Defence Science Journal, Vol. 73, No. 6, November 2023, pp. 666-674, DOI : 10.14429/dsj.73.6.19129
© 2023, DESIDOC

A Predictive Model with Data Scaling Methodologies for Forecasting Spare Parts Demand in Military Logistics

Jae-Dong Kim[#], Ji-Hwan Hwang[$] and Hyoung-Ho Doh[!,*]

[#]*Korea Institute for Defence Analyses (KIDA), Seoul, Republic of Korea*
[$]*Defence Agency For Technology and Quality (DTaQ), Busan, Republic of Korea*
[!]*Department of Industrial and Management Engineering, Hannam University, Daejeon, Republic of Korea*
[*]*E-mail: hojin81@gmail.com*

ABSTRACT

This study addresses the challenge of accurately forecasting demand for maintenance-related spare parts of the K-X tank, influenced by high uncertainty and external factors. Deep learning models with RobustScaler demonstrate significant improvements, achieving an accuracy of 86.90% compared to previous methods. RobustScaler outperforms other scaling models, enhancing machine learning performance across time series and data mining. By collecting eight years' worth of demand data and utilising various consumption data items, this study develops accurate forecasting models that contribute to the advancement of spare parts demand forecasting. The results highlight the effectiveness of the proposed approach, showcasing its superiority in accuracy, precision, recall, and F1-Score. RobustScaler particularly excels in time series analysis, further emphasizing its potential for enhancing machine learning performance on diverse datasets. This study provides innovative techniques and insights, demonstrating the effectiveness of deep learning and data scaling methodologies in improving forecasting accuracy for maintenance spare parts demand.

Keywords: Spare parts; Demand forecasting; Deep learning; Data scaling learning model; Military logistics

1. INTRODUCTION

Demand forecasting is crucial in operations management, serving as the basis for inventory planning[1]. Accurate forecasts are vital for maintaining optimal inventory levels and avoiding negative outcomes like overstocking, understocking, obsolescence, rush orders, and inefficiencies[2]. In the military logistics sector, demand forecasting is even more critical, as inaccurate forecasts can have severe consequences, potentially leading to a national crisis. To enhance demand forecasting accuracy in military logistics, big data analysis technologies, including data mining, are being adopted. These approaches offer new possibilities for improving spare parts demand forecasting.

This study introduces a predictive model that improves spare parts demand forecasting in military logistics by incorporating exogenous variables, including text data. The objective is to develop a forecasting model that identifies critical factors influencing spare parts demand, enhances accuracy, and enables effective inventory management. The proposed model has implications beyond military logistics, benefiting sectors like supply chain management and retail. By leveraging big data analytics and incorporating exogenous variables, the model delivers accurate demand forecasts, mitigating forecasting risks. The study provides valuable insights into spare parts demand in military logistics, facilitating the development of predictive models that enhance accuracy. Military organizations can maintain optimal inventory levels and prepare for future demand using these models. The study's findings contribute to the understanding of demand forecasting in various sectors, aiding in inventory management and preparedness.

The study follows a structured approach with a literature review in Section 2, providing a theoretical foundation and discussing forecasting approaches. Section 3 outlines the methodology, while Section 4 describes experimental studies. Section 5 concludes by presenting implications and future research directions, including deep learning techniques in spare parts demand forecasting.

2. LITERATURE REVIEW

2.1 Time Series

Time series analysis involves studying data points arranged in chronological order. Typically, these data points are collected consecutively and at equal intervals. The underlying assumption in time series analysis is that past patterns tend to repeat in the future. This analytical approach is valuable for understanding natural processes, monitoring changes over time, and evaluating the effectiveness of planned or unplanned interventions[3]. In the context of demand forecasting, time series analysis is widely applied and provides a reliable quantitative method for predicting future patterns based on historical data[4].

Researchers have utilised first- and second-order exponential smoothing techniques to estimate the number of

Received : 18 May 2023, Revised : 26 June 2023
Accepted : 11 July 2023, Online published : 03 November 2023

injuries and deaths resulting from road accidents in Jordan between 1981 and 2016. The accuracy of these results was compared to other methods based on mean absolute percent error, mean absolute deviation, and mean square deviation. The second-order exponential smoothing technique demonstrated notable effectiveness[5]. In the military domain, the demand forecasting model for repair units employs 5-8 time series techniques, with 3 techniques specifically utilised for the Army. South Korean troops, on the other hand, make use of 5-9 time series techniques. This approach involves comparing the actual demand in a given year (*t*) with the demand forecast from the previous year (*t*-1), employing a relatively straightforward demand forecasting technique as a baseline for comparison[6-8].

