Irregular Time-Series Hybrid Model for Enhanced Prognostics of Engine Degradation and Failures

Rohit Deo[1], Aman Yadav[2], Shruti Bharti[3], and Nilesh Powar[4]

[1,2,3,4] *Cummins Technologies India Pvt. Ltd., Pune, MH, 411045, India*

rohit.deo@cummins.com
aman.yadav@cummins.com
shruti.bharti@cummins.com
Nilesh.powar@cummins.com

Abstract

Classification-based prognostics aims to predict the Remaining Useful Life (RUL) of components in diesel engines by identifying failure and degradation stages. This is critical for industries such as automotive, aviation, and manufacturing. Traditional methods rely on classification models trained on historical data from multiple engines to forecast failures based on current engine parameters. However, these global classifiers often struggle with generalization when applied to unseen engines, resulting in poor precision and recall. Moreover, they fail to capture the temporal dependencies inherent in engine degradation, which are crucial for accurate failure prediction. We propose a hybrid model that integrates predictions from global classifiers with time-based memory units to address these limitations, effectively building irregular time-series models. Our approach demonstrates a significant performance improvement, with precision and recall metrics doubling compared to traditional global classifiers.
Keywords – Remaining Useful Life, Prognostics, Predictive Maintenance, Engine Health Monitoring, Downtime reduction, Long short-term memory (LSTM), Convolutional Neural Network (CNN), Recurrent NN (RNN).

1. Introduction

Cummins Inc. is a global corporation specializing in the design, manufacturing, and distribution of engines, filtration systems, and power generation products. Headquartered in Columbus, Indiana, Cummins has been operating since 1919 and serves customers across more than 190 countries and territories. The company emphasizes innovation and sustainability in its products and operations.

Prognostics play a vital role in Cummins' diesel engine operations by enabling predictive maintenance. Through the analysis of sensor data and other performance indicators, prognostics help anticipate when maintenance or repairs are required. This predictive approach allows for planned maintenance, reducing downtime and mitigating the risk of unexpected failures. By leveraging prognostics, Cummins enhances their diesel engines' performance, reliability, and lifespan.

Classification-based prognostics is a predictive maintenance approach where the goal is to estimate the Remaining Useful Life (RUL) of a component or system by categorizing it into predefined failure stages or health states. Instead of predicting an exact time-to-failure, it classifies the system's condition to assess how close it is to a failure event. The component's state is classified as Healthy, Degrading, or Failure Imminent rather than providing a specific time frame for failure. Models are trained on labelled datasets that contain failure histories and operating conditions from multiple components or systems. The dataset consists of diagnostic features derived from sensor data (e.g., temperature, pressure, vibration) to predict the health state. With this approach, it is easier for decision-making since the output is in discrete categories. It also helps us get early indications of degradation stages.

The models predicting the aforementioned discrete categories are referred to as global classifiers. These machine-learning (ML) models utilize historical data from multiple engines. An unseen engine is one whose historical data is available but may not have been included in training the global classifier.

During inference, predictions for an unseen engine rely solely on the current parameter values, reflecting the engine's present condition. Each feature vector in the training set is labelled with discrete categories. Consequently, predictions are based on how similar conditions affected other engines during training, without considering the specific engine's historical data. This approach ignores temporal dependencies in the engine's operational history.

To enhance prediction accuracy, combining global classifier predictions with a memory-based model that leverages historical data can be beneficial. This hybrid approach would incorporate both the engine’s current condition and its historical behavior, leading to more robust predictions. This paper describes how predictions from the global classifier can be effectively combined with those from an individual local memory-based model. We demonstrate that this blended approach enhances prediction accuracy for individual engines and quantifies the potential cost savings resulting from improved performance.

This work focuses on the prediction of component or system failures to enable timely maintenance before abrupt breakdowns occur. Unexpected failures in diesel engines can be severe, often resulting in increased repair costs and extended downtime. Given Cummins' commitment to delivering reliable performance across a wide range of applications, minimizing such disruptions is critical. When an engine experiences an unforeseen failure during operation, it can lead to significant inconvenience, operational delays, and financial losses for customers. These failures may involve key subsystems such as aftertreatment components, oil maintenance systems, and air handling units. Therefore, predictive maintenance strategies are essential to enhance reliability, reduce downtime, and optimize service planning.

