSE 1.34) and the stacking ensemble model [7] (MSE 11.44). For Datasets 4 and 5, where demand is relatively more frequent but still irregular, the proposed model maintains its superior performance with AUC values of 0.56 and 0.58 and MSE values of 1.51 and 1.70, respectively, confirming its robustness across different demand scenarios.

Based on the results, the proposed model demonstrates strong generalizability for forecasting intermittent demand across various datasets. The model consistently outperforms traditional methods, such as SBA and Bootstrapping, and more recent ensemble approaches, like the stacking ensemble model [7] and Chien (2023), regarding AUC and MSE. This indicates its robustness in handling the key challenges of intermittent demand, including sparsity and variability.

The test phase results confirm that the model can effectively address intermittent demand patterns, particularly in datasets with extreme sparsity (e.g., Dataset 2) and high variability in demand magnitudes (e.g., Dataset 3). Additionally, its performance on datasets with a relatively higher frequency of demand occurrences (e.g., Dataset 4 and Dataset 5) further supports its applicability to a wide range of intermittent demand scenarios. These findings not only validate the model's reliability as a forecasting tool for general intermittent demand data but also carry significant managerial implications. The enhanced accuracy of the proposed model enables managers, particularly in industries such as oil and gas, to optimize inventory levels effectively by balancing spare part availability with cost efficiency. This reduces the risks of overstocking and stockouts, improving operational continuity and significant cost savings. Furthermore, the ability to forecast demand with greater precision empowers managers to make informed decisions that enhance service levels and mitigate the financial impacts of unplanned downtime. The model's generalizability across industries underscores its strategic relevance, providing robust solutions for handling complex and irregular demand patterns.

6. Conclusions

This study demonstrates the effectiveness of the proposed ensemble model in addressing the challenges of forecasting intermittent demand. Across diverse datasets with varying levels of sparsity and variability, the proposed model consistently outperformed traditional methods, such as SBA and Bootstrapping, and more advanced approaches, like the stacking ensemble model [7] and Chien (2023). On average, the proposed model improved by 32% in the AUC and 47% in the MSE compared to the benchmark models, highlighting its capability to provide more accurate and reliable forecasts. The integration of focal loss and SMOTE to handle data imbalance and sparsity proved to be a critical factor in achieving these superior results. Notably, the proposed model demonstrated exceptional adaptability to datasets with high sparsity and variability, such as Dataset 2 and Dataset 3, while maintaining strong performance on datasets with relatively higher demand frequencies, such as Dataset 4 and Dataset 5. These findings validate the model's applicability across general intermittent demand scenarios, particularly for industries requiring accurate forecasting to optimize inventory levels, minimize stockouts, and maintain operational continuity.

Despite its strong performance, the model exhibits some limitations. Its effectiveness diminishes slightly in datasets with lower sparsity and higher demand frequencies, as observed in Datasets 4 and 5. This suggests that its design is more suited to handling highly sparse and irregular patterns. Additionally, the computational complexity of the model, due to its ensemble structure and reliance on Genetic Algorithms for hyperparameter optimization, may limit its scalability in real-time applications or resource-constrained environments.

Future research should enhance the model's adaptability to smoother intermittent patterns while maintaining its robustness in highly sparse scenarios. This could involve exploring alternative resampling techniques to improve the representation of minority classes and adaptive loss functions tailored to intermittent data characteristics. Reducing the computational burden through more efficient optimization techniques, such as Particle Swarm Optimization or Bayesian Optimization, could improve the model's practicality for real-time applications.

A key direction for future work is integrating the forecasting model with inventory management systems. This integration would allow for direct evaluation of the relationship between forecasting accuracy and inventory performance metrics, such as service levels, stockouts, and holding costs. By embedding the forecasting model within inventory decision frameworks, researchers could better assess how improvements in forecast accuracy translate to operational and cost efficiencies in inventory systems. This holistic approach would provide deeper insights into the practical utility of the model in real-world industrial applications.