2.2 Data Mining

Data mining is a sophisticated approach that employs statistical and machine learning methods to uncover valuable patterns and insights from extensive data sets. It involves tasks like sourcing relevant data, preparing it by cleaning and organizing, and using algorithms to unveil hidden relationships and trends. This process aims to inform business decisions, enhance processes, and drive innovation across various domains, including customer segmentation, fraud detection, market basket analysis, and predictive modeling.

Data mining has wide-ranging applications across various fields such as healthcare, finance, marketing, and social media analytics. It is a powerful and continuously evolving technology that aids in decision-making and prediction, contributing to advancements in multiple domains[9]. The process involves extracting valuable insights and uncovering relationships from extensive datasets through thorough investigation, analysis, and modeling[10]. Several commonly used data mining techniques include decision trees (DTs), Bayesian networks, and support vector machines (SVMs). Decision trees analyse information based on decision-making rules, organizing it into tree structures to classify relevant classes into multiple categories or predict results[11].

2.3 Deep Learning

Deep learning refers to the use of deep neural networks, which are artificial neural networks with two or more hidden layers, along with algorithms designed for learning these networks. One such algorithm is the Gated Recurrent Unit (GRU), which addresses the issue of gradient vanishing while retaining the benefits of long short-term memory (LSTM)[12]. In a related study, weather data consisting of 12 variables, including hourly photovoltaic power, was obtained from the European medium-range weather forecast center for a specific time period[13].

Deep learning utilises neural networks that mimic the structure of human neural networks and comprise interconnected layers[14-15]. A Multi-Layer Perceptron (MLP) is a type of artificial neural network that consists of multiple layers of single-layer perceptron networks. MLP overcomes the limitations of single-layer perceptrons, which can only handle linearly separable data[16-17]. To enhance the performance of MLP, recurrent neural networks (RNNs) with cyclic structures have been developed. These networks store and transmit information within memory blocks based on predefined schedules. RNNs are particularly suitable for processing sequential data such as text, speech, and time-series data[18]. By incorporating the LSTM algorithm, RNNs can effectively address the issue of long-term dependencies and have found applications across various domains[19]. Recently, Méndez[20], *et al.* reviews 155 scientific publications from 2011-2021 on Deep Learning models for air quality forecasting, covering geography, predicted values, predictor variables, evaluation metrics, and ML models. Also, Abbasimehr[21], *et al.* proposes a demand forecasting method based on multi-layer LSTM networks to address the challenges of accurate prediction in competitive business environments. By conducting experiments and comparing with other methods, the results demonstrate the superior performance of the proposed method in capturing fluctuating demand data. Finally, Xie, *et al.*[22] presents a decomposition-ensemble approach utilising empirical mode decomposition, data characteristic analysis, and Elman's neural network model to enhance the accuracy of tourism demand forecasting in volatile regions like Hong Kong. Through empirical analysis using Hong Kong tourism demand data, the proposed model demonstrates superior performance in both point and interval forecasts for various prediction horizons, highlighting its effectiveness in predicting complex time series data.

Deep learning methodologies have gained traction in demand forecasting, surpassing traditional time series analysis and data mining techniques. They excel in capturing complex non-linear relationships, handling sequential data, automating feature extraction, and effectively managing outliers. While demanding substantial data and computational resources and posing challenges in interpretability, deep learning exhibits promise in enhancing forecasting accuracy and addressing complex patterns.

3. PROPOSAL OF DEMAND FORECASTING MODEL

This study focuses on collecting operational data from the Defense Logistics Integrated Information System (DELIIS) of the Ministry of National Defense of the Republic of Korea. The target of the study is the maintenance data of K-X, an army tank crucial for battlefield operations. The objective is to enhance its superiority and operational potential in times of war.

3.1 Extraction of Variables

This research analyses transaction records to forecast spare parts demand. It categorises 18,476 identified items and extracts data such as consumption, procurement period, unit price, operating time, and maintenance comments. The study focuses on the annual demand data for the K-X tank system, obtained from military maintenance records. Derived variables totaling 32 factors are derived by summing categorical counts of spare parts used between 2010 and 2018. The objective variable is the occurrence of spare parts in 2018, with 9,527 items. To ensure balance, 5,813 unused parts are randomly selected, and 7,428 items are used for experiments. The study compares the predictive performance of time series analysis, data mining, and deep learning techniques in forecasting spare parts usage based on the extracted variables. Table 1 presents an example of extracted data, while Annexure I shows the original data before extraction. In the future, when text mining techniques are organised and available, several potential external variables that can be used are presented.