The paper is organized as follows: Section 2 provides a literature survey. Section 3 elaborates on the proposed hybrid model, followed by results and discussions in section 4. Section 5 gives conclusions and future scope.

2. LITERATURE SURVEY

A new technology in the field of equipment maintenance and support viz.- Prognostics and health management (PHM) improves equipment reliability and safety which enhances equipment support maintenance support capabilities and reduces the maintenance support cost. PHM for diesel engines majorly involves four steps. Data acquisition, data processing, fault diagnosis and health status assessment (Liu et al., 2023). Data acquisition is the basis of the next steps. The acquired data through different sensors in diesel engines are raw signals that are of poor quality. To make the raw signals deterministic, we need data processing. Once processed, fault diagnosis can be done based on knowledge, signal processing or machine learning. And then comes the health status assessment which relies on data processing outcomes, utilizing failure models or intelligent algorithms to evaluate the operational condition of the diesel engine. Later, the research focused on expanding failure studies across all diesel engine subsystems, investigating performance under complex operating conditions, and improving remaining life prediction (RUL) and maintenance decision-making for real-world applications. Machine learning has revolutionized Remaining Useful Life (RUL) prediction for diesel engines by analyzing sensor data and failure patterns to identify wear and degradation trends. These models enable proactive maintenance by forecasting potential failures before they occur. A variety of models are used for RUL prediction in diesel engines, including regression models (e.g., Linear Regression), time-series models (e.g., Long Short-Term Memory), ensemble methods (e.g., Random Forests), deep learning models (e.g., Convolutional Neural Networks), and hybrid physics-informed approaches. The model choice depends on the data's complexity, operating conditions, and failure patterns.

To enable time series classification, (Karim, Majumdar, Darabi, & Harford, 2019) transformed univariate series into multivariate using a fully convolutional block enhanced with a squeeze-and-excitation (SE) block (J. Hu, Shen, & Sun, 2018), which boosts neural networks by emphasizing key features via channel-wise attention. This SE mechanism was adapted for 1D models, improving LSTM-FCN and Attention LSTM-FCN architectures. In financial time series prediction—complex like diesel engine prognostics—(Alhnaity & Abbod, 2020) proposed a hybrid model combining empirical mode decomposition, neural networks, and support vector regression, optimized with genetic algorithms to reduce errors. Similarly, (Zhao, Li, Xu, Fu, & Chen, 2023) used AdaBoost, KNN, and LSTM for financial data classification.

In (W. Hu & Shi, 2020), combining LSTM with Random Forest (LSTM-RF) improves prediction for consumer behavior but is limited to structured temporal data, making it less suitable for irregular patterns in diesel engine prognostics. The CNN-BiGRU model (Xuan et al., 2021) enhances short-term load forecasting by capturing non-linear and temporal patterns, aided by feature selection and model fusion, but still assumes continuous time-series data. (Bieber, Verhagen, & Santos, 2021) were among the first to integrate environmental data for aircraft engine prognostics using raw sensor inputs and Random Forest for RUL prediction. For similar engines (Cheng, Zeng, Wang, & Song, 2023) used an ensemble deep learning approach with models like autoencoders, LSTM, and CNNs, weighted by Mahalanobis distance between health states, to estimate RUL for predictive maintenance.

In (Cheng, Wang, Wu, Zhu, & Lee, 2022), DL-based prognostics models are critiqued for lacking diverse feature extraction and ignoring operating conditions. To address this, they propose MDRNN combined with bi-LSTM to capture temporal dependencies and degradation features, achieving low RMSE on benchmark data. For multivariate time-series, (Yang et al., 2023) introduce attentional Gated Res2Net to detect multi-granular patterns and variable relationships. Though transformers are common in NLP and vision, (Jiang, Liu, & Lian, 2022) apply multi-modal fusion transformers for time-series classification using Gramian Angular Fields and

