Turbofan engine health status prediction with neural network pattern recognition and automated feature engineering

Sławomir Szrama
Aviation Division, Poznan University of Technology, Poznan, Poland

Abstract

Purpose – This study aims to present the concept of aircraft turbofan engine health status prediction with artificial neural network (ANN) pattern recognition but augmented with automated features engineering (AFE).

Design/methodology/approach – The main concept of engine health status prediction was based on three case studies and a validation process. The first two were performed on the engine health status parameters, namely, performance margin and specific fuel consumption margin. The third one was generated and created for the engine performance and safety data, specifically created for the final test. The final validation of the neural network pattern recognition was the validation of the proposed neural network architecture in comparison to the machine learning classification algorithms. All studies were conducted for ANN, which was a two-layer feedforward network architecture with pattern recognition. All case studies and tests were performed for both simple pattern recognition network and network augmented with automated feature engineering (AFE).

Findings – The greatest achievement of this elaboration is the presentation of how on the basis of the real-life engine operational data, the entire process of engine status prediction might be conducted with the application of the neural network pattern recognition process augmented with AFE.

Practical implications – This research could be implemented into the engine maintenance strategy and planning. Engine health status prediction based on ANN augmented with AFE is an extremely strong tool in aircraft accident and incident prevention.

Originality/value – Although turbofan engine health status prediction with ANN is not a novel approach, what is absolutely worth emphasizing is the fact that contrary to other publications this research was based on genuine, real engine performance operational data as well as AFE methodology, which makes the entire research very reliable. This is also the reason the prediction results reflect the effect of the real engine wear and deterioration process.

Keywords Aircraft turbofan engine, Health status prediction, Neural network pattern recognition, Artificial neural network, Prognostic health monitoring, Turbine engine failure analysis

Paper type Research paper

1. Introduction

Engine health status determination is one of the crucial factors affecting aircraft flight safety. Aircraft flight operations are based nowadays on the maintenance preventive and prognostic strategy, which is a combination of preventive maintenance tasks and operations as well as engine monitoring and prediction system. It is quite common to adopt this maintenance strategy based on the engine performance and reliability learnt from the experience. It allows to save a lot of maintenance man-hours, aircraft downtime and in result a great deal of money for the aircraft owner and operator. To be able to enhance aircraft maintenance strategy it is mandatory to have the ability of engine health status prediction. Having such capabilities, allows to change, adopt and schedule engine maintenance and what is even more important is the fact that this allows to predict the moment of engine failure and prevent this scheduling maintenance task to prevent such failure, or remove engine from service if such failure might jeopardize aircraft flight safety. Turbofan engine trending and diagnostics were discussed among others by Szrama (2019), Matuszczak *et al.* (2021). One of the current main lines of research is about artificial intelligence methods applications, especially artificial neural networks (ANNs). Engine predictive strategy supported by ANN was presented for instance by Barad *et al.* (2012), Andrianantara *et al.* (2021), Da Costa *et al.* (2019), Qiao Zijian and Ji-Yu (2022), Wang *et al.* (2024), MA *et al.* (2022), Wang *et al.* (2024), Li *et al.* (2021) or Szrama and Lodygowski (2023). Engine fault detection and isolation with neural network usage were presented by Sadough Vanini *et al.* (2014) or Sina Tayarani-Bathaie *et al.* (2014). Transfer learning concept was implemented by: Tang *et al.* (2019) or Zhong *et al.* (2019).

The current issue and full text archive of this journal is available on Emerald Insight at: **https://www.emerald.com/insight/1748-8842.htm**

Aircraft Engineering and Aerospace Technology
96/11 (2024) 19–26
Emerald Publishing Limited [ISSN 1748-8842]
[DOI 10.1108/AEAT-04-2024-0111]

Received 26 April 2024
Revised 6 June 2024
8 July 2024
Accepted 19 July 2024

There are some more advanced types of neural networks already invented. The deep learning neural networks are often used for the engine remaining useful life estimation. Szrama and Lodygowski (2024) proposed using Convolutional neural network and Long short-term memory-type deep learning network to predict the remaining life of the engine.

Engine health status prediction is a key step in assessment and prediction of the remaining useful life which was discussed by Zhao *et al.* (2022), Wu *et al.* (2019), Sateesh Babu *et al.* (2016), Li *et al.* (2022), Lee and Mitici (2023) or Wang *et al.* (2024). Comparison and the review of the pattern recognition and deep learning technologies and their role in engineering research was elaborated by Serey *et al.* (2023).

The deep learning networks are usually applied for the long-term dependencies in signals. Not always has to be very complex and deep neural network created. In the article, it was decided to create shallow neural network architecture for the pattern recognition in the engine performance data.

