Machine Learning Based Surrogate Model for Fatigue Life Prediction of Aircraft Structures

Felipe Gomes[1], **Lennon Falcão**[1], **Victor Olegário**[1]

[1]Aeronautics Institute of Technology (ITA) – São José dos Campos, SP – Brazil

Abstract. *This paper describes the development of a surrogate model for a complex system used in fatigue life prediction of aircraft structures. Various machine learning techniques and hyperparameters were tested, including K-Nearest Neighbors, Support Vector Machines, Regression Trees, Random Forests, and Artificial Neural Networks. The resulting model achieved a lower computational cost while providing accurate fatigue life predictions, contributing to time savings during the design loop phase in the aeronautical industry. The model's performance was also evaluated using a new dataset that simulates in-production conditions, which helped identify certain limitations and potential areas for improvement.*

Resumo. *Este artigo descreve o desenvolvimento de um modelo substituto para um modelo complexo utilizado na previsão da vida em fadiga em estruturas de aeronaves. Foram testadas diferentes técnicas de aprendizado de máquina e hiperparâmetros: K-vizinhos mais próximos, Máquina de Vetores de Suporte, Árvores de Regressão, Floresta Aleatória e Redes Neurais Artificiais. O modelo obtido apresentou menor custo computacional e boa capacidade de previsão da vida em fadiga, economizando tempo nas decisões de projeto na indústria aeronáutica. O desempenho do modelo também é verificado usando um novo conjunto de dados, simulando a condição de produção, o que permitiu identificar limitações e possíveis melhorias.*

1. Introduction

1.1. Fatigue in Aircraft Structures

The fatigue phenomenon is defined as the process of progressive and localized structural damage that occurs when a material is subjected to cyclic loading, generating crack nucleation and propagation [Suresh 1991]. In general, cracks occur in regions where there are stress concentrations, where stress can be defined as the internal effort of the material. Accidents like the one with the de Havilland Comet model, described in [Cohen and Hall 1954], have proven that this phenomenon can be catastrophic if not properly considered in the aircraft design phase. Figure 1 shows a fatigue crack found after the Commet structural tests.

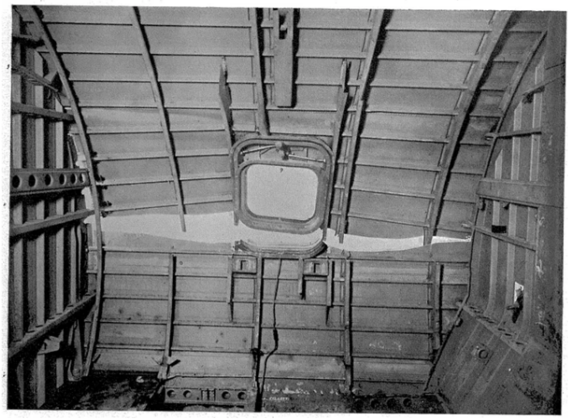

Figure 1. Crack propagation near a window in the Comet after a structural test.

The classical fatigue theory is mainly based on the so-called SN curves, as described in [Schijve 2009], [Budynas and Nisbett 2020] and [Dowling 2013]. This curve is obtained by submitting an standard test article to a cyclic loading for different stress levels, resulting in different life cycles until failure. For an aluminum alloy, an example of this curve can be seen in Figure 2 . In this case, cycles with zero mean stress were performed - a sinusoidal variation, for example). Each point represents an experimental result for a fixed amplitude [Dowling 2013].