Author Contributions: Conceptualization, S.P.K., A.C., A.M. and R.T.C.; methodology, S.P.K., A.C., A.M. and R.T.C.; software, S.P.K.; validation, S.P.K., A.C., A.M. and R.T.C.; formal analysis, S.P.K., A.C., A.M. and R.T.C.; investigation, S.P.K., A.C., A.M. and R.T.C.; resources, S.P.K., A.C., A.M. and R.T.C.; data curation, S.P.K.; writing—original draft preparation, S.P.K.; writing—review and editing, S.P.K., A.C., A.M. and R.T.C.; visualization, S.P.K.; supervision, A.C., A.M. and R.T.C.; funding acquisition, R.T.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Faculty of Industrial Technology, Bandung Institute of Technology (ITB) from ITB research grant under the PPMI 2024 program No.1K/IT1.C07/SKKP/2024.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors upon request.

Acknowledgments: This research was conducted under the guidance, insight, and direction of the late Senator Nur Bahagia. His contributions were deeply appreciated throughout the research. Grammarly was used for grammar and typo assistance, and the content was reviewed and edited by the authors, who accept full responsibility for the final publication.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Affonso, T.B.; Conceição, S.V.; Muniz, L.R.; de Freitas Almeida, J.F.; de Lima, J.C. A New Hybrid Forecasting Method for Spare Part Inventory Management Using Heuristics and Bootstrapping. *Decis. Anal. J.* **2024**, *10*, 100415. [CrossRef]
2. Sareminia, S.; Amini, M. A Reliable and Ensemble Forecasting Model for Slow-Moving and Repairable Spare Parts: Data Mining Approach. *Comput. Ind.* **2023**, *145*, 103827. [CrossRef]
3. Hu, Q.; Boylan, J.E.; Chen, H.; Labib, A. OR in Spare Parts Management: A Review. *Eur. J. Oper. Res.* **2018**, *266*, 395–414. [CrossRef]
4. Molenaers, A.; Baets, H.; Pintelon, L.; Waeyenbergh, G. Criticality Classification of Spare Parts: A Case Study. *Int. J. Prod. Econ.* **2012**, *140*, 570–578. [CrossRef]
5. Hinton, H.L., Jr. *Defense Inventory, Continuing Challenges in Managing Inventories and Avoiding Adverse Operational Effects*; US General Accounting Office: Washington, DC, USA, 1999.
6. Ren, C.; An, N.; Wang, J.; Li, L.; Hu, B.; Shang, D. Optimal Parameters Selection for BP Neural Network Based on Particle Swarm Optimization: A Case Study of Wind Speed Forecasting. *Knowl. Based Syst.* **2014**, *56*, 226–239. [CrossRef]
7. Chien, C.-F.; Ku, C.-C.; Lu, Y.-Y. Ensemble Learning for Demand Forecast of After-Market Spare Parts to Empower Data-Driven Value Chain and an Empirical Study. *Comput. Ind. Eng.* **2023**, *185*, 109670. [CrossRef]
8. Pinçe, Ç.; Turrini, L.; Meissner, J. Intermittent Demand Forecasting for Spare Parts: A Critical Review. *Omega* **2021**, *105*, 102513. [CrossRef]
9. Croston, J.D. Forecasting and Stock Control for Intermittent Demands. *Oper. Res. Q. (1970–1977)* **1972**, *23*, 289. [CrossRef]