Table 1. Variables descriptions

Variables (Number of the unit)	Description
Number of Consumption (8)	Sum of spare parts consumed per item by year (2010 ~2018)
Order Time (8)	Sum of spare parts order times per item by year (2010 ~2018)
Operating Time (8)	Tank operating time
Operating Distance (8)	Tank operating distance

3.2 Proposed Data Scaling Learning System

The proposed data scaling learning system consists of a basic level learning module and a data scaling learning module. The base level learning module was applied to the unscaled dataset. 5-fold cross-validation was performed to prevent overfitting when constructing the dataset. Then, as shown in Fig. 1, the dataset was trained using the Data Scaling Train module to calculate the final result.

3.2.1 Base-Level Learning Module

In our study, we employed various models for time series analysis, machine learning, and deep learning. The time series models used were Arithmetic Moving (AM), Simple Moving Average (SMA), Weighted Moving Average (WMA), Least Moving Average (LMA), and Least Square (LS). The machine learning models included DT, RF, Naive Bayes (NB), Logistic Regression (LR), and SVM. The deep learning models consisted of RNN, LSTM, GRU, and MLP. Our study employed various models for time series analysis, machine learning, and deep learning. Time series models (AM, SMA, WMA, LMA, LS) capture historical patterns, while machine learning models (DT, RF, NB, LR, SVM) handle complex relationships. Deep learning models (RNN, LSTM, GRU, MLP) excel at sequential data and intricate dependencies. This approach ensures comprehensive analysis and accurate demand forecasting.

The GRU, a type of RNN, is designed for sequential data processing and shares similarities with LSTM networks. It incorporates two gates: the reset gate and the update gate, regulating information flow. The reset gate controls forgetting of the previous state, while the update gate determines how much new input is added. The output of the GRU is obtained by applying the hyperbolic tangent activation function to the linear combination of the new candidate state and the current input, weighted by the update gate. This mechanism is valuable for handling sequential data with long-term dependencies, allowing selective retention or omission of previous inputs based on current inputs. For a graphical representation, refer to Fig. 2.

It is important to note that the base-level learning module in our study did not include data scaling learning methods. However, by utilising the diverse range of models mentioned above, we aimed to explore and evaluate their performance in various domains of analysis, from time series to machine learning and deep learning.

3.2.2 Data Scaling Learning Module

The data scaling method used in this study is an important preprocessing step in many machine learning algorithms. This means transforming the data so that all features have the same scale. This is important because some machine learning algorithms are sensitive to measures of input characteristics.