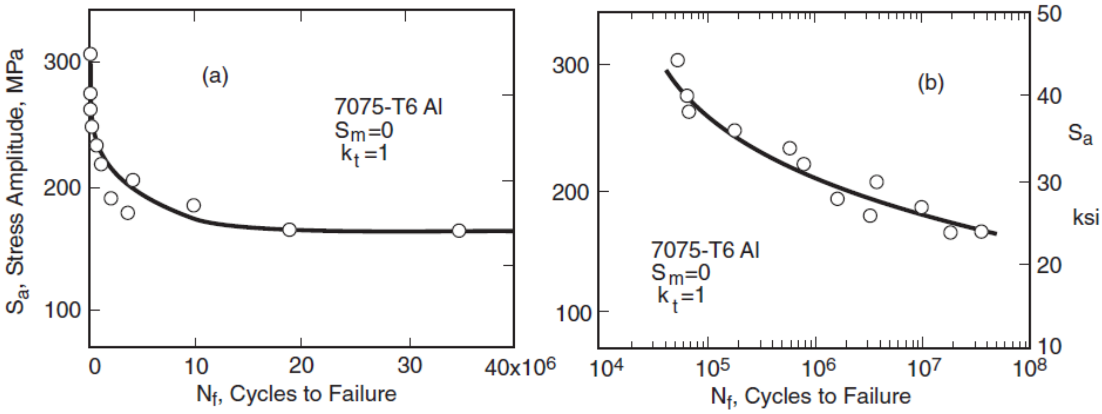

Figure 2. Stress vs life (S-N) curves on a linear scale (a) and on a logarithmic one (b).

However, in real applications, the fatigue cycles may not be that simple. Aircraft are machines that are submitted to several loads during their life cycle: vertical and lateral gusts, pressurization, ground impacts, among others. Essentially, there are three phases for a typical commercial operation: take-off run (including taxi), cruise flight (when the fuselage is full pressurized) and landing. In each phase, the aircraft passes through different conditions depending on altitude, temperature, velocity, etc. Therefore, for a single flight, there are thousands of different types of loading cases. The load variation generates different levels of stresses in the aircraft structure, as shown in Figure 3. As the aircraft accumulates flights during its life, the structure passes through a continuous wear - that is the fatigue damage.

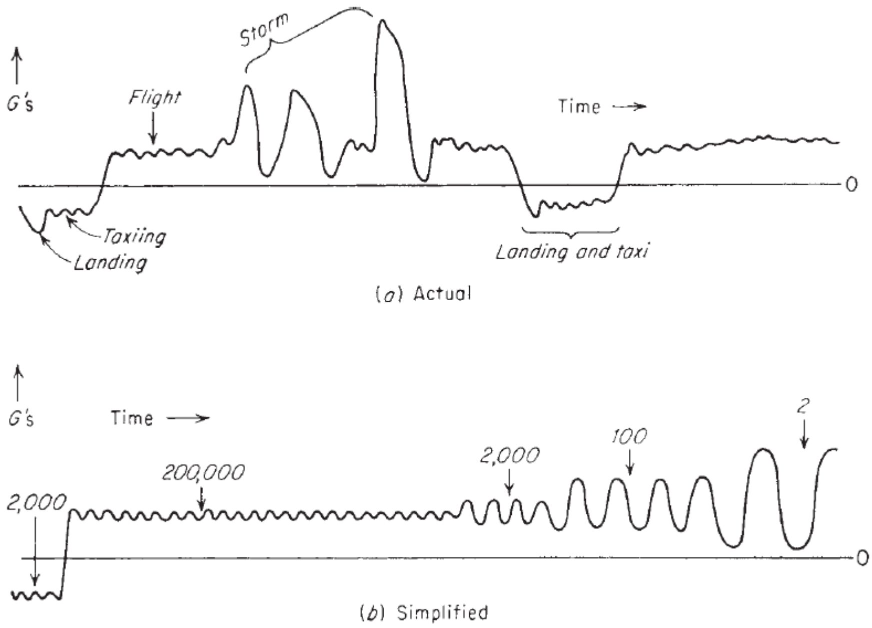

Figure 3. Example of flight load and stress history - the actual variation (a) and the simplied cycle (b) [Dowling 2013].

The process of obtaining stresses based in loads is the first challenge for a fatigue calculation. One common strategy is to simulate each condition using complex models, such as those based on the Finite Element Method, described in [Zienkiewicz and Taylor 2000]. In order to apply the classical theory, it is necessary to simplify the actual load history as shown in Figure 3b. This process of cyclics extraction for a typical flight can be complex since the load variation can not be modeled by a single mathematical periodic function. In order to solve this issue, some algorithms were developed. The most known is the rainflow algorithm [Matsuishi and Endo 1968], that enables to represent complex stress variations by the composition of some simple cycles. After obtaining the cycles, a methodology as described by [Palmgreen 1924] and [Miner 1945] can be applied in order to obtain the total damage.

