Predictive Maintenance Using CNN LSTM With Attention

Aishwarya S*
Sri Eshwar College of Engineering
Anna University
Chennai, Tamil Nadu, India.
aishwarya.s2020cse@sece.ac.in

Abstract

Predictive maintenance in the aerospace industry, typically relying on manual inspections and basic statistical models, is crucial for reducing downtime, costs, and ensuring safety. This study introduces and compares two advanced deep learning models, a CNN-LSTM and an Attention-based CNN-LSTM, designed to improve predictive accuracy for Remaining Useful Life (RUL) estimation. Using NASA's C-MAPSS dataset, these models capture spatial-temporal dependencies and focus on critical features. The results show that attention mechanisms significantly enhance both interpretability and accuracy, making these models widely applicable for more reliable and actionable insights in predictive maintenance.

1 Introduction

In the aerospace industry, where high-performance machinery like turbofan engines is crucial, accurate Remaining Useful Life (RUL) prediction is essential for ensuring safety and cost-effectiveness (Zhao et al., 2017; Babu et al., 2016). Machine learning, particularly deep learning, has shown great potential in predictive maintenance by modeling complex, non-linear relationships in large datasets. CNNs are effective in extracting spatial features from time-series data, while LSTMs capture long-term dependencies, making them ideal for analyzing sensor data from aerospace systems (Zheng et al., 2017). This paper compares two deep learning architectures for RUL prediction: a CNN-LSTM model and a CNN-LSTM model with attention layers. Using NASA's C-MAPSS dataset, we evaluate how attention mechanisms affect predictive accuracy and model interpretability. The goal is to assess whether attention improves RUL predictions and provides better insights into critical operational factors (Saxena et al., 2008, Selvamurugan, A. et al.).

2 Proposed Methodology

The CNN-LSTM model leverages the strengths of both CNNs and LSTMs to capture spatial and temporal patterns from time-series data. To improve the interpretability and performance of the CNN-LSTM model, we propose an enhanced architecture that incorporates attention mechanisms. Attention layers are added to allow the model to dynamically focus on critical parts of the input sequence, which can improve the model's ability to predict RUL more accurately.

2.1 Dataset and Preprocessing

The NASA C-MAPSS dataset (Saxena et al., 2008) is a widely used benchmark for predictive maintenance, simulating turbofan engine sensor data under varying operational conditions and degradation modes. The Remaining Useful Life (RUL) is calculated by subtracting the current time step from the total number of cycles for each engine. After performing correlation analysis, highly correlated features like sm_9 were removed to reduce redundancy. Selected sensor features (e.g., sm_2, sm_3, sm_7) were normalized using a MinMaxScaler to ensure consistent scaling between 0 and 1, improving model convergence. The dataset was split by engine IDs, with 80% used for training and 20% reserved for testing, ensuring that each engine's data was fully contained within either set to avoid data leakage. For sequence generation, sensor readings were divided into overlapping sequences of 50-time steps, paired with corresponding RUL labels, creating the necessary 3D input for the CNN-LSTM model, with dimensions for sequences, time steps, and features.

3 Implementation

The proposed CNN-LSTM and CNN-LSTM with Attention models were implemented using Tensor- Flow and trained on the NASA C-MAPSS dataset to predict the Remaining Useful Life (RUL) of turbofan engines. The model architecture was compiled with the Adam optimizer, using a learning rate of 0.001 and a Mean Squared Error (MSE) loss function.

The input sequence length was set to 50-time steps, and each input sample contained 13 features representing the selected sensor measurements. The CNN layers extract spatial features from the sensor data, while the LSTM layers capture temporal dependencies. In the attention-based model, attention layers were added to allow the model to focus on important time steps, enhancing interpretability and potentially improving predictive performance.