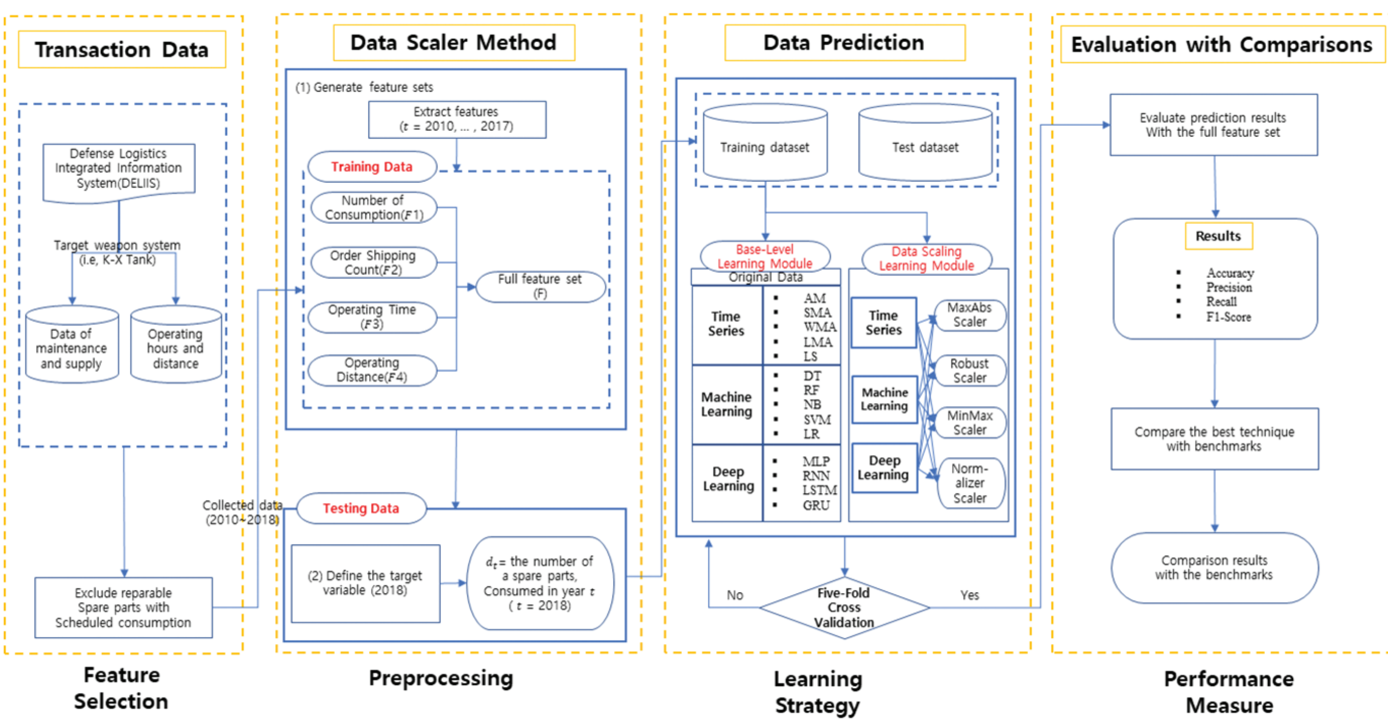

Figure 1. Overview of demand forecasting mechanism based on the data scaling learning model.

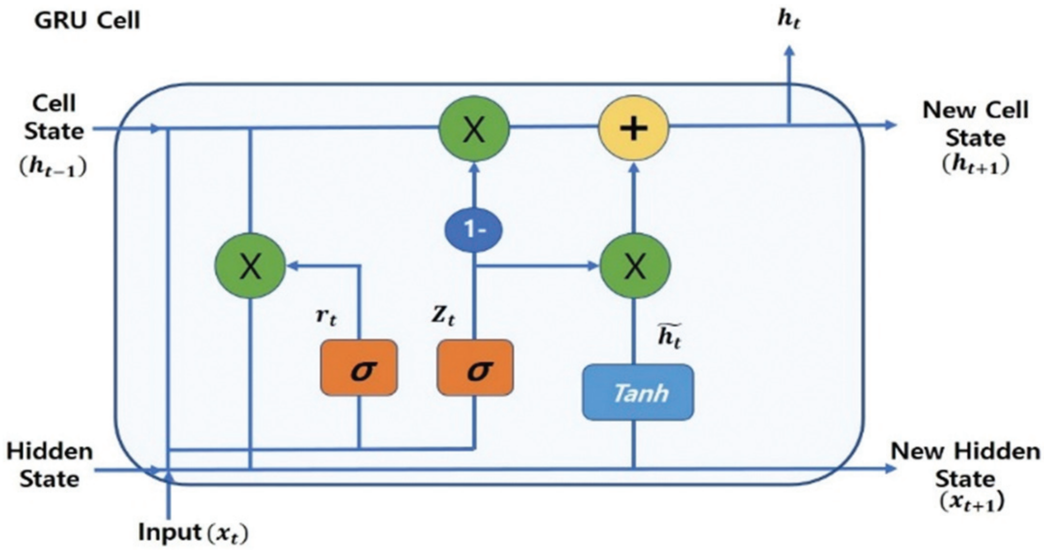

Figure 2. Overview of GRU mechanism.

Using out-of-scale data may result in less-than-optimal performance. In this paper, four data scaling methods are considered: StandardScaler, RobustScaler, MinMaxScaler and MaxAbsScaler[23].

- StandardScaler: This method scales the data so that it has zero mean and unit variance. This means that each feature will have a mean of zero and a standard deviation of one. The formula for standardisation is (*x*-*mean*)/*std*. This is the most commonly used scaling method and works well when the data is normally distributed
- RobustScaler: This method scales the data using the interquartile range (IQR) instead of the mean and standard deviation. This makes it more robust to outliers in the data. The formula for RobustScaler is (*x*-*mean*)/*IQR*
- MinMaxScaler: This method scales the data so that it is between a specified minimum and maximum value, usually 0 and 1. The formula for MinMaxScaler is (*x*-*min*)/(*max*-*min*)
- MaxAbsScaler: This method scales the data so that the absolute values of each feature are between 0 and 1. The formula for MaxAbsScaler is *x*/*max*(*abs*(*x*)).

Considering the distribution of data and the presence of outliers is important when selecting a scaling method. StandardScaler can be a suitable default choice, but RobustScaler might be more effective in the presence of outliers.

4. PROPOSAL OF DEMAND FORECASTING MODEL

4.1 Experiment Design

To capture the time series pattern in the data, the study split the training data by year. Input variables from 2010 to 2017 were used to train the model for forecasting spare parts demand in 2018. For evaluation, spare parts demand in 2017 was predicted using input variables from 2011 to 2016. Performance assessment employed classification metrics like accuracy, precision, recall, and F1-score.

4.2 Classification Results of the Base Model

The study utilised a Base model that integrated time series, machine learning, and deep learning techniques to develop a predictive model for spare parts demand forecasting. Historical data spanning six years was used, and a binary representation (0 or 1) was employed to indicate the occurrence or non-occurrence of spare parts. Cross-validation with five randomly classified folds was conducted to evaluate the model.