One of the biggest challenges for the fatigue failure prevention is that every aircraft has thousands of structural details. Theoretically, every detail should be evaluated in terms of fatigue damage. In practice, the fatigue calculation can be very time and computational consuming since it demands the use of several calculation process in series. This issue forces engineers to apply criteria (like stress in a typical load case) for choosing the most critical regions that will be selected for a detailed fatigue verification. The problem is that the criteria may not be accurate enough, since a great amount of data must be evaluated.

Here is where Machine Learning can be a powerful tool for this type of problem. With results from structural elements that has already passed through a complete fatigue calculation process, a predictive supervised task can be applied for predicting damage for new structural details.

1.2. Surrogate Models

Surrogate models or meta-models are simplified approximations of more complexes models. Several techniques have been developed for surrogate modeling, requiring a systematic approach for selecting which technique may be appropriate for an application [Williams and Cremaschi 2021].

Example of surrogate models are described by [Forrester et al. 2008], with focus in engineer and design, covering methods like kriging, artificial neural networks (ANN) and support vector regressors. [Queipo et al. 2005] discuss several surrogate modeling techniques, mainly in aerospace applications. This article also highlights the use of such modeling strategy for making sensitivity and optimization studies. [Liu et al. 2020] evaluates neural network models, support vector regression, Gaussian processes and hybrid methods. Support vector regression is also used by [Shi et al. 2020]. In this context, it is shown an application of kernel functions for developing a linear input-output correlation. In this case, the root means square error (RMSE) is used for model evaluation. [Sun et al. 2019] focus on the integration of ML and surrogate models. In the structural context, [Khandelwal et al. 2021] develops a ML-based model to predict deformation of microstructures. In this study, the mean average error and the coefficient of determination (R^2) as in [Williams and Cremaschi 2021] are used to assess the model performance. [Williams and Cremaschi 2021] also presents how to use a Random Forest based tool to make predictions on the models' performance, which is very useful for a more systematic technique selection. For civil structures, [Eslamlou and Huang 2022] describes how to use ANN for developing surrogate models in health monitoring. In the aerodynamics field, [Sun and Wang 2019] reviews the use of ANN for bringing new perspectives to design optimization problems.

2. Materials and Methods

2.1. Dataset Description

The available dataset consists of results from elements from a fuselage structure. The results are for two component type (stringer and skin) and three different metallic alloys. Figure 4 shows an example of these structures in a typical fuselage. For each instance, the attributes are also stress results from six loading conditions: taxi, landing, vertical gust, lateral gust (right side and left side) and pressurization. The target attribute is a dimensionless value that represents the number of flights until failure in relation to the aircraft design service life (DSL). For example, a value of 0.5 means that the component fails in 20,000 flights, for a DSL of 40,000.