10. Boylan, J.E.; Syntetos, A.A. On the Bias of Intermittent Demand Estimates. *Int. J. Prod. Econ.* **2001**, *71*, 457–466. [CrossRef]
11. Hua, Z.; Zhang, B. A Hybrid Support Vector Machines and Logistic Regression Approach for Forecasting Intermittent Demand of Spare Parts. *Appl. Math. Comput.* **2006**, *181*, 1035–1048. [CrossRef]
12. Kourentzes, N. Intermittent Demand Forecasts with Neural Networks. *Int. J. Prod. Econ.* **2013**, *143*, 198–206. [CrossRef]
13. Lolli, F.; Gamberini, R.; Regattieri, A.; Balugani, E.; Gatos, T.; Gucci, S. Single-Hidden Layer Neural Networks for Forecasting Intermittent Demand. *Int. J. Prod. Econ.* **2017**, *183*, 116–128. [CrossRef]
14. Zhuang, X.; Yu, Y.; Chen, A. A Combined Forecasting Method for Intermittent Demand Using the Automotive Aftermarket Data. *Data Sci. Manag.* **2022**, *5*, 43–56. [CrossRef]
15. Rožanec, J.M.; Fortuna, B.; Mladenić, D. Reframing Demand Forecasting: A Two-Fold Approach for Lumpy and Intermittent Demand. *Sustainability* **2022**, *14*, 9295. [CrossRef]
16. Estabrooks, A.; Jo, T.; Japkowicz, N. A Multiple Resampling Method for Learning from Imbalanced Data Sets. *Comput. Intell.* **2004**, *20*, 18–36. [CrossRef]
17. Kumari, R.; Singh, J.; Gosain, A. SmS: SMOTE-Stacked Hybrid Model for Diagnosis of Polycystic Ovary Syndrome Using Feature Selection Method. *Expert. Syst. Appl.* **2023**, *225*, 120102. [CrossRef]
18. Huang, C.; Huang, X.; Fang, Y.; Xu, J.; Qu, Y.; Zhai, P.; Fan, L.; Yin, H.; Xu, Y.; Li, J. Sample Imbalance Disease Classification Model Based on Association Rule Feature Selection. *Pattern Recognit. Lett.* **2020**, *133*, 280–286. [CrossRef]
19. Bazarnovi, S.; Mohammadian, A.K. Addressing Imbalanced Data in Predicting Injury Severity after Traffic Crashes: A Comparative Analysis of Machine Learning Models. *Procedia Comput. Sci.* **2024**, *238*, 24–31. [CrossRef]
20. Chamseddine, E.; Mansouri, N.; Soui, M.; Abed, M. Handling Class Imbalance in COVID-19 Chest X-Ray Images Classification: Using SMOTE and Weighted Loss. *Appl. Soft Comput.* **2022**, *129*, 109588. [CrossRef]
21. Sun, J.; Li, H.; Fujita, H.; Fu, B.; Ai, W. Class-Imbalanced Dynamic Financial Distress Prediction Based on Adaboost-SVM Ensemble Combined with SMOTE and Time Weighting. *Inf. Fusion* **2020**, *54*, 128–144. [CrossRef]
22. Jiang, G.; Yue, R.; He, Q.; Xie, P.; Li, X. Imbalanced Learning for Wind Turbine Blade Icing Detection via Spatio-Temporal Attention Model with a Self-Adaptive Weight Loss Function. *Expert. Syst. Appl.* **2023**, *229*, 120428. [CrossRef]
23. Zhao, X.; Yao, J.; Deng, W.; Jia, M.; Liu, Z. Normalized Conditional Variational Auto-Encoder with Adaptive Focal Loss for Imbalanced Fault Diagnosis of Bearing-Rotor System. *Mech. Syst. Signal Process* **2022**, *170*, 108826. [CrossRef]
24. Boylan, J.E.; Syntetos, A.A. Spare Parts Management: A Review of Forecasting Research and Extensions. *IMA J. Manag. Math.* **2010**, *21*, 227–237. [CrossRef]

