Proceedings of 8th International Conference on Mechanical, Industrial and Energy Engineering 2024
02-04 January, 2025, Khulna, BANGLADESH
SciEn Conference Series: Engineering Vol. 3, 2025, pp 197-200 https://doi.org/10.38032/scse.2025.3.51

Predicting Metal Fatigue Life Under Multiaxial Loading Using Machine Learning: A Comparative Performance Analysis

Showrov Dev Nath, Shafin Ahmed, Md. Abu Mowazzem Hossain*

Department of Mechanical Engineering, Chittagong University of Engineering & Technology, Chittagong-4349, Bangladesh

ABSTRACT
The aim of this research is to increase the accuracy of multiaxial fatigue life prediction using several machine learning models and examining their results. The dataset used in this study consists of 1,167 data points of 40 distinct materials which include some important mechanical parameters and load cases. Support Vector Machine (SVM), Random Forest (RF), Decision Tree Regressor (DTR), AdaBoost and Stacking Regressor were among the studied machine learning models. To begin with, the models were subjected to training and testing on an 80:20 data proportion. Hyperparameter tuning using GridSearchCV was performed for SVM and DTR as their R^2 scores were relatively low. Moreover, the paired t-tests were applied in order to evaluate the difference in the models statistically, while the prediction intervals were calculated in order to illustrate the level of assurance on the predictions. Among all models, the Stacking Regressor achieved the highest R^2 score of 0.84, outperforming RF, DTR, and AdaBoost. Its superior performance demonstrates its potential for real-world applications in predicting fatigue life, such as enhancing reliability in mechanical systems and reducing costs associated with experimental testing. This study underscores the value of ensemble techniques like stacking in addressing complex engineering problems and improving prediction accuracy.

Keywords: SVM, RF, DTR, AdaBoost, Stacking regressor

1. Introduction

Materials and structures which are subjected to repeated multiaxial loading, can lead to fatigue failures. Predicting multiaxial fatigue life accurately under various loading condition is a significant and challenging task in fatigue analysis [1]. Half of all mechanical failures are caused by fatigue [2]. Thus, accurate estimation of fatigue life is a critical aspect of engineering design. Conventional fatigue prediction methods require a lot of experimental data. These methods are expensive and time-consuming. Using machine learning models, the limitations of traditional methods can be surpassed. Machine Learning models learn patterns directly from data, enabling them to approximate complex physical processes [3]. Multiaxial fatigue life is particularly complex to predict compared to uniaxial fatigue due to the simultaneous action of stresses in multiple directions, leading to complex stress-strain interactions and material responses. The effects of stress path, loading history, and phase differences between stresses further complicate multiaxial fatigue behavior. Traditional models often fail to capture these intricate relationships, necessitating advanced approaches such as machine learning. In recent years, a variety of machine learning methods have been developed to predict fatigue life more precisely. Deep learning has been applied by Yang et al. (2021) to get a better prediction result for multiaxial fatigue life [1]. Gao et al. have implemented adaptive neuro-fuzzy inference system (ANFIS) using hybrid models [4]. A year later, Yang et al. (2022) have extended this work by enhancing prediction under varying conditions [5]. Chen et al. have combined machine learning methods with physical principles to create a neural network that was able to make accurate predictions with a small amount of data [6]. Developing a novel technique, Gulgec et al. have evaluated fatigue by measuring strain from acceleration data using deep learning [7]. Collectively, these studies show how machine learning is becoming more and more important for predicting fatigue life and how different models can improve accuracy. Despite these advancements, existing studies often lack versatility in handling highly variable datasets, such as those encountered in multiaxial fatigue life prediction. Ensemble methods, which combine the strengths of multiple models, offer a promising solution by reducing bias and variance. This study aims to address these gaps by applying and evaluating several machine learning models, including an ensemble-based Stacking Regressor, to improve prediction accuracy and reliability.

2. Methodology

This study systematically prepared the dataset, developed predictive models, tuned their parameters, and conducted statistical evaluations to assess the models' reliability. First, the dataset used for this study was introduced, detailing its composition and significance for fatigue life prediction. Next, the models employed were discussed, including the reasons for selecting various machine learning algorithms. Subsequently, the process of hyperparameter tuning was explained to optimize the performance of the predictive models. The statistical analysis followed, where paired t-tests were used to compare different models performances and draw meaningful conclusions. Lastly, confidence interval analysis was conducted to understand the prediction uncertainty,

*Corresponding Author Email Address: devnathshowrov03@gmail.com

providing a more comprehensive evaluation of the models' predictive capabilities. This structured approach ensured a robust methodology for achieving accurate fatigue life predictions.