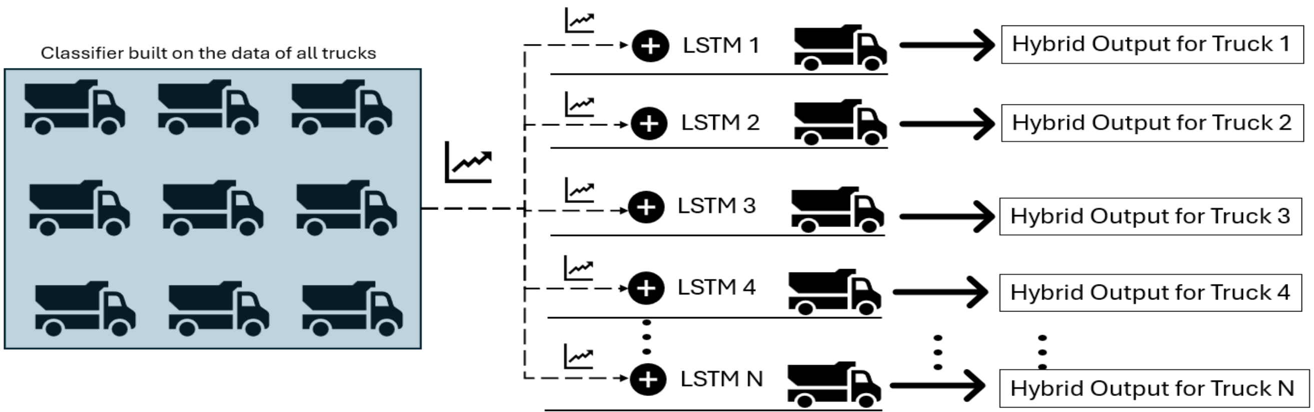

Figure 1. Block diagram of the proposed hybrid model where the output of the globally trained classifier is fed as one of the inputs to the individual LSTM models, built for individual trucks.

CNNs, followed by ResNet. In cybersecurity, (Karat et al., 2024) use CNN-LSTM to tackle evolving threats, while (Roshanzadeh, Choi, Bidram, & Martínez-Ramon, 2024) apply a 1D convolutional autoencoder for cyberattack detection in AC microgrids, combining CNN for local features and LSTM for temporal modeling. (Gong, Li, & Wang, 2022) propose TCN-BiGRU with attention and VMD to fuse multi-sensor data, optimize IMFs, and refine hyperparameters via grid search and cross-validation, achieving strong RUL prediction on C-MAPSS and sealing ring datasets.

S. Hamida and et. al. build a hybrid system (Hamida et al., 2023) developed a hybrid system for skin disease classification using Random Forest for feature selection and CNN for classification. To handle sequential and non-sequential data, (Haibin & Yongliang, 2023) fed LSTM outputs into decision trees for time-dependent predictions. The LSTM-RF-PSO model (Chen, Chen, & Gu, 2023) integrates LSTM for autoregression, RF for environmental analysis, and PSO for optimization, outperforming standalone models in tourist flow forecasting. (Almaghrabi, Rana, Hamilton, & Rahaman, 2024) proposed a multilevel fusion framework with spatial encoding via 3D autoencoder, temporal encoding with residual connections, and prediction processing using sequence analyzers. This is termed MF-NBEA (Khan & Kumar, 2024). For ECG data, a GRU-CNN hybrid model with reductive bias (RB-GRU-CNN) fuses past residuals with current predictions to reduce error.

To predict wind power, (Pu et al., 2024) use VMD to reduce dimensionality, followed by BiLSTM and Random Forest, whose weighted outputs form the final prediction. For power load forecasting, (Jin, Weiqing, Bingcun, & Xiaobo, 2024) apply a hybrid of LSTM and genetic algorithm. In meteorology, (Vallejo & Manzione, 2024) use neural network auto-regression for multidimensional precipitation forecasting. Overall, while hybrid models show promise, challenges like individualized learning and irregular time-series data—especially in diesel engine prognostics—remain. This study proposes a global classifier trained across trucks, fused with individualized LSTMs, to reduce unnecessary repairs and downtime for Cummins.