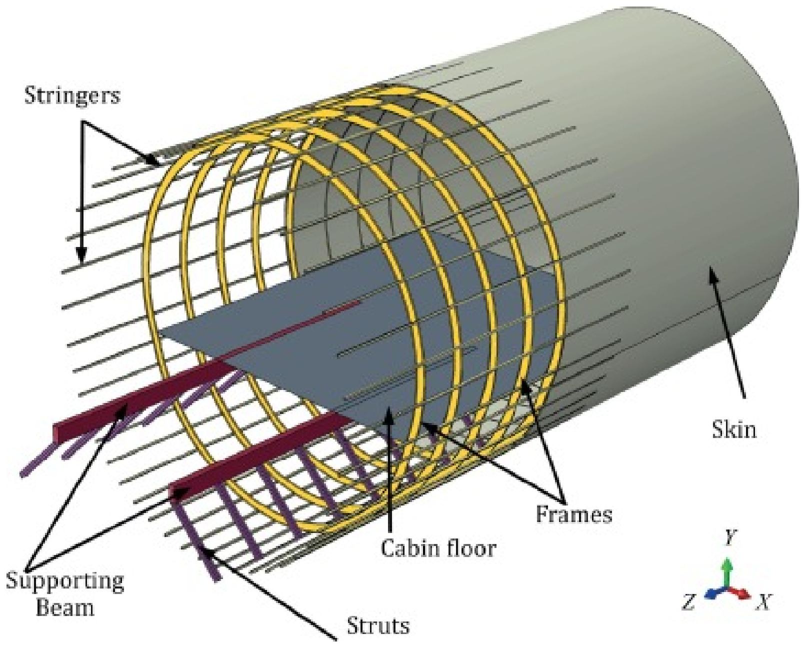

Figure 4. Typical fuselage structural components [Mehreganian and Fallah 2021].

As can be seen in Figure 2, it is expected a logarithmic relation between stress and fatigue life. Therefore, the first preprocessing technique was the application of the logarithm function to the fatigue life.

Initially, a exploratory data analysis (EDA) was performed using Python. The logarithm of life distribution is shown in Figure 5. Here it can be seen that the values are right-skewed (skewness of 1.64) and there are values above 10.

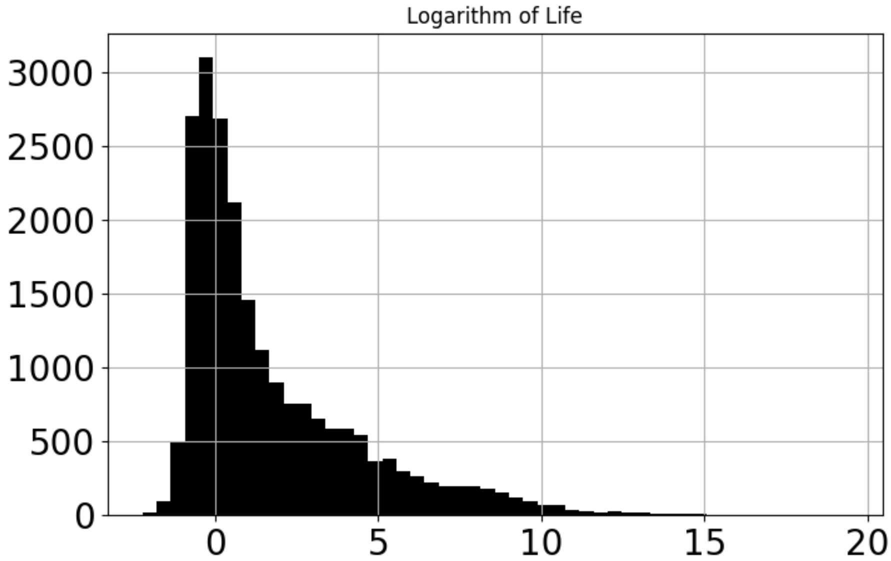

Figure 5. Logarithm of Life histogram.

During the EDA, it was verified correlations between the attributes and the fatigue life. Figure 6 shows an example.