2.1. Dataset

The dataset [8] underwent several preprocessing steps to ensure suitability for machine learning. Missing values were imputed using mean imputation, and Min-Max Scaling was applied to standardize feature magnitudes within the range [0,1]. Log transformations normalized skewed features like stress and strain, while categorical variables, such as loading paths, were converted into numerical representations using one-hot encoding. Outliers were detected via the IQR method, and extreme cases (< 2%) were removed to maintain data quality. The dataset was then split into training and testing sets (80:20 ratio) using stratified sampling to preserve the distribution of key variables. Finally, the data was vectorized into 241 features to capture complex interdependencies effectively.

2.2. Model development and training

In this study, various machine learning models were used to predict multiaxial fatigue life. The goal was to compare different algorithms and understand their performance in capturing the complex patterns and dependencies within data. Each model was selected for its unique characteristics and strengths in handling different aspects of regression task. The models include Support vector machine (SVM), Random Forest (RF), Decision tree regressor (DTR), AdaBoost regressor, and Stacking regressor. These algorithms were chosen to provide a diverse set of approaches, from linear regression techniques to complex ensemble methods, enabling a comprehensive evaluation of their capabilities in fatigue life prediction. SVM was included because of its ability to handle both linear and non-linear regression tasks effectively. It is particularly well-suited for datasets with fewer features and complex relationships, which are common in fatigue life predictions. Previous studies, such as Chen et al. [6], demonstrated its potential in small-data scenarios, making it a relevant choice for this research. RF is an ensemble learning method that reduces overfitting by combining predictions from multiple decision trees. It was chosen due to its proven success in handling non-linear relationships and high-dimensional data, as shown in studies like Gulgec et al. [7], where RF captured complex physical processes efficiently. The DTR model was selected to explore the dataset's underlying patterns due to its simplicity and interpretability. While DTR is prone to overfitting, its inclusion provided insights into baseline performance and the potential benefits of ensemble techniques. Adaptive Boosting (AdaBoost) was employed for its ability to improve prediction accuracy by sequentially combining weak learners. AdaBoost has been effective in addressing noisy data and has shown success in fatigue life prediction under complex loading conditions, as highlighted by Yang et al. [5]. Stacking was selected as a meta-learning approach, combining predictions from base models to enhance overall performance. Its ability to leverage the strengths of diverse models aligns well with the need to predict multiaxial fatigue life, a highly non-linear and variable-dependent task. Stacking's success in hybrid methods, as noted in Gao et al. [4], made it a compelling choice for this study.

2.3. Hyperparameter Tuning

After initial results, SVM and DTR model showed poor performance. Therefore, GridSearchCV was applied to find the optimal hyperparameters, which improved their performance. For SVM, the regularization parameter (C), kernel type, and kernel coefficient (gamma) were tuned. For DTR max_depth, min_samples_split, and min_samples_leaf were optimized.

2.4 Statistical analysis

Paired t-tests were conducted to statistically compare the performance of the machine learning models. A paired t-test is a statistical method used to determine whether there is a significant difference between the means of two related datasets. In this case, the predictions of different models. It evaluates whether the differences observed are due to random chance or reflect genuine performance differences. This approach was critical for identifying whether the superior performance of certain models, such as the Stacking Regressor, was statistically significant. The Stacking Regressor statistically outperformed the other models due to its ability to combine the strengths of multiple base models (SVM, RF, DTR, AdaBoost) into a single meta-model.

2.5 Confidence interval analysis

Confidence intervals were computed for each model to visually present the uncertainty and confidence in their model predictions.