3. PROPOSED METHODOLOGY

We propose a generalized (which could be any classifier with best-suited hyperparameters and performance) classifier that predicts the degrading stage for prognostics. In our case, the XGBoost is the classifier and then the output is blended with LSTM for each truck, described below. The block diagram is depicted in Figure 1.

3.1. Data Description

We at Cummins, have built an embedded software feature that summarizes and stores information about its powertrain via 400+ parameters. This proprietary data can be viewed, analyzed and downloaded using tools. Field Performance Analytics (FPA) at Cummins involves analyzing real-world data to evaluate the performance and reliability of engines and components under actual operating conditions. This includes metrics like fuel efficiency, operational hours, failure rates, and maintenance trends, along with environmental and application-specific factors. The insights gained help Cummins enhance product design, optimize maintenance schedules, and improve customer satisfaction by ensuring engines perform reliably across diverse environments. FPA also supports predictive maintenance and compliance with global regulations, driving innovation and quality in Cummins' offerings.

The FPA data is collected via the Electronic Control Module (ECM), which records information from various sensors and actuators embedded within the engine. The ECM stores data at a fixed sampling frequency and aggregates it into fewer rows corresponding to each trip, defined by a key-on and key-off cycle. This aggregation is necessary due to the limited memory capacity of the ECM. Additionally, data extraction occurs at irregular intervals, typically when the engine is brought in for service and accessed via a datalogger device. However, the ECM retains only the most recent trip data up to its memory limit, resulting in the loss of older trip records if the engine has not been serviced for an extended period.

3.2. Data Preparation

This section outlines the data preparation steps for both the Global classifier and the LSTM model. Figure 2 illustrates the preprocessing pseudo-code for the overall data.

Step 1: Drop parameters with >40% missing values.

Step 2: Compute RUL for each row.

Step 3: Create Target: If RUL ≤ 10,000 → Target = 1 (Faulty) Else → Target = 0 (Healthy).

Step 4: Standardize the parameters.

Step 5: Remove highly correlated features.

Step 6: Handle missing values.

Figure 2. Data Preparation Method.

3.3. Global Classifier

Instead of predicting an exact time-to-failure, the classifier classifies the system's condition to assess how close it is to a failure event. As stated above, 400+ features together constitute the feature space for the global classifier – XGBoost.

In the proposed methodology, we employ XGBoost as the global classifier to predict the Remaining Useful Life (RUL) of diesel engine components, utilizing historical Field Performance Analytics (FPA) data. The FPA data, which includes both static engine parameters (e.g., load, fuel efficiency) and aggregated performance metrics from diverse operating conditions, serves as input to the global classifier. XGBoost, known for its ability to handle large datasets and complex feature interactions, is trained on these features to identify degradation patterns and predict failure stages across multiple engines. By leveraging this approach, the global classifier captures generalized trends in engine behavior, forming a foundation for identifying potential failures. However, to improve generalization and account for temporal dependencies in the degradation process, this global classification output is further enhanced by time-based memory units in our hybrid model. The Global classifier was optimized through hyperparameter tuning using the method from the scikit-learn library.

3.4. LSTMs

Long Short-Term Memory (LSTM) networks are a type of recurrent neural network designed to capture long-term dependencies in sequential data. They overcome the vanishing gradient problem, enabling them to retain important information over long sequences. LSTMs are ideal for tasks involving temporal dynamics, such as time-series prediction and pattern recognition.

3.5. Hybrid Model

The engines used to train the global classifier are kept separate from the unseen engines. For each unseen engine, a global classifier is applied to the data of each trip, and its predictions are added to the feature space formed by the Field Performance Analytics (FPA) data. These global classifier predictions (0 = healthy, 1 = faulty) are then incorporated into the feature space on which an individual LSTM model is trained for each engine. By leveraging both the temporal dependencies specific to each engine's operational history and the global degradation patterns captured by the classifier, this hybrid model effectively integrates global and temporal information. The performance of this approach is evaluated in the subsequent section.