Very few studies were performed with the neural network pattern recognition system, and this was the reason it was decided to conduct the research presented in the article. In addition to this, none of the publications were based on the real-life engine operational data. All the publications including the proposed ones were used with the engine simulated data provided by National Aeronautics and Space Administration and called C-MAPSS Aircraft Engine Simulator Data. As it was proved in the authors' article mentioned above, there is a huge difference in neural network performance while training and testing on simulated data and real-life data.

In this article, authors proposed the novel concept of neural network pattern recognition application augmented with automated feature engineering (AFE) for the turbofan engine health status prediction.

2. Neural network pattern recognition architecture and methodology

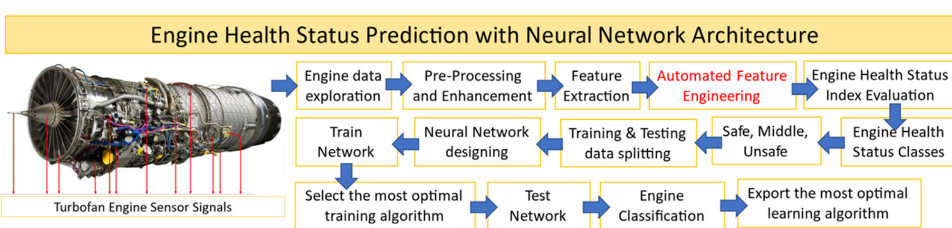
Pattern recognition can be described as a process of finding regularities and similarities in data using machine learning algorithms and architectures. Novel pattern recognition system architecture and methodology are presented in Figure 1.

Data exploration allows to analyse engine data, engine parameters it comprises and to determine which data is absolutely necessary. The accuracy of recognition highly depends on the quality of the data set. Pre-processing is coupled with enhancement. It involves smoothing and normalization process of the data. At the next stage, the input data is transformed into a feature vector, a reduced representation of a set of features. In the data extraction features step, it is being decided, which engine-sensed parameters are crucial for engine degradation assessment. Then extracted features are compared to a similar pattern stored in the database. When patterns are eventually matched to the stored data, the classification of input data is performed. As a part of the novel approach to the neural network performance enhancement, it was decided to implement new step called automated features engineering (AFE). During this step, the designed function automatically generates 13 new (additional) features from the original training data, and then applies the same transformations to the test data. The following step is the Engine Health Status Index Evaluation. In our case, Engine Health Index results from performance margin (PMAR) and specific fuel consumption margin (SMAR) engine parameters (thoroughly explained by Szrama and Lodygowski, 2023). PMAR is used to define and determine engine degradation level and is based on the engine exhaust gas temperature. SMAR expresses the fuel consumption rate in relation to the engine thrust (power). Based on the analysis performed in the mentioned article, engine health status classes were set based on the PMAR and SMAR being: safe, middle and unsafe. Then the complete engine data was split into training set and testing set (usually we take two/thirds of the complete set as training data and one/third of the testing data). The next step was to design neural network architecture which in this case was the shallow feedforward network with sigmoid hidden neurons and three output neurons. The next step was to train engine training data set using various network architectures. Network training results were compared and checked. Not only was the learning accuracy measured but also what is the cost of the misprediction. The crucial misprediction is when for the engine which has reached unsafe engine health condition, the prediction value is safe. Another crucial step was to evaluate the designed neural network architecture on the separate engine testing data set. As a result, it was possible to confirm neural network performance and accuracy.

The study was conducted for the two-layer feed-forward network, with 10, 20 or 50 sigmoid hidden layers and three output neurons which were trained to classify input vectors into three different classes, being: Safe, Middle and Unsafe. They were defined as the following:

- Safe– [1 0 0];
- Middle– [0 1 0]; and

Figure 1 Neural network pattern recognition system architecture and methodology



Source: Authors' own creation

- Unsafe– [0 0 1] vectors.

To understand how the novel methodology step affected the neural network performance, it was decided to measure network performance without AFE function and then compare the results to the results achieved with the newly generated features.

2.1 Automated features engine data

To improve the neural network performance, it was decided to design new neural network pattern recognition network with AFE algorithm. Artificial network pattern recognition architecture with AFE for the turbofan engine data set was presented in Figure 2.