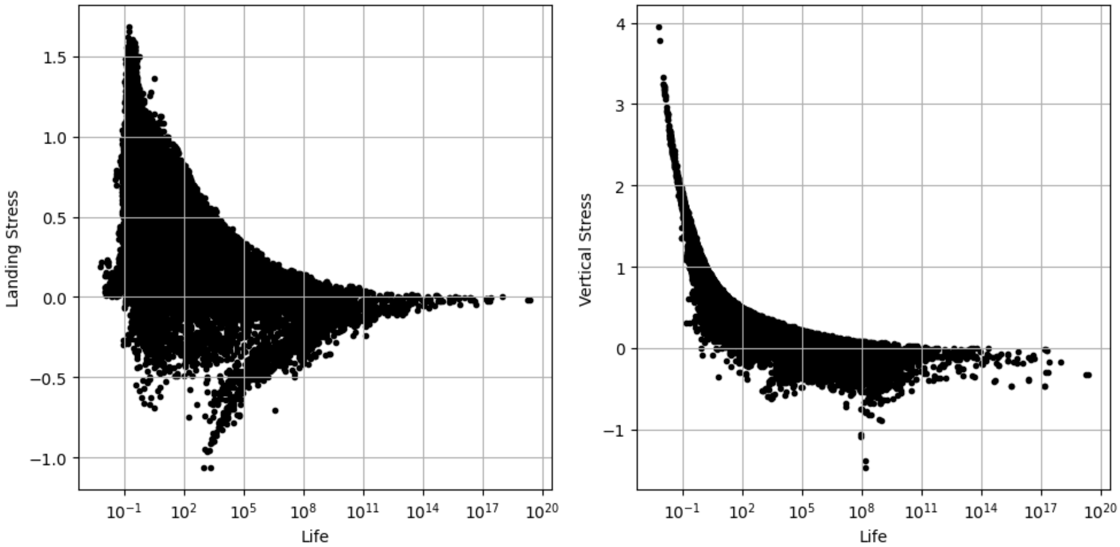

Figure 6. Stress in landing and in vertical gust case versus life.

Compared to Figure 2, it can be seen that the vertical gust case shows less dispersion in terms of predicted life. For both cases, it is possible to see a typical SN curve shape in the upper region of the each graph. After the EDA, it was possible to define the preprocessing strategies in order to apply the learning techniques properly.

2.2. Machine Learning Techniques

In this study, a variety of machine learning algorithms were implemented and evaluated to identify the most effective approach for predicting the structural component lifetime. These techniques were selected due to their diverse learning paradigms and demonstrated performance in handling complex, high-dimensional datasets. Each model was carefully trained and validated using the same data to ensure a fair comparison of their predictive capabilities. The techniques and hyperparameters considered are described below:

- K-Nearest Neighbors (KNN) [Cover and Hart 1967]: number of neighbors and Minkowski distance metric;
- Regression Trees [Breiman et al. 1984]:maximum depth, minimum samples split and criterion;
- Support Vector Machines (SVM): regularization parameter (C) and kernel coefficient (γ);
- Random Forest [Breiman 2001]: number of trees, maximum depth, and the minimum samples split;
- Artificial Neural Networks (ANN): hidden layer sizes, activation functions, learning rate and strength of L2 regularization (α).

2.3. Data Pipeline

In summary, the applied methodology was based on a typical machine learning project data pipeline, as shown in Figure 7.

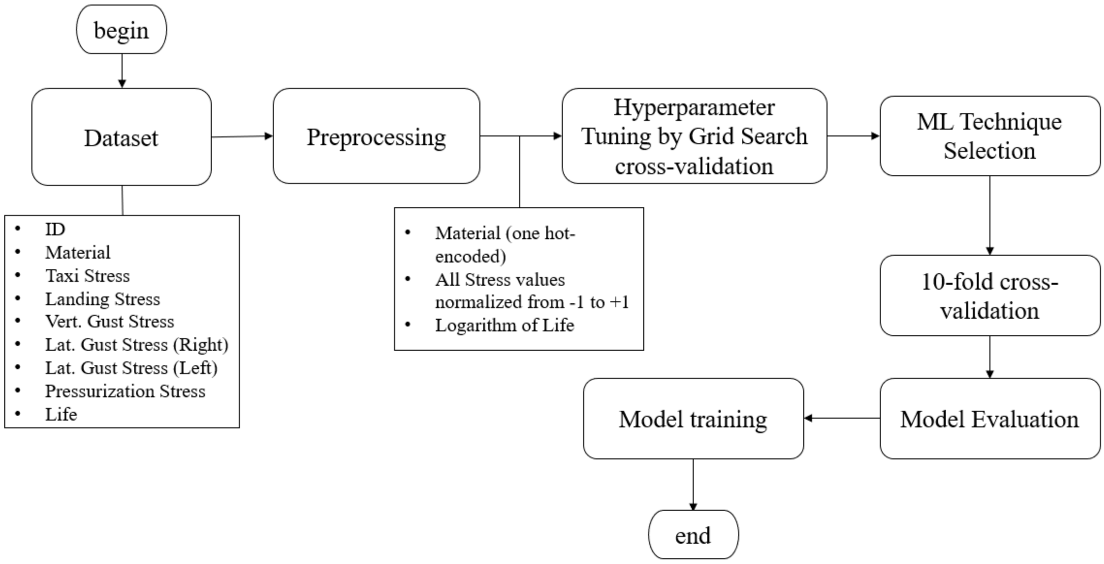

Figure 7. Data pipeline applied to the project.