3. Result and Discussion

3.1. Model performance comparison

The SVM model generated an R^2 score of 0.59, which suggests that the model explains 59 % of the variance in fatigue life. The higher MSE and MAE than that yielded by the other models it seems to suggest SVM had challenges making accurate predictions, pointing towards its sensitivity with complex data relationships or further fine-tuning if required. The Random Forest model had a much better score of R^2= 0.83 explaining 83% of the variance in the target variable. RF also resulted in a low MSE of 0.10 and MAE of 0.25, suggesting more accurate predictions. This behavior can be attributed to the ability of RF to deal with intricate relationships using their ensemble and that it circumvents those deviations in predictions. The R^2 score from the Decision Tree Regressor which is 0.68, and this one is lower than RF or AdaBoost, but higher than SVM. It still had relatively low MSE and MAE which suggests it was flexible enough to be able to get in some of the non-linear patterns but due to end up overfit, especially for without ensemble techniques helps fix these issues. The AdaBoost model performed similarly as RF with the same R^2 of 0.82 meaning that RF and AdaBoost are highly accurate in explaining the variance. Low MSE and MAE represent a higher accuracy of the model and indicates that it can predict fatigue life very close to actual. This significant improvement over the DTR performance can be explained by AdaBoost training weak learners one after another so they correct their mistakes, unlike a single stand-alone decision tree. The highest R^2 score (0.84) and the lowest MSE (0.10) were obtained using a Stacking Regressor with Ridge as the meta-model for this configuration, an MAE somewhat closer to RF and AdaBoost was of 0.25. By grouping it makes concatenation of various versions and thus giving better accuracy with ensemble modeling. Through the aggregation of SVM, RF, DTR and AdaBoost predictions that will allow to reduce bias and variance to generate better accuracy than other models. In conclusion, The Random Forest, AdaBoost, and the Stacking Regressor models were found as the best performing models to predict fixed acidity which outperformed R^2> 0.80 of three these models with significantly lower MSE and MAE. The Stacking Regressor significantly outperformed the other models as it could take advantage of the best of both worlds among multiple algorithms. At the same time, SVM and DTR had significantly lower performance compared to KNN but gained improvements after hyperparameter tuning. This demonstrates the advantages of using ensemble techniques for fatigue life prediction especially in complicated datasets where a single model may miss to capture all important patterns. To better illustrate this, the following bar chart shows how well the stack model performs compared to others in terms of fatigue life, with it being distinctly favored.

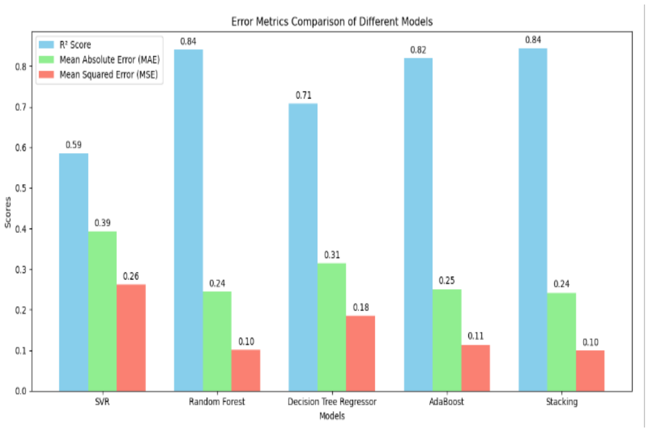

Fig.1 Model performance comparison.

3.2. Confidence Interval Visualizations

The effectiveness of several machine learning techniques, SVM, Random Forest, Decision Tree Regressor, AdaBoost and Stacking Regressor was evaluated based on prediction intervals providing more insights into the accuracy and uncertainty. Concerning the accuracy of the predictions, the maximum agreement between the actual values and forecasts was observed in Random Forest and Stacking Regressor models, which were also characterized by very narrow confidence intervals. AdaBoost was capable of good performance as well, whereas SVM and Decision Tree Regressor had noticeable wider intervals, denoting more uncertainty and somewhat less accurate predictions. In general, the best result efficiency versus confidence levels was achieved through the use of Stacking Regressor.

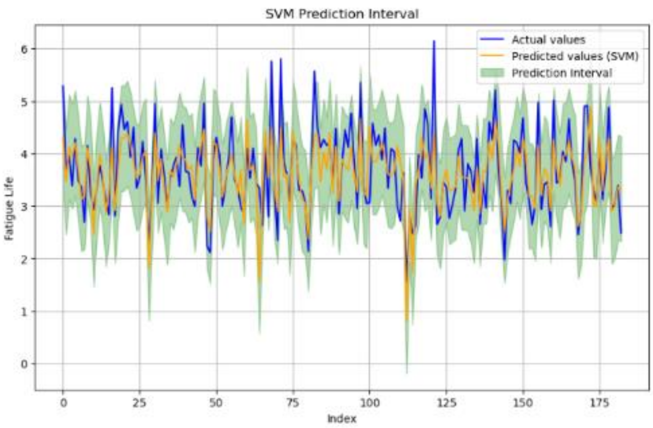

Fig.2 Confidence interval for SVM.