Based on the newly designed network architectures, new features were generated to check the network performance and accuracy. Before passing the original engine training data to a classifier, new additional features were generated from the predictors in the engine data set. The returned data was used to train the classifier. As a part of the predictors generation, the minimum redundancy maximum relevance (MRMR) features selection method was implemented. The MRMR algorithm finds an optimal set of features that is mutually and maximally dissimilar and can represent the response variable effectively. The algorithm minimizes the redundancy of a feature set and maximizes the relevance of a feature set to the response variable. The algorithm quantifies the redundancy and relevance using the mutual information of variables-pairwise mutual information of features and mutual information of a feature and the response. As a part of the MRMR feature selection, predictors were ranked and then included in the requested number of top-ranked features in newTrain engine data set. The most important predictors selected were: Tt2 (Total Temperature at Engine Inlet), FTIT (Fan Turbine Inlet Temperature) and Pt6 (Total Pressure at the Engine Exhaust). Tt2 is the parameter which is the key engine input and sensed parameter. Around this feature, all the engine operational maps are designed, and engine operation is controlled by a Full Authority Engine Controller based on the Tt2. FTIT is the crucial engine operational and safety parameter. This parameter defines the highest temperature at the inlet of the fan turbine, which is safe for engine construction. Pt6 is also the key parameter which is used by full authority digital engine control to control engine pressure ratio, which is the engine performance parameter.

As a result of AFE, it was decided to generate 13 new features as a combination of the existing ones. The automated features were added to the original data to enhance the neural network performance and engine health status simulation. Some generated features are a combination of multiple transformations. For example, five features were generated by converting the variable to the categorical variable with three categories and then calculating the frequency of the categories. As a result, the new neural network architecture was designed with 59 input neurons, and trained, validated and tested in accordance with scaled conjugate gradient algorithm.

3. Neural network performance metrics

One of the most important neural network performance metrics is the cross-entropy loss function (PRF). The function returns result that heavily penalizes outputs that are extremely inaccurate ($\tilde{y}_i \sim 1 - y_i$), with very little penalty for fairly correct classifications ($\tilde{y}_i \sim y_i$). Minimizing cross-entropy allows to converge the classification model.

Cross-entropy loss function (PRF) could be calculated in accordance with equation (1):

$$CrossEntropy = PRF(\tilde{y}, y) = -\sum_{i=1}^{N} y_i ln \tilde{y}_i \tag{1}$$

where:
y_i = the following target value; and
$\tilde{y}_i$ = the following predicted value.

In addition to the previously mentioned performance metric, additional metrics were proposed as a part of the comparison between the results achieved for the simulated and real-life engine data.

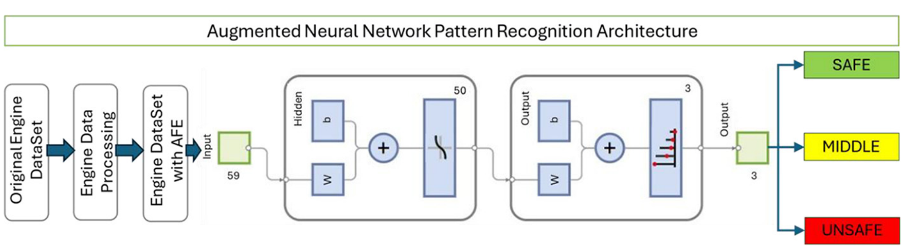
Accuracy performance metric. This performance metric is usually used in case of classification neural networks, where it is calculated as a ratio of the sum of True Positive and True Negative values divided by the sum of True Positives, True Negatives, False Positives and False Negatives. It could be calculated in accordance with equation (2):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

The next performance metric is the Mean absolute percentage error (MAPE) which is calculated between the predicted values and the actual values. It can be defined as:

$$MAPE = \frac{1}{N}\sum_{i=1}^{N} \frac{y_i - \tilde{y}_i}{y_i} \cdot 100\% \tag{3}$$

Figure 2 Artificial network pattern recognition architecture with automated features engineering for the turbofan engine data set with 50 sigmoid hidden layers and 3 output neurons trained to classify input vectors into three different classes: safe, middle and unsafe



Source: Authors' own creation

Another performance metric, which is used in Recurrent Neural Networks is the coefficient of determination R-squared. In the context of regression, it is a statistical measure of how well the regression line approximates the actual data. R-squared coefficient can be calculated in accordance with equation (4):

$$R^2 = 1 - \frac{SSR}{SST} = 1 - \frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \tag{4}$$

where:

N = number of observations;
SSR = sum squared regression is the sum of the residuals [1] squared;
SST = total sum of squares is the sum of the distance the data is away from the mean all squared;
y_i = the following target value;
$\tilde{y}_i$ = the following predicted value; and
$\bar{y}$ = the mean of the predicted value.