Table 2. Performance values of base models.

	Original data					
		In per cent		**In per cent**		**In per cent**
Accuracy	AM	**69.50**	DT	82.00	MLP	**84.30**
	SMA	68.60	RF	**84.50**	RNN	82.40
	WMA	51.30	NB	65.80	LSTM	81.00
	LMA	64.70	SVM	81.50	GRU	79.90
	LS	52.30	LR	82.20		
Precision	AM	97.00	DT	82.00	MLP	**82.00**
	SMA	97.00	RF	80.00	RNN	73.00
	WMA	**99.00**	NB	32.00	LSTM	68.00
	LMA	96.00	SVM	**90.00**	GRU	65.00
	LS	97.00	LR	68.00		
Recall	AM	**63.00**	DT	82.00	MLP	86.00
	SMA	62.00	RF	88.00	RNN	90.00
	WMA	51.00	NB	**99.00**	LSTM	92.00
	LMA	59.00	SVM	77.00	GRU	**93.00**
	LS	51.00	LR	95.00		
F1-Score	AM	**76.00**	DT	82.00	MLP	**84.00**
	SMA	75.00	RF	**84.00**	RNN	81.00
	WMA	67.00	NB	48.00	LSTM	78.00
	LMA	73.00	SVM	83.00	GRU	76.00
	LS	67.00	LR	79.00		

The results presented in Table 2 showed that the machine learning-based model outperformed the traditional univariate time series model. Among the models, RF achieved the highest accuracy score of 84.5 per cent, closely followed by the MLP model at 84.3 per cent. Other models like DT and LR also performed well with accuracy scores of 82.0 per cent and 82.2 per cent, respectively. On the other hand, the Worst Model (WM) had a poor accuracy score of 51.3 per cent.

In terms of precision, the WM model obtained the highest score of 99.0 per cent, while the NB model had the lowest score of 32.0 per cent. The RF and MLP models achieved precision scores of 80.0 per cent and 82.0 per cent, respectively. The RNN and GRU models displayed the lowest precision scores at 73.0 per cent and 65.0 per cent, respectively. For recall, the GRU model performed the best with a score of 93.0 per cent, followed by the LSTM model at 92.0 per cent, and the RNN model at 90.0 per cent.

Table 3. Performance values of data scaling models

(a) Result with time seires

		MaxAbsScaler	RobustScaler	MinMaxScaler	StandardScaler
		In per cent	In per cent	In per cent	In per cent
Accuracy	AM	50.10	78.80	50.10	50.00
	SMA	50.10	78.90	50.10	50.00
	WMA	**54.10**	**84.30**	**54.10**	51.50
	LMA	50.20	81.00	50.20	**66.80**
	LS	50.00	49.90	50.00	50.00
Precision	AM	0.00	61.00	0.00	0.00
	SMA	0.00	61.00	0.00	0.00
	WMA	8.00	85.00	8.00	99.00
	LMA	0.00	67.00	0.00	73.00
	LS	**100.00**	**99.00**	**100.00**	**100.00**
Recall	AM	**100.00**	95.00	**100.00**	0.00
	SMA	100.00	**96.00**	100.00	0.00
	WMA	100.00	84.00	100.00	51.00
	LMA	94.00	93.00	94.00	**65.00**
	LS	50.00	50.00	50.00	50.00
F1-Score	AM	0.00	74.00	0.00	0.00
	SMA	0.00	74.00	0.00	0.00
	WMA	15.00	**84.00**	15.00	67.00
	LMA	1.00	78.00	1.00	**69.00**
	LS	**67.00**	66.00	**67.00**	67.00

(b) Result with data mining

		MaxAbsScaler	RobustScaler	MinMaxScaler	StandardScaler
		In per cent	In per cent	In per cent	In per cent
Accuracy	DT	80.10	78.60	80.10	80.10
	RF	**84.00**	82.60	**84.00**	**84.00**
	NB	61.10	65.30	61.10	81.70
	SVM	50.90	**84.30**	50.90	83.60
	LR	58.50	82.10	58.50	83.00
Precision	DT	78.00	83.00	78.00	80.00
	RF	**79.00**	**84.00**	**79.00**	79.00
	NB	22.00	31.00	22.00	81.00
	SVM	2.00	78.00	2.00	86.00
	LR	17.00	68.00	17.00	**87.00**
Recall	DT	81.00	76.00	81.00	80.00
	RF	88.00	82.00	88.00	**88.00**
	NB	100.00	**99.00**	**100.00**	82.00
	SVM	100.00	89.00	**100.00**	82.00
	LR	**100.00**	95.00	**100.00**	81.00
F1-Score	DT	80.00	79.00	80.00	80.00
	RF	**83.00**	83.00	**83.00**	83.00
	NB	36.00	47.00	36.00	82.00
	SVM	4.00	**83.00**	4.00	84.00
	LR	29.00	79.00	29.00	**84.00**