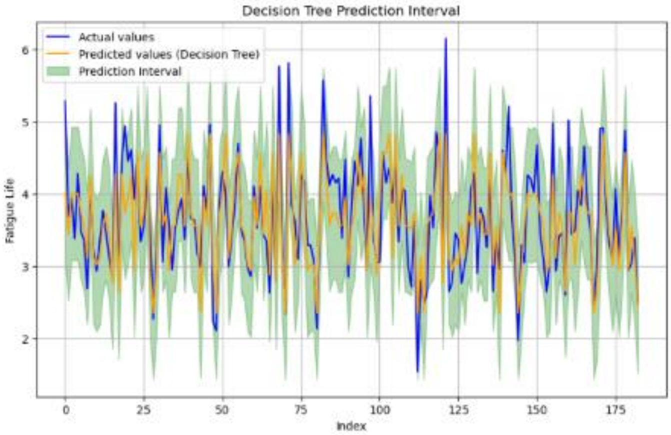

Fig.3 Confidence interval for DTR.

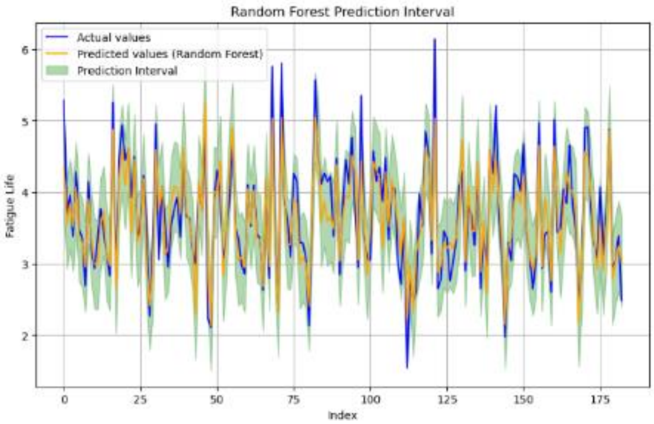

Fig.4 Confidence interval for RF.

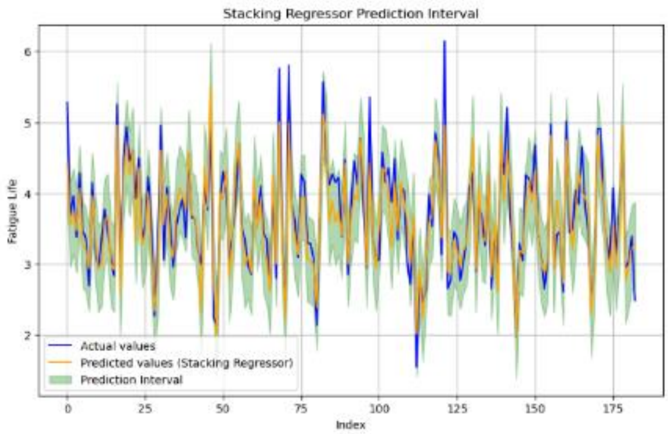

Fig.5 Confidence interval for Stacking regressor.

3.3. Statistical Significance Analysis

The paired t-test analysis compares the performance of various machine learning models (SVM et al.) based on their predictive capabilities for fatigue life. The results indicate that the differences between SVM, Random Forest, Decision Tree, and AdaBoost are not statistically significant, with p-values greater than 0.05 in all comparisons involving these models. This suggests that these models perform similarly in predicting fatigue life based on the dataset used. However, the Stacking Regressor shows statistically significant improvement over both Random Forest (p = 0.0072) and

AdaBoost ($p = 0.0056$), highlighting its superior predictive power. This can be attributed to the Stacking Regressor's ability to combine the strengths of multiple models, such as SVM, Random Forest, Decision Tree, and AdaBoost, into a meta-model (Ridge), leading to better overall performance. The statistically significant results for Stacking suggest that ensemble methods like Stacking offer an advantage in predictive accuracy for complex tasks like fatigue life prediction. Meanwhile, the other models, although effective, may only capture part of the complexity of the data as successfully as the Stacking Regressor. In practical terms, a significant p-value implies that the model's performance is reliably better, which can be crucial for real-world applications, such as selecting the best predictive model for fatigue life estimation in engineering design.