Percentage of errors %Error which the sum of the mispredictions divided by the number of elements in the data set:

$$\%Error = \sum_{i=1}^{N} \frac{(y_i \sim = \ \tilde{y}_i)}{N} \cdot 100\% \tag{5}$$

Another comparison of the results might be performed by comparing relative accuracy RA, which could be calculated as a ratio of predictions to the actual values:

$$RA = \frac{\tilde{y}_i}{y_i} \cdot 100\% \tag{6}$$

Precision is the metric which presents how accurate the positive predictions are, and could be calculated as follows:

$$Precision = \frac{TP}{TP + FP} \tag{7}$$

Recall is a metric that measures how often a machine learning model correctly identifies positive instances (true positives) from all the actual positive samples in the data set and is calculated as follows:

$$Recall = \frac{TP}{TP + FN} \tag{8}$$

The F1 score is calculated as the harmonic mean of the precision and recall scores, as shown below. It ranges from 0% to 100%, and the higher F1 score denotes the better-quality classifier:

$$F1Score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{9}$$

4. First case study on engine health status index

For the first case study, a new engine data set was created, which consisted of the engine performance data. Engine performance data was collected for the same type of aircraft turbofan engine, which was a low by-pass high-performance engine with mixer and afterburner. Engine data collection selected for this study consisted of real-life engine operational data. This amount of data has been collected for over twelve years. A total of 70% of the full set (20,999 observations) were randomly assigned as training data. A total of 15% of the complete data (4,500 observations) were selected for the validation process and the rest 15% (4,500 observations) were dedicated to test neural networks. Test results achieved for the 50-layer network are presented in Table 1.

In this case, the best-achieved cross-entropy was 0.0147 after 226 epochs. To analyse, what was the artificial network performance, another analysis was performed based on the confusion matrices. For all the processes: training, validation and tests, the overall performance ranged from 96.7% for Class 3 to 99.2% for Class 2. It is important to notice was the fact that only in very minor number of cases (about 3.3%) the prediction of the most serious class being the unsafe engine health condition was predicted as middle and none of the predictions for Class 3 were classified as Class 1 being safe. In all the cases, negative prediction rate was about 1.0%.

The final neural network performance analysis tool is the Error Histogram from which it was deduced that all the outputs (engine health status predictions) were assigned to two bins. The first one is around −0.04943 error value, whereas the other one is around 0.0496 value. It is worth noting that both bins are located close to the zero error line. It means that the network response is working properly with the Error range extending from −0.9407 to 0.9409.

The same engine data set was applied to the new neural network with pattern recognition and AFE. This resulted in a significant improvement of the neural network performance as well as other performance metrics. For all the generated numbers of the hidden layers, accuracy, precision, recall, F1Score and R^2 coefficient achieved the maximum values, while MAPE, mean square error (MSE), %Error and RMSE equalled zero value.

5. Second case study on Engine Health Status Index

In this case, Engine Health Status Index was calculated based on the SMAR, used to define engine health status and

Table 1 Original engine dataset neural network pattern recognition results for PMAR, SMAR and engine failure data for 50-layer neural network architecture

EH data	PRF	% error	Accuracy	Precision	Recall	F1Score	RA	MAPE	MSE	RMSE	R^2
PMAR	0.0147	1.0400	0.9930	0.9888	0.9912	0.9900	99.6177	0.6561	0.0104	0.1019	0.9722
SMAR	0.1896	0.2833	0.8111	0.7017	0.8321	0.7614	61.0637	26.2103	0.2938	0.5420	−0.6301
EF data	0.0005	0.0000	1.0000	1.0000	1.0000	1.0000	80.0000	0.0000	0.0000	0.0000	1.0000

Source: Author's own creation

condition. It is evident that specific fuel consumption rate increases when the engine compressor compression rate decreases. If the engine compressor compression rate decreases in time for the same power level requirement, it means that engine compressor efficiency has degraded. Engine data set was prepared in the equivalent way, as it was generated for the PMAR parameter. The results achieved by the neural network pattern recognition process are presented in Table 1. The achieved results for this case study are worse with at least the order of magnitude. What might be the reason for this? The most probable explanation is the fact that the correlation between engine performance data pattern and the SMAR response was not as simple as it could be presumed. Generated by the neural network pattern did not reflect the response in the way as it was achieved in the real life. What was the best performance validation check achieved? The best-achieved accuracy was not as high as it was for the PMAR parameter, achieving 0.1896 at epoch 185.

What was the artificial network performance based on the confusion matrices? For all the processes: training, validation and test, the correct prediction was slightly above 70%. What is worse there were 214 mispredictions between critical engine health status (unsafe) and middle or safe class.