25. Syntetos, A.A.; Boylan, J.E.; Croston, J.D. On the Categorization of Demand Patterns. *J. Oper. Res. Soc.* **2005**, *56*, 495–503. [CrossRef]
26. Willemain, T.R.; Smart, C.N.; Schwarz, H.F. A New Approach to Forecasting Intermittent Demand for Service Parts Inventories. *Int. J. Forecast.* **2004**, *20*, 375–387. [CrossRef]
27. Syntetos, A.A.; Boylan, J.E. The Accuracy of Intermittent Demand Estimates. *Int. J. Forecast.* **2005**, *21*, 303–314. [CrossRef]
28. Jiang, A.; Tam, K.L.; Guo, X.; Zhang, Y. A New Approach to Forecasting Intermittent Demand Based on the Mixed Zero-truncated Poisson Model. *J. Forecast.* **2019**, *39*, 69–83. [CrossRef]
29. Babai, M.Z.; Chen, H.; Syntetos, A.A.; Lengu, D. A Compound-Poisson Bayesian Approach for Spare Parts Inventory Forecasting. *Int. J. Prod. Econ.* **2021**, *232*, 107954. [CrossRef]
30. Tian, X.; Wang, H.; Erjiang, E. Forecasting Intermittent Demand for Inventory Management by Retailers: A New Approach. *J. Retail. Consum. Serv.* **2021**, *62*, 102662. [CrossRef]
31. Snyder, R.D.; Ord, J.K.; Beaumont, A. Forecasting the Intermittent Demand for Slow-Moving Inventories: A Modelling Approach. *Int. J. Forecast.* **2012**, *28*, 485–496. [CrossRef]
32. Yang, K.; Wang, Y.; Fan, S.; Mosleh, A. Multi-criteria Spare Parts Classification Using the Deep Convolutional Neural Network Method. *Appl. Sci.* **2021**, *11*, 7088. [CrossRef]
33. Baisariyev, M.; Bakytzhanuly, A.; Serik, Y.; Mukhanova, B.; Babai, M.Z.; Tsakalerou, M.; Papadopoulos, C.T. Demand Forecasting Methods for Spare Parts Logistics for Aviation: A Real-World Implementation of the Bootstrap Method. *Procedia Manuf.* **2021**, *55*, 500–506. [CrossRef]
34. Ye, Y.; Lu, Y.; Robinson, P.; Narayanan, A. An Empirical Bayes Approach to Incorporating Demand Intermittency and Irregularity into Inventory Control. *Eur. J. Oper. Res.* **2022**, *303*, 255–272. [CrossRef]
35. Ahmadov, Y.; Helo, P. Deep Learning-Based Approach for Forecasting Intermittent Online Sales. *Discov. Artif. Intell.* **2023**, *3*, 45. [CrossRef]
36. Wang, S.; Kang, Y.; Petropoulos, F. Combining Probabilistic Forecasts of Intermittent Demand. *Eur. J. Oper. Res.* **2024**, *315*, 1038–1048. [CrossRef]
37. Fernández, A.; García, S.; Herrera, F.; Chawla, N.V. SMOTE for Learning from Imbalanced Data: Progress and Challenges, Marking the 15-Year Anniversary. *J. Artif. Intell. Res.* **2018**, *61*, 863–905. [CrossRef]
38. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-Sampling Technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [CrossRef]
39. Tang, Y.; Zhang, Y.Q.; Chawla, N.V. SVMs Modeling for Highly Imbalanced Classification. *IEEE Trans. Syst. Man. Cybern. Part. B Cybern.* **2009**, *39*, 281–288. [CrossRef]
40. Li, H.; Hong, L.; Mo, Y.; Zhu, B.; Chang, P. Restructuring Performance Prediction with a Rebalanced and Clustered Support Vector Machine. *J. Forecast.* **2018**, *37*, 437–456. [CrossRef]
41. Prado, E.M.G.; de Souza Filho, C.R.; Carranza, E.J.M.; Motta, J.G. Modeling of Cu-Au Prospectivity in the Carajás Mineral Province (Brazil) through Machine Learning: Dealing with Imbalanced Training Data. *Ore Geol. Rev.* **2020**, *124*, 103611. [CrossRef]
42. Feng, S.; Keung, J.; Yu, X.; Xiao, Y.; Zhang, M. Investigation on the Stability of SMOTE-Based Oversampling Techniques in Software Defect Prediction. *Inf. Softw. Technol.* **2021**, *139*, 106662. [CrossRef]
43. Zhen, X.; Ning, Y.; Du, W.; Huang, Y.; Vinnem, J.E. An Interpretable and Augmented Machine-Learning Approach for Causation Analysis of Major Accident Risk Indicators in the Offshore Petroleum Industry. *Process Saf. Environ. Prot.* **2023**, *173*, 922–933. [CrossRef]
44. Zhang, R.; Qi, Y.; Kong, S.; Wang, X.; Li, M. A Hybrid Artificial Intelligence Algorithm for Fault Diagnosis of Hot Rolled Strip Crown Imbalance. *Eng. Appl. Artif. Intell.* **2024**, *130*, 107763. [CrossRef]
45. Qazi, N.; Raza, K. Effect of Feature Selection, Synthetic Minority Over-Sampling (SMOTE) and under-Sampling on Class Imbalance Classification. In Proceedings of the 2012 14th International Conference on Modelling and Simulation, UKSim, Cambridge, UK, 28–30 March 2012; pp. 145–150.
46. Ullah, I.; Mahmoud, Q.H. A Two-Level Hybrid Model for Anomalous Activity Detection in IoT Networks. In Proceedings of the 2019 16th IEEE Annual Consumer Communications & Networking Conference (CCNC), Las Vegas, NV, USA, 11–14 January 2019; IEEE: New York, NY, USA; pp. 1–6.
47. Karatas, G.; Demir, O.; Sahingoz, O.K. Increasing the Performance of Machine Learning-Based IDSs on an Imbalanced and Up-to-Date Dataset. *IEEE Access* **2020**, *8*, 32150–32162. [CrossRef]
48. Dong, H.; He, D.; Wang, F. SMOTE-XGBoost Using Tree Parzen Estimator Optimization for Copper Flotation Method Classification. *Powder Technol.* **2020**, *375*, 174–181. [CrossRef]
49. Zhang, H.; Huang, L.; Wu, C.Q.; Li, Z. An Effective Convolutional Neural Network Based on SMOTE and Gaussian Mixture Model for Intrusion Detection in Imbalanced Dataset. *Comput. Netw.* **2020**, *177*, 107315. [CrossRef]
50. Gamel, S.A.; Ghoneim, S.S.M.; Sultan, Y.A. Improving the Accuracy of Diagnostic Predictions for Power Transformers by Employing a Hybrid Approach Combining SMOTE and DNN. *Comput. Electr. Eng.* **2024**, *117*, 109232. [CrossRef]