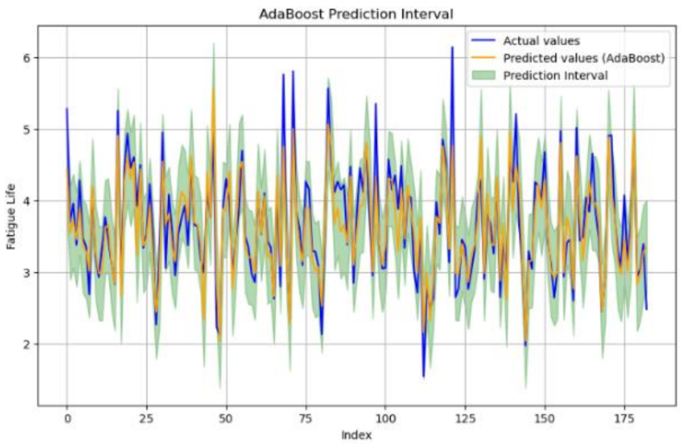

Fig.6 Confidence interval for AdaBoost regressor.

4. Conclusion

This research study was able to illustrate that there are several machine learning models capable of predicting multiaxial fatigue life, with the Staking Regressor being the most effective of all. The Staking Regressor, through the use of a number of base models, achieved the greatest score for R^2, while at the same time minimizing prediction uncertainty. The other models such as SVM, Random Forest and Decision Tree, and AdaBoost even though were able to give good results, did not match the performance of the ensemble method. The findings support the effectiveness of machine learning in predicting fatigue life, more so that of its ensemble models which greatly enhances the predictions. This research provides a basis upon which hybrid and ensemble methods can be researched in respect to the problems encountered in fatigue assessment. The practical implications of this research are noteworthy. By enabling accurate fatigue life predictions, these models can reduce reliance on costly experiments, benefiting industries like aerospace and automotive. However, limitations include the use of a single dataset, which may restrict generalizability, and the computational cost of ensemble methods like Stacking. Future research should focus on expanding datasets, exploring additional material properties, and optimizing models for resource-constrained environments. Furthermore, integrating techniques like transfer learning and hybrid physics-informed approaches could enhance prediction robustness.

References

[1] Yang, J., Kang, G., Liu, Y., & Kan, Q. (2021). A novel method of multiaxial fatigue life prediction based on deep learning. *International Journal of Fatigue*, *151*, 106356.

[2] Kamal, M., & Rahman, M. M. (2018). Advances in fatigue life modeling: A review. *Renewable and Sustainable Energy Reviews*, *82*, 940-949.

[3] Zheng, Z., Li, X., Sun, T., Huang, Z., & Xie, C. (2023). Multiaxial fatigue life prediction of metals considering loading paths by image recognition and machine learning. *Engineering Failure Analysis*, *143*, 106851.

[4] Gao, J., Heng, F., Yuan, Y., & Liu, Y. (2024). A novel machine learning method for multiaxial fatigue life prediction: Improved adaptive neuro-fuzzy inference system. *International Journal of Fatigue*, *178*, 108007.

[5] Yang, J., Kang, G., & Kan, Q. (2022). A novel deep learning approach of multiaxial fatigue life-prediction with a self-attention mechanism characterizing the effects of loading history and varying temperature. International Journal of Fatigue, 162, 106851.

[6] Chen, D., Li, Y., Liu, K., & Li, Y. (2023). A physics-informed neural network approach to fatigue life prediction using small quantity of samples. *International Journal of Fatigue*, *166*, 107270.

[7] Gulgec, N. S., Takáč, M., & Pakzad, S. N. (2020). Structural sensing with deep learning: Strain estimation from acceleration data for fatigue assessment. *Computer-Aided Civil and Infrastructure Engineering*, *35*(12), 1349-1364.

[8] S. Cooh, "Metal Multiaxial Fatigue Life Prediction Using Deep Learning," GitHub Repository. [Online]. Available: https://github.com/stupid-cooh/Metal-Multiaxial-Fatigue-Life-Prediction-Using-Deep-Learning/tree/main. [Accessed: Sept. 20, 2024].