4. EXPERIMENTS, RESULTS, AND DISCUSSIONS

For experimentation, we have access to the FPA (Field Performance Analytics) dataset, which consists of approximately 150,000 engines, with 3,900 having previously failed. To build our model, we applied a bootstrapping technique to downsize the dataset, ensuring that both failed and non-failed engines were well-represented. The ratio of faulty to healthy engines was set at 1:1, with healthy engines randomly selected from a pool of approximately 140,000 non-failed engines and faulty engines from 3900 failed engines. The bootstrapping approach ensured random selection and representative sampling of both healthy and failed engine populations. Statistical validation was performed using a z-test across all parameters, confirming similarity between the selected subset and the overall population with a p-value ≥ 0.90.

The global classifier was trained on a subset of 882 engines, which were kept separate from the LSTM models. The LSTM was developed using data from 301 engines, where predictions from the global classifier were included as a critical input feature. In the classification task, failure was annotated as 1 and non-failure as 0. On average, the dataset exhibits a substantial class imbalance with a ratio of 99.3:0.7 between non-failure and failure instances (Healthy v/s faulty) across individual engines. We assessed the performance of our hybrid model by comparing it to a baseline model that excludes the LSTM component. The results reveal significant improvements: precision doubled, recall increased by 10%, and overall accuracy improved by 20%. To validate the robustness of our methodology, we also applied it to the C-MAPPS dataset by NASA [24]. The dataset consists of multiple multivariate time series, each representing data from a different jet engine in a fleet of the same type, with varying degrees of initial wear and manufacturing variation. The engines operate normally at the start of each time series, developing a fault at some point, with the training set capturing fault growth until failure, and the test set ending before failure. The goal is to predict the RUL in the test set, i.e., the number of cycles an engine will continue to operate before failure, with true RUL values also provided for evaluation.

We obtain similar performance improvements. These findings highlight the effectiveness of combining temporal dependencies with global classification insights for more accurate failure predictions in engine prognostics. The comparison of precision and recall for each of the engines in testing is depicted in Figure 3 and Figure 4 for FPA using the **sm-pre-lstm**(Smoothened Precision of LSTM), **sm-pre-cla**(Smoothened Precision of Classification), **sm-recall-cla**(Smoothened Recall of Classification), **sm-recall-lstm**(Smoothened Recall of LSTM). The average improvement in precision and recall is mentioned in Table 1.

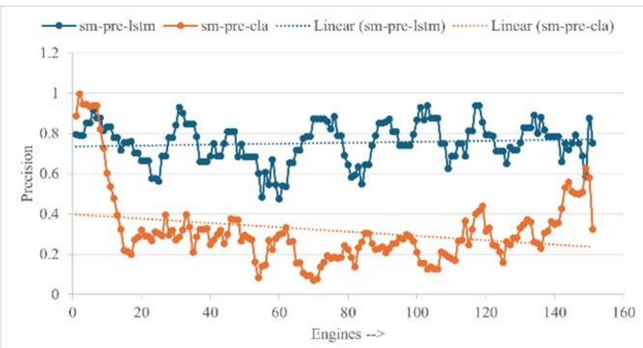

Figure 3. Engine-wise comparison of precision on FPA data.

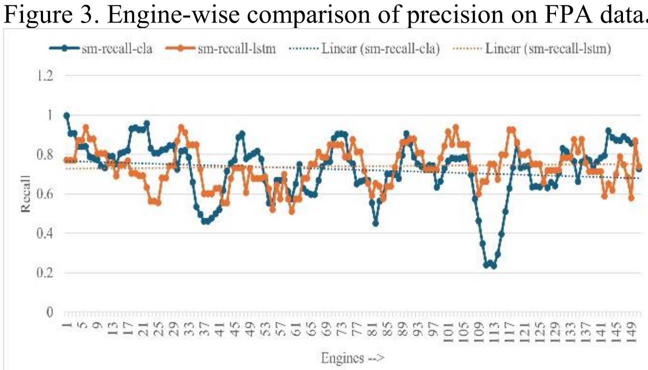

Figure 4. Engine-wise comparison of recall on FPA data.