The preprocessing phase consisted in transforming categorical attributes to numerical values by using the one-hot encoding technique. For the numerical attributes, it was applied a normalization between -1 and 1, in order to consider negative stress values.

In order to optimize the hyperparameters for each technique, it was used the Grid Search method using cross-validation with three folds. With the best parameters for each case, it was performed an evaluation using a new cross-validation with ten folds. By the verification of each technique score, it was possible to select the best one. Finally, the selected method was trained using all the available dataset.

3. Results

In Figure 8, the vertical axis shows the obtained score (RMSE) after the hyperparameters tuning. The values are presented as negative numbers so that higher scores represent better results.

It can be seen that the KNN had the best mean score, followed by the ANN. On the other hand, SVM was the technique with the lowest mean score. The regression tree had the greater interquartile range. As expected, the Random Forest outperformed this model.

The graph from Figure 9 gives a better view how the KNN performed in relation to entire dataset. One can notice that the error margin increases with life. This is a good behavior, since higher lives are less critical. It is also useful to focus on lifespans below three. This range is critical since it defines if a design modification will be necessary to guarantee safety for the entire DSL. Figure 10 shows the distribution for each material.

It can be observed that Material 1 exhibits greater dispersion relative to the ideal line. Furthermore, the predictions for Material 3 are well-fitted for values between 0 and 1.5.

Considering the model application, the KNN was chosen due to several advantages, such

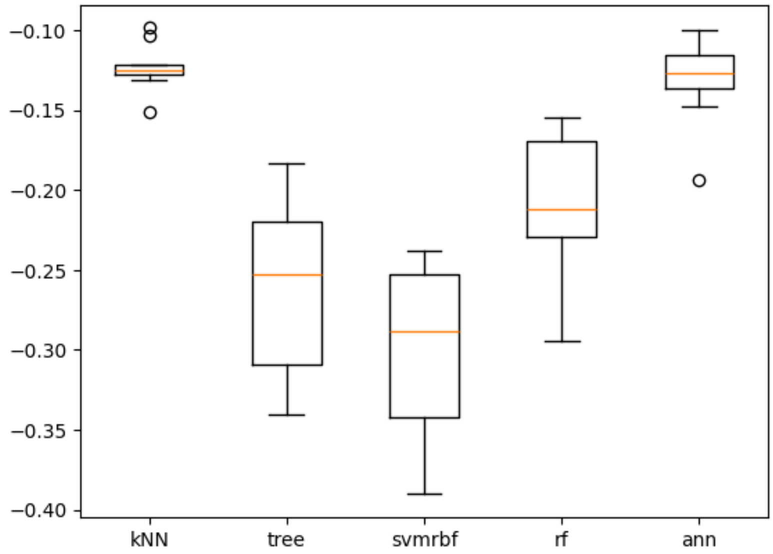

Figure 8. Obtained scores for each technique.

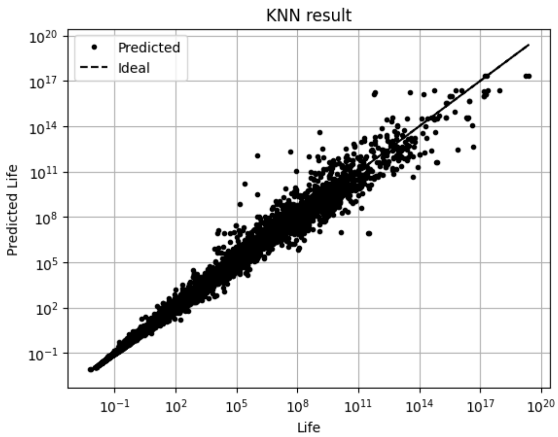

Figure 9. Predicted life vs life in log scale for KNN.