51. Xiao, P.; Liu, Z.; Zhao, G.; Pan, P. Novel Stacking Models Based on SMOTE for the Prediction of Rockburst Grades at Four Deep Gold Mines. *Undergr. Space* **2024**, *19*, 169–188. [CrossRef]
52. Chen, J.; Huang, H.; Cohn, A.G.; Zhang, D.; Zhou, M. Machine Learning-Based Classification of Rock Discontinuity Trace: SMOTE Oversampling Integrated with GBT Ensemble Learning. *Int. J. Min. Sci. Technol.* **2022**, *32*, 309–322. [CrossRef]
53. Ibrahim, B.; Ahenkorah, I.; Ewusi, A.; Majeed, F. A Novel XRF-Based Lithological Classification in the Tarkwaian Paleo Placer Formation Using SMOTE-XGBoost. *J. Geochem. Explor.* **2023**, *245*, 107147. [CrossRef]
54. Bradley, J.; Rajendran, S. Developing Predictive Models for Early Detection of Intervertebral Disc Degeneration Risk. *Healthc. Anal.* **2022**, *2*, 100054. [CrossRef]
55. Chen, X.; Zhao, Z.; Guo, F.; Tan, S.; Wang, J. Diagnosis Method of Transformer Winding Mechanical Deformation Fault Based on Sliding Correlation of FRA and Series Transfer Learning. *Electr. Power Syst. Res.* **2024**, *229*, 110173. [CrossRef]
56. Chen, W.; Zhang, L. Building Vulnerability Assessment in Seismic Areas Using Ensemble Learning: A Nepal Case Study. *J. Clean. Prod.* **2022**, *350*, 131418. [CrossRef]
57. Nguyen, Q.D.; Thai, H.T. Crack Segmentation of Imbalanced Data: The Role of Loss Functions. *Eng. Struct.* **2023**, *297*, 116988. [CrossRef]
58. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 318–327. [CrossRef] [PubMed]
59. Pasupa, K.; Vatathanavaro, S.; Tungjitnob, S. Convolutional Neural Networks Based Focal Loss for Class Imbalance Problem: A Case Study of Canine Red Blood Cells Morphology Classification. *J. Ambient. Intell. Humaniz. Comput.* **2023**, *14*, 15259–15275. [CrossRef]
60. Roy, S.; Tyagi, M.; Bansal, V.; Jain, V. SVD-CLAHE Boosting and Balanced Loss Function for COVID-19 Detection from an Imbalanced Chest X-Ray Dataset. *Comput. Biol. Med.* **2022**, *150*, 106092. [CrossRef]
61. Gökkan, O.; Kuntalp, M. A New Imbalance-Aware Loss Function to Be Used in a Deep Neural Network for Colorectal Polyp Segmentation. *Comput. Biol. Med.* **2022**, *151*, 106205. [CrossRef]
62. You, X.-X.; Liang, Z.-M.; Wang, Y.-Q.; Zhang, H. A Study on Loss Function against Data Imbalance in Deep Learning Correction of Precipitation Forecasts. *Atmos. Res.* **2023**, *281*, 106500. [CrossRef]
63. Büttner, M.; Schneider, L.; Krasowski, A.; Pitchika, V.; Krois, J.; Meyer-Lueckel, H.; Schwendicke, F. Conquering Class Imbalances in Deep Learning-Based Segmentation of Dental Radiographs with Different Loss Functions. *J. Dent.* **2024**, *148*, 105063. [CrossRef] [PubMed]
64. Mushava, J.; Murray, M. A Novel XGBoost Extension for Credit Scoring Class-Imbalanced Data Combining a Generalized Extreme Value Link and a Modified Focal Loss Function. *Expert Syst. Appl.* **2022**, *202*, 117233. [CrossRef]