Table 3. Performance values of data scaling models

(c) Result with deep learning

		MaxAbsScaler	**RobustScaler**	**MinMaxScaler**	**StandardScaler**
		In per cent	**In per cent**	**In per cent**	**In per cent**
Accuracy	MLP	**84.70**	83.40	**84.30**	81.60
	RNN	82.10	84.60	82.10	83.20
	LSTM	80.30	86.10	74.60	83.50
	GRU	76.90	**86.90**	79.50	**84.10**
Precision	MLP	**77.00**	82.00	**81.00**	83.00
	RNN	74.00	79.00	72.00	81.00
	LSTM	48.00	**83.00**	51.00	79.00
	GRU	61.00	**83.00**	64.00	**84.00**
Recall	MLP	90.00	84.00	86.00	81.00
	RNN	89.00	**89.00**	90.00	85.00
	LSTM	**96.00**	88.00	**96.00**	**87.00**
	GRU	94.00	89.00	93.00	84.00
F1-Score	MLP	**83.00**	83.00	**84.00**	82.00
	RNN	81.00	84.00	80.00	83.00
	LSTM	64.00	86.00	67.00	83.00
	GRU	74.00	**86.00**	76.00	**84.00**

Regarding the F1-score metric, the RF and MLP models achieved the highest scores of 84.0 per cent and 84.00 per cent, respectively, while the NB model had the lowest score of 48.0 per cent. The LSTM, SVM, and GRU models achieved F1-scores of 78.0 per cent, 83.0 per cent, and 76.0 per cent, respectively.

Overall, the RF and MLP models exhibited strong performance across multiple metrics, while the NB and WM models performed poorly. The GRU model showed high recall but lower precision and F1-score values. Notably, the RF demonstrated favorable overall performance in terms of accuracy and F1-score.

4.3 Classification Results Obtained Using the Data Scaling Learning Model

The study employed various scaling techniques, including MaxAbsScaler, RobustScaler, MinMaxScaler, and StandardScaler, across three different types of models: time series, machine learning, and deep learning. Performance evaluation metrics such as accuracy, precision, recall, and F1-score were utilized, and the results are presented in Table 3.

RobustScaler emerged as the top-performing scaling model, consistently demonstrating strong performance across time series, data mining, and deep learning. In the realm of time series, RobustScaler outperformed other scaling models in terms of accuracy, precision, recall, and F1-score. Notably, it achieved the highest accuracy in the WMA category and precision scores for the LS model consistently approached 99-100 per cent. Additionally, RobustScaler, MaxAbsScaler, and MinMaxScaler achieved recall results ranging from 96 per cent to 100 per cent in the SM and AM categories. RobustScaler also delivered a satisfactory F1-score of 84 per cent in the WMA category. In data mining, the scaling models exhibited similar performance levels in terms of accuracy, precision, recall, and F1-score. RF and SVM consistently attained comparable accuracy scores across different scaling models. Precision scores for RF ranged from 79 per cent to 84 per cent, and MaxAbsScaler and MinMaxScaler yielded recall results of 99 per cent to 100 per cent in various models. F1-score results were also consistent among RF, SVM, and LR, ranging from 83 per cent to 84 per cent across multiple scaling models.

In the domain of deep learning, RobustScaler demonstrated superior performance compared to other scaling models across accuracy, precision, recall, and F1-score. The GRU model achieved the highest accuracy of 86.9 per cent when combined with RobustScaler. Precision results ranged from 83 per cent to 84 per cent for GRU with both RobustScaler and StandardScaler. MaxAbsScaler and MinMaxScaler achieved a recall of 96 per cent in the LSTM model. The best F1-score of 86 per cent was achieved by both RobustScaler and GRU.

The exceptional accuracy of RobustScaler can be attributed to its robustness to outliers. By using the median and interquartile range to scale each data point, RobustScaler ensures accurate scaling even in the presence of outliers. This characteristic makes RobustScaler highly valuable in datasets that contain outliers, ultimately enhancing model performance.

The results clearly indicate that in the context of military equipment, there are challenges such as missing data and imbalanced data distribution. To address these challenges and improve prediction accuracy, the use of data scaling models has been demonstrated in this study. Data scaling techniques, such as normalisation or standardisation, can help mitigate the impact of missing data and handle imbalanced datasets effectively. By applying data scaling, the features in the dataset are transformed to a common scale, reducing the influence of missing values and enabling a more balanced representation of the data. This normalisation process ensures that each feature contributes equally to the prediction model, avoiding biases caused by imbalanced data. Additionally, data scaling can enhance the convergence speed and stability of machine learning algorithms, improving the overall performance of the prediction model. The incorporation of data scaling models in this study highlights their effectiveness in dealing with the unique challenges of military equipment data, leading to more reliable and accurate predictions. The findings suggest that utilising data scaling techniques can be a valuable approach in similar domains where missing data and imbalanced distributions are prevalent.