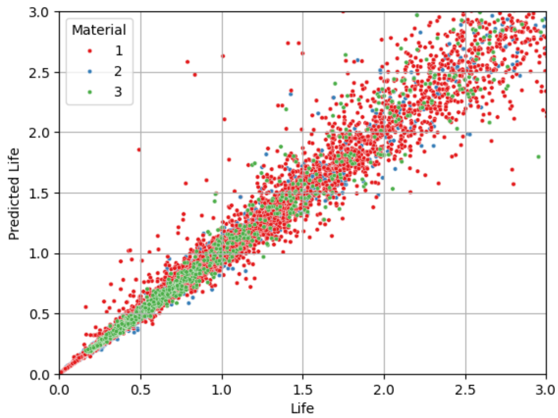

Figure 10. Predicted life vs original life for KNN.

as its simplicity, interpretability, and the fact that it makes no assumptions about the data distribution. The ANN demonstrated good robustness and flexibility; however, its limited interpretability is not a desirable characteristic when supporting well-founded engineering decisions.

In order to simulate an in-production condition, the model's performance was evaluated using a new dataset. Figure 11 shows the result.

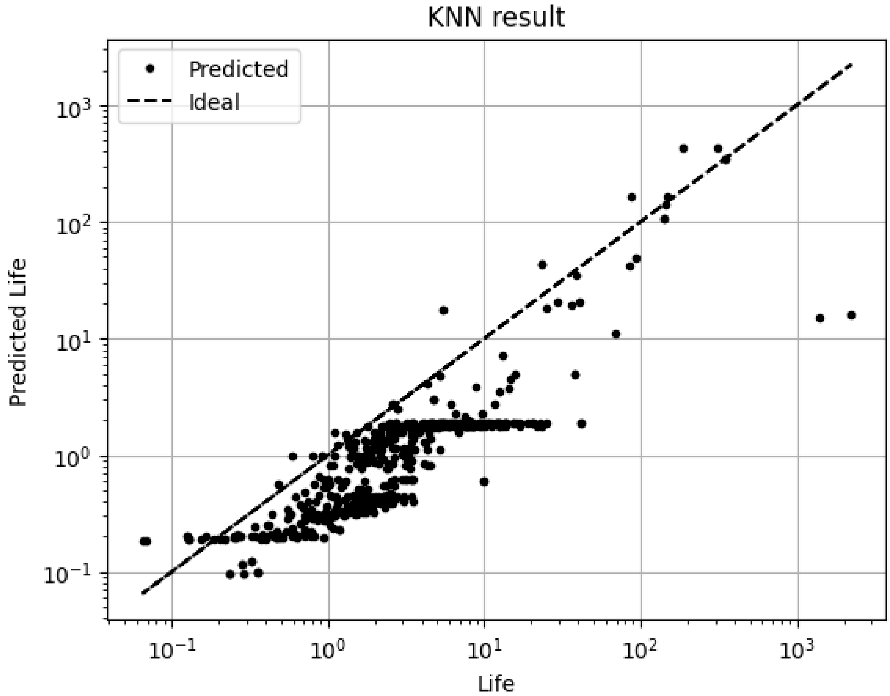

Figure 11. Result from final verification using a new dataset

It can be observed that the predicted lives fall below the ideal line, indicating that the model is conservative compared to the original one. This behavior may be related to the distribution of the target variable used during training, which is right-skewed. However, the model is expected to improve with the adoption of the following strategies:

- Use numerical material properties obtained from tests, instead of treating them as categorical attributes;
- Acquire additional data that covers the entire aircraft structure;
- Develop dedicated models for each material type or component type;

Another important result concerns the processing time compared to the original model. While the complex model typically requires dozens of hours to run, the developed model produces results in just a few seconds, with significantly lower computational cost.