65. Hu, C.; Quan, Z.; Chong, W.F. Imbalanced Learning for Insurance Using Modified Loss Functions in Tree-Based Models. *Insur. Math. Econ.* **2022**, *106*, 13–32. [CrossRef]
66. Wirth, R.; Hipp, J. CRISP-DM: Towards a Standard Process Model for Data Mining. In Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, Manchester, UK, 11–13 April 2000.
67. Kim, Y.-S.; Kim, M.K.; Fu, N.; Liu, J.; Wang, J.; Srebric, J. Investigating the Impact of Data Normalization Methods on Predicting Electricity Consumption in a Building Using Different Artificial Neural Network Models. *Sustain. Cities Soc.* **2024**, *118*, 105570. [CrossRef]
68. Pedregosa, F.; Michel, V.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Vanderplas, J.; Cournapeau, D.; Pedregosa, F.; Varoquaux, G.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
69. Kaya, M.; Çetin-Kaya, Y. A Novel Ensemble Learning Framework Based on a Genetic Algorithm for the Classification of Pneumonia. *Eng. Appl. Artif. Intell.* **2024**, *133*, 108494. [CrossRef]
70. Fawcett, T. An Introduction to ROC Analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [CrossRef]
71. Gifford, M.; Bayrak, T. A Predictive Analytics Model for Forecasting Outcomes in the National Football League Games Using Decision Tree and Logistic Regression. *Decis. Anal. J.* **2023**, *8*, 100296. [CrossRef]
72. Joseph, R.V.; Mohanty, A.; Tyagi, S.; Mishra, S.; Satapathy, S.K.; Mohanty, S.N. A Hybrid Deep Learning Framework with CNN and Bi-Directional LSTM for Store Item Demand Forecasting. *Comput. Electr. Eng.* **2022**, *103*, 108358. [CrossRef]
73. Chandriah, K.K.; Naraganahalli, R.V. RNN/LSTM with Modified Adam Optimizer in Deep Learning Approach for Automobile Spare Parts Demand Forecasting. *Multimed. Tools Appl.* **2021**, *80*, 26145–26159. [CrossRef]
74. Akyuz, A.O.; Uysal, M.; Bulbul, B.A.; Uysal, M.O. Ensemble Approach for Time Series Analysis in Demand Forecasting: Ensemble Learning. In Proceedings of the 2017 IEEE International Conference on INnovations in Intelligent Systems and Applications (INISTA), Gdynia, Poland, 3–5 July 2017; IEEE: New York, NY, USA; pp. 7–12.
75. Wu, H.; Levinson, D. The Ensemble Approach to Forecasting: A Review and Synthesis. *Transp. Res. Part. C Emerg. Technol.* **2021**, *132*, 103357. [CrossRef]
76. Jordan, M.; Kleinberg, J.; Schölkopf, B. *Pattern Recognition and Machine Learning*; Springer: Berlin/Heidelberg, Germany, 2006.

77. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 52.
78. Raja Santhi, A.; Muthuswamy, P. Pandemic, War, Natural Calamities, and Sustainability: Industry 4.0 Technologies to Overcome Traditional and Contemporary Supply Chain Challenges. *Logistics* **2022**, *6*, 81. [CrossRef]