CNN-LSTM Architecture	CNN-LSTM with Attention Architecture
Layer (Type)	**Layer (Type)**
Input Layer (50, 13)	Input Layer (50, 13)
Conv1D (32 filters, kernel size=3)	Conv1D (32 filters, kernel size=3)
Conv1D (32 filters, kernel size=3)	Conv1D (32 filters, kernel size=3)
Conv1D (64 filters, kernel size=3)	Conv1D (64 filters, kernel size=3)
Conv1D (64 filters, kernel size=3)	Conv1D (64 filters, kernel size=3)
Conv1D (128 filters, kernel size=3)	Conv1D (128 filters, kernel size=3)
Flatten	Flatten
Reshape (50, 128)	Reshape (50, -1)
LSTM (50 units)	LSTM (50 units)
LSTM (50 units)	LSTM (50 units)
LSTM (50 units)	Attention (LSTM output with itself)
Dense (50, 1)	Concatenate (LSTM output and Attention output)
	Dense (50 units, ReLU)
	Dropout (0.2)
	Attention (Dropout output with itself)
	Concatenate (Dropout and Attention output)
	Dense (1 unit)

Table 1: Comparison of CNN-LSTM and CNN-LSTM with Attention Architectures

For training, the dataset was split into training and validation sets, with 80% of the data used for training and 20% for validation. The model was trained for 250 epochs with a batch size of 512. Early stopping was used to prevent overfitting, monitoring the validation loss and stopping the training when there was no further improvement.

Predictive Maintenance NASA turbofan

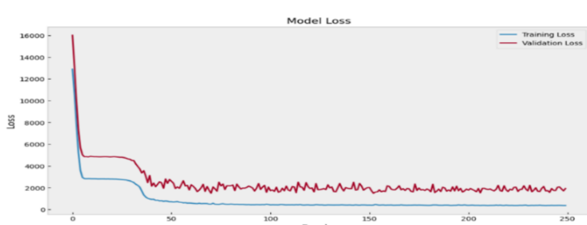

Figure 1: CNN LSTM with Attention model performance

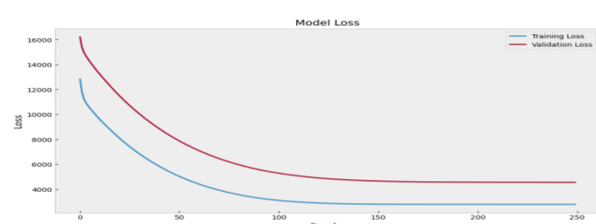

Figure 2: CNN LSTM model performance

Hyperparameter	CNN-LSTM + Attention Model
Learning Rate	0.001
Attention Mechanisms	2 attention layers
Dropout Rate	0.2
Optimizer	Adam
Loss Function	Mean Squared Error (MSE)

Table 2: Hyperparameters for CNN-LSTM + Attention Model

	RMSE	
Model	**Train**	**Test**
CNN LSTM	56.38	70.37
CNN LSTM with Attention	25.11	41.17

Table 3: Performance Metrics

Result

The CNN LSTM Attention model outperforms the CNN LSTM model, with a training loss of 25.11 and test loss of 41.17, compared to 56.38 and 70.37, respectively, for the CNN LSTM model. The attention mechanism improves performance by helping the model focus on the most relevant features in the time series data, leading to better prediction accuracy and lower loss. This selective focus enhances the model's ability to generalize, resulting in more effective predictions of the Remaining Useful Life (RUL) in predictive maintenance scenarios.

References

Babu, G. S., Zhao, P., & Li, X. L. Deep convolutional neural network-based regression approach for estimation of remaining useful life. *Proceedings of the International Conference on Database Systems for Advanced Applications*, 214-228, 2016.

Saxena, A., Goebel, K., Simon, D. and Eklund, N. Damage propagation modeling for aircraft engine run- to-failure simulation. *Proceedings of the 2008 International Conference on Prognostics and Health Management (PHM),* 2008.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł. and Polosukhin, I. Attention is all you need. *Advances in Neural Information Processing Systems*, 30, 5998-6008, 2017.

Wang, T., Yu, J. and Zhu, J. Remaining useful life prediction of machinery based on a convolutional neural network and long short-term memory. *Advances in Mechanical Engineering*, 10(12), 168781401881173, 2018.

Selvamurugan, A. et al. 2024. CNN-LSTM-Based Nonlinear Model Predictive controller for temperature trajectory tracking in a batch reactor. *ACS Omega*. (Nov. 2024). DOI:https://doi.org/10.1021/acsomega.4c07893.

Zheng, S., Ristovski, K., Farahat, A.and Gupta, C. Long short-term memory network for remaining useful life estimation. *Proceedings of the 2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*, pp. 88-95, 2017.

Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P. and Gao, R. X. Deep learning and its applications to machine health monitoring: A survey. *Mechanical Systems and Signal Processing*, 115, 213-237, 2017.