To analyse the neural network results for SMAR engine data set, Error Histogram with 20 Bins plot for the 50-layer size network and SMAR parameter was created. In this case scenario, even though the positive prediction rate was not as high as it was for the PMAR parameter, still the differences between the target values and the outputs for all 20 bins ranged from −0.8558 to 0.9449. What is worth mentioning is the error distribution. There are outputs which were assigned to the bins not only close to zero error line but were distributed among other bins.

6. Third case study on engine failure data

To confirm the results achieved in the two previous case studies, it was decided to create another engine test data set. This data set was created based on engine failure data collected for 4 years between 2019 and 2022. This failure data consisted of 5,280 records (observations). What were the engine failures selected from the engine failure data? They were the most common engine failures concerning turbofan engine data. Among them it is worth mentioning: Augmentor/Afterburner Blowout, Engine Stall, Engine Stagnation, Engine Surge, Engine Hot Start, No Air Start, No Ground Start, Rear Compressor Variable Vanes. As it was previously mentioned, engine failure data records were mixed with the sample engine performance data and the final engine test data set counted 17,452 records. Results of the neural network pattern recognition for the engine failure data were presented in Table 1.

Network performance defined by cross-entropy for this case study was incredibly low for all the network sizes, reaching 5E-04 value for 50-layer size. Prediction error was also equal or remarkably close to zero value. All the calculations confirmed the fact, that defined network is recognizing engine status pattern with absolutely great accuracy. For this case study, the best validation performance, being the lowest cross-entropy was achieved after 121 epochs being 4.6491E-08.

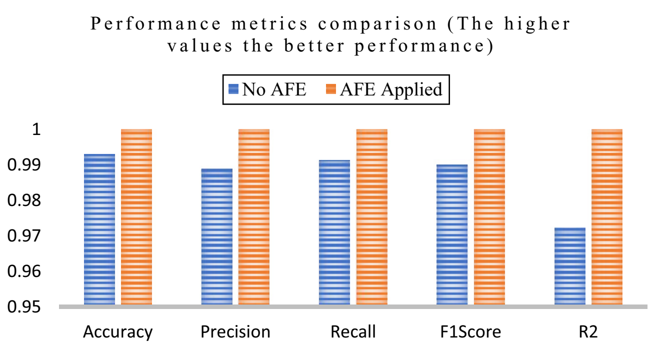
Having analysed the best validation performance plot for this case study, it was noticed that for the 50-layer size, neural network becomes overfitted, test results with the increment of the epoch are getting worse and the test curve is moving away from the training and validation.

From Error Histogram with 20 bins for the additional engine failure data, it was deduced that all the predictions were fitted into two bins with very minor errors ranging between 2.886E-02 and 2.886E-02 and remarkably close to the zero error line.

Comparison of the neural network pattern recognition performance metrics with and without AFE were presented in Figure 3.

Having analysed the neural network performance with and without AFE, it was concluded that for some performance metrics such as accuracy, precision, recall, F1Score, RA or $R2$ the improvement was about 1%, whereas for the others such as

Figure 3 ANN pattern recognition performance metrics comparison with and without the AFE



Source: Authors' own creation

Table 2 Results of the classification learners without AFE

Model type	PRF	% error	Accuracy	Precision	Recall	F1Score	RA	MAPE	MSE	RMSE	R^2
Tree	1.0000	0.0000	1.0000	1.0000	1.0000	1.0000	50.0000	0.0000	0.0000	0.0000	1.0000
Discriminant	0.9343	0.0412	0.9725	0.9516	0.9734	0.9624	49.4399	2.3813	0.0412	0.2030	0.8848
Naive bayes	0.7805	0.2155	0.8563	0.8171	0.7492	0.7817	91.6262	11.0190	0.2160	0.4647	0.6504
SVM	0.2485	0.9050	0.3966	0.0903	0.0629	0.0741	66.6667	47.1681	1.0842	1.0412	−0.6197
KNN	0.6883	0.1810	0.8792	0.8619	0.8342	0.8479	49.4399	11.3233	0.1857	0.4310	0.5912
Ensemble	1.0000	0.0001	0.9998	0.9997	0.9998	0.9998	50.0000	0.0054	0.0001	0.0128	0.9995
Kernel	0.8991	0.0695	0.9536	0.9360	0.9444	0.9402	49.4399	4.5709	0.0756	0.2750	0.8005

Source: Author's own creation

%Error, MAPE, MSE and RMSE, the improvement was reaching 100%.

7. Neural network and classification learner comparison

As a final validation of the neural network pattern recognition, it was decided to compare the results of the network to the machine learning classification algorithms being: decision tree for multiclass classification, discriminant analysis classifier, multiclass Naive Bayes model, multiclass models for support vector machines or other classifiers, k-nearest neighbour classifier, ensemble of learners for classification, binary Gaussian kernel classifier using random feature expansion.