Table 1 with Figure 5 and Figure 6 using **sm-cl-pre**(Smoothened Classification Precision), **sm-lstm-pre**(Smoothened LSTM Precision), **sm-cl-recall**(Smoothened Classification Recall), **sm-lstm-recall**(Smoothened LSTM Recall) also mentions how our methodology works on the C-MAPPS dataset. The graphs and tables clearly indicate the improvement in the performance after using hybrid model we proposed.

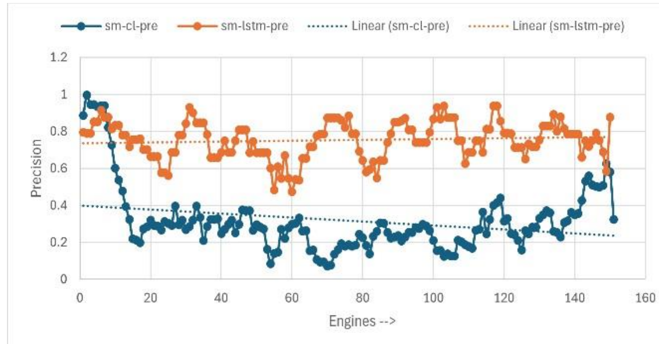

Figure 5. Engine-wise comparison of precision on C-MAPPS data.

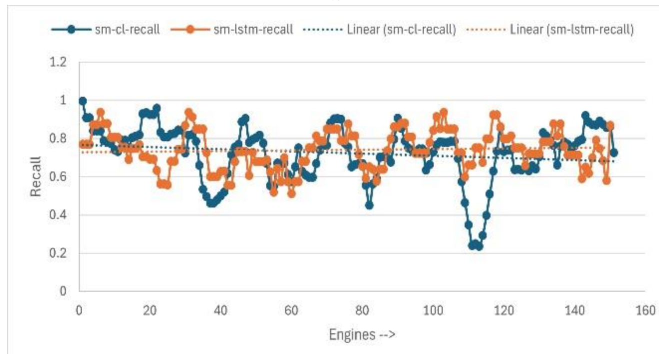

Figure 6. Engine-wise comparison of recall on C-MAPPS data.

5. CONCLUSION

Temporal dependencies are crucial for accurate prognostics of components in diesel engines. Ignoring these dependencies can lead to unnecessary repairs and increased unplanned downtime costs. The hybrid model we proposed enhances prognostic accuracy by 30% and doubles the precision, significantly reducing the risk of false positives that would have occurred without it. While the recall has shown improvement, it can be further optimized with additional input from domain experts.

Table 1. Comparison between FPA and C-MAPPS with and without using our hybrid model.

Parameters/Dataset	FPA data – without the proposed methodology	FPA data – with the proposed methodology	C-MAPPS data – without the proposed methodology	C-MAPPS data – with the proposed methodology
Average Precision	0.32	0.75	0.32	0.68
Average Recall	0.72	0.75	0.78	0.79
Average F1-score	0.44	0.74	0.46	0.62
Average Accuracy	0.66	0.81	0.76	0.77

REFERENCES

Alhnaity, B., & Abbod, M. (2020). A new hybrid financial time series prediction model. Engineering Applications of Artificial Intelligence, 95, 103873.

Almaghrabi, S., Rana, M., Hamilton, M., & Rahaman, M. S. (2024). Multivariate solar power time series forecasting using multilevel data fusion and deep neural networks. Information Fusion, 104, 102180.

Bieber, M., Verhagen, W. J., & Santos, B. F. (2021). Data-driven prognostics incorporating environmental factors for aircraft maintenance. In 2021 annual reliability and maintainability symposium (rams) (pp. 1–6).

Chen, J., Chen, P., & Gu, H. (2023). Lstm-rf-pso combined algorithm for short-term passenger flow forecasting in scenic areas. In 2023 international conference on electronics and devices, computational science (icedcs) (pp. 544–548).

Cheng, Y., Wang, C., Wu, J., Zhu, H., & Lee, C. K. (2022). Multi-dimensional recurrent neural network for remaining useful life prediction under variable operating conditions and multiple fault modes. Applied Soft Computing, 118, 108507.