4. Conclusion

It can be concluded that the KNN algorithm demonstrated good performance for the described problem. Due to its simplicity and ease of interpretability, it was chosen over the ANN. Potential improvements can be pursued during the aircraft development phase, using data generated from the various structural analyses conducted by engineers throughout the design cycle. As new data becomes available, additional training can be performed, which will likely lead to improved model performance.

References

Breiman, L. (2001). Random forests. *Machine Learning, 45(1), 5-32.*

Breiman, L., Friedman, J., Olshen, R. A., and Stone, C. J. (1984). *Classification and Regression Trees*. Chapman and Hall/CRC, 1st edition.

Budynas, R. and Nisbett, J. (2020). *Shigley's Mechanical Engineering Design*. New York: McGraw-Hill Education, 11. edition.

Cohen, L. and Hall, A. (1954). Civil aircraft accident. Technical report, Ministry of Transport and Civil Aviation.

Cover, T. and Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory, 13(1), 21–27*.

Dowling, N. E. (2013). *Mechanical Behavior of Materials*. Pearson, 4th edition.

Eslamlou, A. D. and Huang, S. (2022). Artificial-neural-network-based surrogate models for structural health monitoring of civil structures: A literature review. *Soft Computing for Structural Health Monitoring*.

Forrester, A. I. J., Sobester, A., and Keane, A. J. (2008). *Engineering Design via Surrogate Modelling: A Practical Guide*. John Wiley & Sons,, 1st edition.

Khandelwal, S., Basu, S., and Patra, A. (2021). A machine learning-based surrogate modeling framework for predicting the history-dependent deformation of dual phase microstructures. *Materials Today Communications, 29, 451-456*.

Liu, D., Zhang, W., and Yang, Y. (2020). A review of surrogate models and their applications to multi-fidelity modeling and optimization. *Engineering with Computers, 36, 393–415*.

Matsuishi, M. and Endo, T. (1968). Fatigue of metals subjected to varying stress. *Proceedings of the Kyushu Branch, Japan Society of Mechanical Engineers (JSME),Vol. 68, pp. 37–40*.

Mehreganian, N. and Fallah, A. S. (2021). Blast loading effects on aircraft fuselage. *Multiphysics Simulations in Automotive and Aerospace Applications, 239-285*.

Miner, M. A. (1945). Cumulative damage in fatigue. *Journal of Applied Mechanics,Transactions of the ASME, Vol. 12, pp. A159–A164*.

Palmgreen, A. (1924). Die lebensdauer von kugellagern. *Zeitschrift des Vereins Deutscher Ingenieure (VDI),Vol. 68, pp. 339–341*.

Queipo, N. V., Haftka, R. T., Shyy, W., Goel, T., Vaidyanathan, R., and K.Tucker, P. (2005). Surrogate-based analysis and optimization. *Progress in Aerospace Sciences, 41(1), 1–28*.

Schijve, J. (2009). *Fatigue of Structures and Materials*. Dordrecht: Springer, 2nd edition.

Shi, M., Lv, L., Sun, W., and Song, X. (2020). A multi-fidelity surrogate model based on support vector regression. *Structural and Multidisciplinary Optimization, 61, 2363-2375*.

Sun, G. and Wang, S. (2019). A review of the artificial neural network surrogate modeling in aerodynamic design. *Journal of Aerospace Engineering, 233, 2363-2375*.

Sun, H., Wang, H., Liu, H., and Liu, Y. (2019). Surrogate model-based optimization: Recent advances and applications. *Frontiers of Computer Science,13(4), 769–784*.

Suresh, S. (1991). Fatigue of materials: past, present and future. *Acta Metallurgica et Materialia, v. 39, n. 5, p. 1075–1093*.

Williams, B. and Cremaschi, S. (2021). Novel tool for selecting surrogate modeling techniques for surface approximation. *Computer Aided Chemical Engineering, 50, 451-456*.

Zienkiewicz, O. C. and Taylor, R. L. (2000). *The Finite Element Method: Volume 1, The Basis*. Butterworth-Heinemann, 5th edition.