These algorithms were used to train on the original features as well as on the new automated features and the results are presented in Table 2.

As a following step, the same machine learning algorithms with the same settings were applied to the engine data set enhanced with the generated new predictors. Having analysed the results, it might be deduced that in the case of the decision tree for multiclass classification, multiclass Naive Bayes model as well as ensemble of learners for classification, the results achieved for the engine data set with AFE were better. The greatest improvement achieved was for the Naïve Bayes algorithm, which ranged from 14.5% to 35%.

8. Summary and conclusion

The main goal of this article was to present the concept of aircraft turbofan engine health status prediction, taking advantage of real-life engine flight data. The novel concept of neural network pattern recognition application augmented with AFE for the turbofan engine health status prediction was presented. The additional goal was to check how such neural network works on real-life engine data and how the neural network recognizes patterns in real-life engine data.

The main achievement of this elaboration was the presentation of the complete methodology for the engine status prediction with neural network pattern recognition but augmented by AFE. Presented engine health status prediction was based on two ideas. The first one was based on the engine performance and safety parameters such as performance margin (PMAR) and specific fuel consumption margin (SMAR). This idea confirmed the thesis, that the pattern recognition network is working properly for both engine parameters. Especially satisfactory results were achieved for the PMAR engine health status parameter. For the SMAR parameter, even though prediction errors were not remarkably high, the error distribution was extremely wide, and there were 214 mispredictions. To improve the neural network performance, it would require SMAR engine health status index correction. It should reflect more real engine condition status. Unfortunately, the real engine status was incorrectly assigned to the wrong class, wrong pattern recognition and in results improper predictions. The second case study idea was based on the supervised engine failure data. All the failure data was confirmed as the actual engine problem. The results of the neural network pattern recognition were more than satisfying. For all the processes: training, validation and test predictions rate was 99.9%. It means that almost all the engine failures were properly predicted and every engine health status being: safe and unsafe (no middle class existed) was correctly classified. Even the predicted responses were close to zero error line with a very minor deviation reaching 0.02886. One of the important conclusions from the conducted study was the fact that it is not always a promising idea to increase the number of hidden layers (the size of the neural network), as it might result in the artificial network overfitting. Another important conclusion is the fact that neural network performance strongly depends on the input data and the size of the input data. If the input data is properly prepared and supervised, the results (engine health status predictions) were predicted correctly in almost 100% of cases.

Having analysed the results of the neural network pattern recognition augmented with AFE one important conclusion might be deduced. There is no doubt that AFE is a powerful tool to enhance either neural networks or machine learning algorithm augmented with AFE. For the proposed neural network pattern recognition architecture, the average improvement achieved for all the given performance metrics was about 46%.

The question might be raised, what the advantage of the proposed neural network pattern recognition with AFE is, in comparison to the machine learning algorithms, if the performance results in some cases are to some extent quite similar. The answer is that the greatest advantage of the proposed application is the fact that neural networks work excellently in every possible case while some of the machine learning algorithms might fail to converge. The reason for this is that neural networks work with diverse types of predictors. They work with both numerical data, vectors and categorical data, while ML algorithms (discriminant and KNN types) in some case might fail. The same problem might happen when the automated feature generation results in excessively big or infinite numerical data. Also, in this case scenario ML algorithms either fail to converge or result in very "weak" (with low performance) prediction models. The greatest issue in NN pattern recognition might be the newly generated engine data, which sometimes might reach infinite values. In this case scenario, even the NN will not be able to converge, and the results will not be correct. So why not to implement AFE for any type of neural network or machine learning algorithm. During the multiple and complex training, validations and tests it turned out that specific features were sometimes randomly generated. Five of the features generated from the mathematical conversions resulted in exceptionally large or even infinite variables. This, in turn, resulted in very weak classification models and neural network performance results.

Based on the research results it might be concluded that Neural Network Pattern Recognition is a strong tool which can help to solve engine health status detection and classification problems. One of the most valuable applications of pattern recognition is that such networks can analyse engine data observations and correlate multiple patterns across enormous amounts of data. Thanks to this, such AI models can make perfectly accurate engine health status predictions. The question might be raised if there are any disadvantages of neural network pattern recognition applications. One of the issues is the fact that such a model requires a great deal of engine data, which sometimes might not be available. In addition to this, heavy and based on the vast amount of data, training is required to train networks for pattern analysis. Another particularly critical issue is related to the data quality. Training data for machine learning algorithms should come from reliable sources. It should be free from bias and noise that hamper inherent pattern identification and decision-making capabilities of the neural network. That is the reason, AFE should be followed with the data analysis and noise removal. The following research works could be focused on the new neural network architectures with AFE and variable value sensitivity.