5. CONCLUSIONS

This study addresses the challenge of demand forecasting for maintenance-related spare parts of the K-X tank, which poses high uncertainty due to their external nature. To improve accuracy in spare part demand forecasting, the study explores various analyses and modeling approaches.

To develop an accurate forecasting model, the study collects eight years' worth of demand data for K-X tank spare parts and extracts various consumption data items. Deep learning models with data scaling techniques are developed for solving demand forecasting problem. The best result is achieved using RobustScaler and the GRU model, with an accuracy of 86.90 per cent. The base model demonstrates accuracies of 84.5 per cent in data mining and 84.3 per cent in deep learning. When RobustScaler is applied, data mining maintains a similar accuracy to the base model (84.3 %), while deep learning shows a 2.6 per cent improvement, reaching an accuracy of 86.9 per cent. RobustScaler outperforms other scaling models in terms of accuracy, precision, recall, and F1-Score across time series, data mining, and deep learning. Particularly in time series analysis, RobustScaler achieves the highest accuracy and recall results, while also demonstrating excellent accuracy and F1-Score in deep learning. RobustScaler can enhance machine learning performance on diverse datasets, including those containing outliers. Future research can focus on developing advanced scaling techniques to further improve performance on datasets with various types of outliers.

The study highlights the potential for improved performance in the future, as evidenced by the M4 competition, where Convolutional Neural Network (CNN) models leveraging unstructured data showcased superior results. This emphasizes the growing trend of utilizing unstructured data for prediction and underscores the need for advanced model development.

REFERENCES

1. Boylan, J.E. & Syntetos, A.A. Spare parts management: a review of forecasting research and extensions. *IMA J. Manag. Math.*, 2010, **21**(3), 227-237. doi:10.1093/imaman/dpp016
2. Nenni, M.E.; Giustiniano, L. & Pirolo, L. Demand forecasting in the fashion industry: A review. *Int. J. Eng. Bus. Manag.*, 2013, **5**, 37. doi:10.5772/56840
3. Hamilton, J.D. Time Ssries analysis. Princeton: Princeton university press, USA, 2020.
4. Chatfield, C. Time-series forecasting. CRC Press, USA, 2000. doi:10.1201/9781420036206
5. Rahamneh, A. Using single and double exponential smoothing for estimating the number of injuries and fatalities resulted from traffic accidents in Jordan (1981-2016), *Middle East J. Sci. Res.*, 2017, **25**(7), 1544-1552. doi:10.5829/idosi.mejsr.2017.1544.1552
6. Kim, J.-D.; Lee, H.-J.; & Choi, S.-B. Machine learning based approach for demand forecasting anti-aircraft missiles, 2018 5th IEEE International Conference on Industrial Engineering and Applications (ICIEA), Singapore, 2018, 367-372. doi:10.1109/IEA.2018.8387126
7. Lee, H.-J., & Kim, J.-D. A predictive model for forecasting spare parts demand in military logistics, 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Bangkok, Thailand, 2018, 1106-1110. doi:10.1109/IEEM.2018.8607801
8. Kim, J.-D. Text mining-based approach for forecasting spare parts demand of K-X tanks. *In* 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Bangkok Thailand, 2018, 1652-1656. doi:10.1109/IEEM.2018.8607632
9. Liao, S.H.; Chu, P.H.; & Hsiao, P.Y. Data mining techniques and applications-A decade review from 2000 to 2011. *Expert Syst. Appl.*, 2012, **39**(12), 11303-11311. doi:10.1016/j.eswa.2012.02.063
10. Hand, D.J. Principles of data mining. *Drug Saf.*, 2007, **30**(7), 621-622. doi:10.2165/00002018-200730070-00010
11. Kass, G.V. An exploratory technique for investigating large quantities of categorical data. *J.R. STAT. SOC. C-A PPL.*, 1980, **29**(2), 119-127.
12. Chung, J.; Gulcehre, C.; Cho, K.; & Bengio, Y. Empirical evaluation of gated recurrent neural networks on sequence modeling. In NIPS 2014 Workshop on Deep Learning, December 2014. doi:10.48550/arXiv.1412.3555
13. Hong, T.; Pinson, P.; Fan, S.; Zareipour, H.; Troccoli, A. & Hyndman, R.J. Probabilistic energy forecasting: Global energy forecasting competition 2014 and beyond. *Int. J. Forecast.*, 2016, **32**(3), 896-913. doi:10.1016/j.ijforecast.2016.02.001
14. Bengio, Y.; Courville, A. & Vincent, P. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013, **35**(8), 1798-1828. doi:10.1109/tpami.2013.50
15. Schmidhuber, J. Deep learning in neural networks: An overview. *Neural Netw.*, 2014, **61**, 85–117. doi:10.1016/j.neunet.2014.09.003
16. Werbos, P.J. Backpropagation through time: What it does and how to do it. *Proc. IEEE*, 1990, **78**(10), 1550–1560.