Cheng, Y., Zeng, J., Wang, Z., & Song, D. (2023). A health state-related ensemble deep learning method for aircraft engine remaining useful life prediction. Applied Soft Computing, 135, 110041.

Gong, R., Li, J., & Wang, C. (2022). Remaining useful life prediction based on multisensor fusion and attention tcn-bigru model. IEEE Sensors Journal, 22(21), 21101–21110.

Haibin, C., & Yongliang, H. (2023). A hybrid lstm and decision tree model: A novel machine learning architecture for complex data classification. In 2023 ieee international conference on sensors, electronics and computer engineering (icsece) (pp. 1441–1446).

Hamida, S., El Gannour, O., Lamalem, Y., Saleh, S., Lamrani, D., & Cherradi, B. (2023). Efficient medical diagnosis hybrid system based on rf-dnn mixed model for skin diseases classification. In 2023 3rd international conference on innovative research in applied science, engineering and technology (iraset) (pp. 01–08).

Hu, J., Shen, L., & Sun, G. (2018). Squeeze-and-excitation networks. In Proceedings of the ieee conference on computer vision and pattern recognition (pp. 7132– 7141).

Hu, W., & Shi, Y. (2020). Prediction of online consumers' buying behavior based on lstm-rf model. In 2020 5th international conference on communication, image and signal processing (ccisp) (pp. 224–228).

Jiang, H., Liu, L., & Lian, C. (2022). Multi-modal fusion transformer for multivariate time series classification. In 2022 14th international conference on advanced computational intelligence (icaci) (pp. 284–288).

Jin, S., Weiqing, W., Bingcun, S., & Xiaobo, X. (2024). Research on time series prediction of hybrid intelligent systems based on deep learning. Intelligent Systems with Applications, 23, 200419.

Karat, G., Kannimoola, J. M., Nair, N., Vazhayil, A., VG, S., & Poornachandran, P. (2024). Cnn-lstm hybrid model for enhanced malware analysis and detection. Procedia Computer Science, 233, 492–503.

Karim, F., Majumdar, S., Darabi, H., & Harford, S. (2019) Multivariate lstm-fcns for time series classification. Neural networks, 116, 237–245.

Khan, S., & Kumar, V. (2024). A novel hybrid gru-cnn and residual bias (rb) based rb-gru-cnn models for prediction of ptb diagnostic ecg time series data. Biomedical Signal Processing and Control, 94, 106262.

Liu, Z., Zhang, C., Dong, E., Wang, R., Li, S., & Han, Y. (2023). Research progress and development trend of prognostics and health management key technologies for equipment diesel engine. Processes, 11(7), 1972.

Pu, W., Liu, K., Yang, F., Lu, J., Tan, H., & Zhang, C. (2024). Wind power ultra-short-term power prediction based on vmd and weighted combination model of bilstm and rf. In 2024 ieee 2nd international conference on power science and technology (icpst) (pp. 1507–1512).

Roshanzadeh, B., Choi, J., Bidram, A., & Mart´ınez-Ramon, ´ M. (2024). Multivariate time-series cyberattack detection in the distributed secondary control of ac microgrids with convolutional neural network autoencoder ensemble. Sustainable Energy, Grids and Networks, 38, 101374.

Vallejo, C. A. M., & Manzione, R. L. (2024). Multidimensional forecasting of precipitation and potential evapotranspiration in the paranapanema river basin using neural network time series. Journal of South American Earth Sciences, 142, 104961.

Xuan, Y., Si, W., Zhu, J., Sun, Z., Zhao, J., Xu, M., & Xu, S. (2021). Multi-model fusion short-term load forecasting based on random forest feature selection and hybrid neural network. Ieee Access, 9, 69002–69009.

Yang, C., Wang, X., Yao, L., Long, G., Jiang, J., & Xu, G. (2023). Attentional gated res2net for multivariate time series classification. Neural Processing Letters, 55(2), 1371–1395.

Zhao, H., Li, X., Xu, J., Fu, X., & Chen, J. (2023). Financial time series data prediction by combination model adaboost-knn-lstm. In 2023 international joint conference on neural networks (ijcnn) (pp. 1–8).