Presented methodology could be implemented into the aircraft (engine) maintenance management computer system, which could allow to automate engine health status analysis and improve engine maintenance management. Such methodology could help propulsion maintenance management in engine big data analysis and avoid any situations when the engine health status degraded below an acceptable level, especially for large engine fleets. Engine health status prediction based on the ANN augmented with AFE is an extremely strong tool in aircraft accident and incident prevention.

Note

1 Residual value = actual y value − predicted y value

References

Andrianantara, M., Ghazi, G. and Botez, R. (2021), "Aircraft engine performance model identification using artificial neural networks", *AIAA Propulsion and Energy*, doi: 10.2514/6.2021-3247.

Babu, G.S., Zhao, P. and Li, X.L. (2016),"Deep convolutional neural network based regression approach for estimation of remaining useful life", *Database Systems for Advanced Applications* DASFAA 2016. Lecture Notes in Computer Science 9642, doi: 10.1007/978-3-319-32025-0_14.

Barad, S., Ramaiah, P.V., Giridhar, R.K. and Krishnaiah, G. (2012), "Neural network approach for a combined performance and mechanical health monitoring of a gas turbine engine", *Mechanical Systems and Signal Processing*, Vol. 27, pp. 729-742, doi: 10.1016/j.ymssp.2011.09.011. ISSN 0888-3270

Costa, F., Domingues, P., Freire, R.Z., Coelho, L.S., Tavakolpour-Saleh, A.R. and Ayala, H.V. (2019), "Genetic algorithm for topology optimization of an artificial neural network applied to aircraft turbojet engine identification", 2019 IEEE Congress on Evolutionary Computation (CEC), *Wellington, New Zealand*, pp. 1095-1101, doi: 10.1109/CEC.2019.8790171

Lee, J. and Mitici, M. (2023), "Deep reinforcement learning for predictive aircraft maintenance using probabilistic remaining-useful-life prognostics", *Reliability Engineering & System Safety*, Vol. 230, p. 108908, doi: 10.1016/j.ress.2022.108908.

Li, B., Zhao, Y. and Chen, Y. (2021), "Unilateral alignment transfer neural network for fault diagnosis of aircraft engine", *Aerospace Science and Technology*, Vol. 118, pp. 1270-9638, doi: 10.1016/j.ast.2021.107031.

Li, H., Wang, Z. and Li, Z. (2022), "An enhanced CNN-LSTM remaining useful life prediction model for aircraft engine with attention mechanism", *PeerJ Computer Science*, Vol. 8, p. e1084, doi: 10.7717/peerj-cs.1084.

Matuszczak, M., Żbikowski, M. and Teodorczyk, A. (2021), "Predictive modelling of turbofan engine components condition using machine and deep learning methods", *Eksploatacja i Niezawodność – Maintenance and Reliability*, Vol. 23 No. 2, pp. 359-370, doi: 10.17531/ein.2021.2.16.

Serey, J., Alfaro, M., Fuertes, G., Vargas, M., Durán, C., Ternero, R., Rivera, R. and Sabattin, J. (2023), "Pattern recognition and deep learning technologies, enablers of industry 4.0, and their role in engineering research", *Symmetry*, Vol. 15 No. 2, p. 535, doi: 10.3390/sym15020535.

Sina Tayarani-Bathaie, S., Sadough Vanini, Z.N. and Khorasani, K. (2014), "Dynamic neural network-based fault diagnosis of gas turbine engines", *Neurocomputing*, Vol. 125, pp. 153-165, doi: 10.1016/j.neucom.2012.06.050.

Szrama, S. and Lodygowski, T. (2024), "Aircraft engine remaining useful life prediction using neural networks and real-life engine operational data", *Advances in Engineering Software*, Vol. 192, p. 103645, doi: 10.1016/j.advengsoft.2024.103645.,

Szrama, S. (2019), "F-16 turbofan engine monitoring system", *Combustion Engines*, Vol. 177 No. 2, pp. 23-35, doi: 10.19206/CE-2019-205.

Szrama, S. and Lodygowski, T. (2023), "Maintenance strategy supervised by machine learning on real engine flight data", *Journal of Aerospace Part-G, In-Publication*.