doi:10.1109/5.58337

17. Kumpati, S.N. & Kannan, P. Identification and control of dynamical systems using neural networks. *IEEE Trans. Neural Netw.*, 1990, **1**(1), 4-27. doi:10.1109/72.80202
18. Mikolov, T.; Karafiát, M.; Burget, L.; Cernocký, J. & Khudanpur, S. Recurrent neural network based language model. *Interspeech*, 2010, **2**(3), 1045–1048.
19. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.*, 1997, **9**(8), 1735–1780. doi:10.1162/neco.1997.9.8.1735
20. Méndez, M.; Merayo, M.G. & Núñez, M. Machine learning algorithms to forecast air quality: A survey. *Artif. Intell. Rev*., 2023. doi.org/10.1007/s10462-023-10424-4
21. Abbasimehr, H.; Shabani, M. & Yousefi, M. An optimised model using LSTM network for demand forecasting. *Comput. Ind. Eng.*, 2020, **143**, 106435. doi:10.1016/j.cie.2020.106435.
22. Xie, G.; Qian, Y. & Wang, S. A decomposition-ensemble approach for tourism forecasting. *Ann. Tour. Res.*, 2020, **81**, 102891. doi:10.1016/j.annals.2020.102891
23. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O. & Duchesnay, E. Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.*, 2011, **12**, 2825-2830. doi:10.48550/arXiv.1201.0490

ACKNOWLEDGMENTS

This research was performed in the Defense Agency of Technology and Quality(DTaQ), and DTaQ verified that it did not contain any information related to military security.

CONTRIBUTORS

Dr Jae-Dong Kim is a Research Fellow at the Center for Defense Resource Management, The Korea Institute for Defense Analyses, Seoul, South Korea. He obtained his PhD from Korea University, School of Industrial and Management Engineering, South Korea. His areas of research include: Production planning, supply chain management, artificial intelligence, and data mining.

His contribution in the current study: Suggested the idea for the study, developed the framework and algorithm for the research, contributed to the implementation of the proposed approach, and played a significant role in the creation of the original manuscript.

Dr Ji-Hwan Hwang is a Research Fellow at the Naval system Team 4, the Defense Agency for Technology and Quality, Busan, South Korea. He obtained his PhD from Korea University, School of Electrical and Electronic Engineering, South Korea. His areas of research include: Control using artificial intelligence, data mining, signal processing.

His contribution in the current study: Conducted data collection and preprocessing, provided critical analysis and interpretation of the results, and participated in the revision and improvement of the manuscript.

Dr Hyoung-Ho Doh is a Research Fellow at the School of Industrial and Management Engineering College of Engineering, Hannam University, Daejeon, South Korea. He obtained his PhD from Hanyang University, Department of Industrial Engineering, South Korea. His areas of research include: Production planning, inventory control, supply chain management, and data mining.

His contribution in the current study: Conducted to the literature review and background research, assisted in experimental design and methodology, contributed to the writing and editing of the manuscript.

Annexure I

Examples of the transaction data

Equipment name	Equipment NSN	Equipment registration number	Maintenance date	NSN	Troops	Unit of issue	Unit exhausted	Number of consumption	Operator name	Number of operator	Recoverability code
Equipment A	A8111	2000A1	20130221	A92345	X Division	EA	EA	1	O O O	1	Y
Equipment B	B5111	2000B1	20130419	B92345	X LSC	EA	EA	1	O O O	1	Y
…	…	…	…	…	…	…	…	…	…	…	…
Equipment identification code	Represent identification code	QPA (Quantity Per Assembly)	Unit standard	The newest unit	PROLT (Procurement Lead Time)	OST (order Shipping Time)	Acquisition year	Maintenance enforcement year	Maintenance history		
N	Y	2	9000 Won	12500 Won	150 days	3 days	2002 year	2013 year	Equipment registration number: XX, inspector comments: the equipment was replaced due to emulsification and leakage of oil.		
N	Y	3	5700 Won	6900 Won	100 days	15 days	2004 year	2013 year	inspector comments: daily maintenance and repair parts replacement.		