Tang, S., Tang, H. and Chen, M. (2019), "Transfer-learning based gas path analysis method for gas turbines", *Applied Thermal Engineering*, Vol. 155, pp. 1-13, doi: 10.1016/j.applthermaleng.2019.03.156. ISSN 1359-4311

Vanini, S., Khorasani, K. and Meskin, N. (2014), "Fault detection and isolation of a dual spool gas turbine engine using dynamic neural networks and multiple model approach", *Information Sciences*, Vol. 259, pp. 234-251, doi: 10.1016/j.ins.2013.05.032. ISSN 0020-0255.

Wang, Z., Zhu. and Zhao, X. (2024), "Dynamic predictive maintenance strategy for system remaining useful life prediction via deep learning ensemble method", *Reliability Engineering & System Safety*, Vol. 245, p. 110012, doi: 10.1016/j.ress.2024.110012.

Wu, J., Hu, K., Cheng, Y., Wang, J., Deng, C., et al (2019), "Ensemble recurrent neural network-based residual useful life prognostics of aircraft engines", *Structural Durability & Health Monitoring*, Vol. 13 No. 3, pp. 317-329, doi: 10.32604/sdhm.2019.05571.

Zhao, S., Pang, Y., Chen, J. and Liu, J. (2022), "Predication of remaining useful life of aircraft engines based on multi-head attention and LSTM", 2022 IEEE 6th Information Technology and Mechatronics Engineering Conference (ITOEC), *Chongqing, China*, pp. 1530-1534, doi: 10.1109/ITOEC53115.2022.9734660.

Zhong, S., Fu, S. and Lin, L. (2019), "A novel gas turbine fault diagnosis method based on transfer learning with CNN", *Measurement*, Vol. 137, pp. 435-453, doi: 10.1016/j.measurement.2019.01.022. ISSN 0263-2241.

Further reading

Tayarani-Bathaie, S. and Khorasani, K. (2015), "Fault detection and isolation of gas turbine engines using a bank of neural networks", *Journal of Process Control*, Vol. 36, pp. 22-41, doi: 10.1016/j.jprocont.2015.08.007. ISSN 0959-1524.

Fentaye, A.D., Gilani, S.I., Baheta, A.T. and Li, Y.-G. (2018), "Performance-based fault diagnosis of a gas turbine engine using an integrated support vector machine and artificial neural network method, proc", *Inst. Mech. Eng. A, J. Power Energy*, doi: 10.1177/0957650918812510.

Gurrola, M. and Botez, R.M. (2022), "Improved local scale generic cycle model for aerothermodynamic simulations of gas turbine engines for propulsion", *Designs*, Vol. 6 No. 5, p. 91, doi: 10.3390/designs6050091.

Khan, S. and Yairi, T. (2018), "A review on the application of deep learning in system health management", *Mechanical Systems and Signal Processing*, Vol. 107, pp. 241-265, doi: 10.1016/j.ymssp.2017.11.024.

Kothari, S.C. and Oh, H. (1993),*Neural Networks for Pattern Recognition*, in Marshall, C. (Ed.), Yovits, Advances in Computers, Elsevier, Vol. 37, pp. 119-166, ISSN 0065-2458, ISBN 9780120121373, doi: 10.1016/S0065-2458(08)60404-0.

Metlek, S. (2023), "A new proposal for the prediction of an aircraft engine fuel consumption: a novel CNN-BiLSTM deep neural network model", *Aircraft Engineering and Aerospace Technology*, Vol. 95 No. 5, pp. 838-848, doi: 10.1108/AEAT-05-2022-0132.

Silva, F., Grinet, M. and Silva, A. (2022), "A machine learning approach to forecasting turbofan engine health using real flight data", *AIAA 2022-0491. AIAA SCITECH 2022 Forum*, doi: 10.2514/6.2022-0491.

Yu, B., Shu, W. and Cao, C. (2018), "A novel modeling method for aircraft engine using nonlinear autoregressive exogenous (NARX) models based on wavelet neural networks", *International Journal of Turbo & Jet-Engines*, Vol. 35 No. 2, pp. 161-169, doi: 10.1515/tjj-2017-0005.

Zhao, Y., Huang, G., Hu, G., Tan, J., Wang, J. and Yang, Z. (2019), "Soft extreme learning machine for fault detection of aircraft engine", *Aerospace Science and Technology*, Vol. 91, pp. 70-81, doi: 10.1016/j.ast.2019.05.021. ISSN 1270-9638.

Zhao, L., Mo, C., Sun, T. and Wei, T. (2020), "Aero engine Gas-Path fault diagnose based on multimodal deep neural networks", *Wireless Communications and Mobile Computing*, Vol. 2020, pp. 1-10, doi: 10.1155/2020/8891595.

Corresponding author

Sławomir Szrama can be contacted at: slawomir.szrama@put.poznan.pl