Structural health monitoring of aircraft through prediction of delamination using machine learning

Rajeswari D Corresp., [1] , **Osamah Ibrahim Khalaf** [2] , **Srinivasan R** [3] , **Pushpalatha M** [3] , **Habib Hamam** [4]

[1] Department of Data Science and Business Systems, School of computing, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Tamilnadu, India

[2] Department of Solar, Al-Nahrain Research Center for Renewable Energy, Al-Nahrain University, Jadriya, Baghdad, Iraq

[3] Department of Computing Technologies, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India

[4] Université de Moncton, Moncton, Canada

Corresponding Author: Rajeswari D
Email address: drajiit@gmail.com

Background. Structural Health Monitoring (SHM) is a regular procedure of monitoring and recognizing changes in the material and geometric qualities of aircraft structures, bridges, buildings, and so on. The structural health of an airplane is more important in aerospace manufacturing and design. Inadequate structural health monitoring causes catastrophic breakdowns, and the resulting damage is costly. There is a need for an automated SHM technique that monitors and reports structural health effectively. The dataset utilized in our suggested study achieved a 0.95 R2 score earlier. **Methods.** The suggested work employs SVM + Extra Tree + Gradient Boost + Ada Boost + Decision Tree approaches in an effort to improve performance in the delamination prediction process in aircraft construction. **Results.** The stacking ensemble method outperformed all the technique with 0.975 R2 and 0.023 RMSE for old coupon and 0.928 R2 and 0.053 RMSE for new coupon. It shown the increase in R2 and decrease in root mean square error (RMSE).

Structural Health Monitoring of Aircraft through Prediction of Delamination using Machine Learning

Rajeswari D[1], Osamah Ibrahim Khalaf[2], Srinivasan R[3], Pushpalatha M[3], Habib Hamam [4]

[1]Department of Data Science and Business Systems, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India.
[2] Department of Solar, Al-Nahrain Research Center for Renewable Energy, Al-Nahrain University, Jadriya, Baghdad, Iraq
[3] Department of Computing Technologies, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India.
[4] Department of Electrical Engineering, Université de Moncton, Moncton, Canada.

Corresponding Author:
Rajeswari D[1]
[1]Department of Data Science and Business Systems, School of Computing, College of Engineering and Technology, SRM Institute of Science and Technology, Kattankulathur-603203, Tamil Nadu, India
Email address: drajiit@gmail.com

Abstract

Background. Structural Health Monitoring (SHM) is a regular procedure of monitoring and recognizing changes in the material and geometric qualities of aircraft structures, bridges, buildings, and so on. The structural health of an airplane is more important in aerospace manufacturing and design. Inadequate structural health monitoring causes catastrophic breakdowns, and the resulting damage is costly. There is a need for an automated SHM technique that monitors and reports structural health effectively. The dataset utilized in our suggested study achieved a 0.95 R2 score earlier.

Methods. The suggested work employs SVM + Extra Tree + Gradient Boost + Ada Boost + Decision Tree approaches in an effort to improve performance in the delamination prediction process in aircraft construction.

Results. The stacking ensemble method outperformed all the technique with 0.975 R2 and 0.023 RMSE for old coupon and 0.928 R2 and 0.053 RMSE for new coupon. It shown the increase in R2 and decrease in root mean square error (RMSE).

Keyword: Structural health monitoring, delamination, prediction, stack ensemble, machine learning.

Introduction

The structure of the aircraft is made up of composite materials because of its well-known properties like excellent resistance to fatigue, high strength, weight, high modulus, and stiffness. The carbon composite materials are widely used for manufacturing the aircraft structure [1]. However, the composite materials in the structure are damaged due to aging, fatigue, dynamic load, and cyclic load. Structural Health Monitoring (SHM) plays a vital role in identifying these damages. Inadequate SHM leads to catastrophic failures and the damages caused by catastrophic failure is costly [2]. The factors to be considered for SHM are strain pattern, fiber failure, matrix cracking, delamination, and skin stiffener [3]. This work concentrates on delamination.

The lamination is a collection of laminae. In the structure of aircraft, the lamination has 14 laminas each has 1m.m. to 1.1 m.m. thickness. Laminae is a positioning of unidirectional or woven fibers in a matrix. The fibers are act as a load carrying agent and commonly strong and stiff. The purpose of the matrix is to protect the fibers by distributing the load across it. The layers of amination are made by same matrix material. Due to some interlaminar stresses available on the structure, the lamination starts to delaminate. Gradually it spreads to entire structure. The extant of delamination worsen the characteristics of composite material and finally leads to the failure in aircraft structure [4].

To overcome this failure, the structural health of aircraft should be monitored on timely basis. The current SHM techniques are complicated and time consuming as it is a manual process. This manual SHM requires more resources, i.e., human resources, time and cost. Also, the disassembling and assembling of aircraft structure increases downtime. There is a need for an automated SHM techniques that monitors and reports structural health efficiently. This work focuses on automated SHM techniques and optimizes design, increases safety, reduces downtime, maintenance time and cost.

This work monitors the guided lamb waves in piezoelectric sensor network to identify the delamination in the aircraft structure. Sensor signal features like frequency, load, cycle, Time of Flight (ToF) and Power Spectral Density (PSD) are used to quantify the damage. Machine Learning (ML) algorithm is one of the best techniques to predict the delamination size [5]. In this work, ML based prediction approaches are used to calculate the delamination size. Stack ensemble with linear regression as a meta model and nearest neighbour, extra tree, Ada boost, gradient boost and decision tree algorithm is implemented. The work contributes the following things:

1) The sensor features like PSD, ToF and Interrogation Frequency are calculated for various composite coupons from given actuator and sensor signal.
2) The delamination size was calculated from given X-ray images of multiple composite coupons and considered as ground truth.

3) An ensemble regression technique is used with five base level models to predict the size of delamination.

Literature Review

Toyama et al. presented the variation in stiffness of Carbon Fiber Reinforced Polymers (CFRP) laminates using guided lamb waves [6]. The quantitative damage of laminates was calculated by in situ quantification of the wave velocity. It provides more location accuracy than other conventional technique. Johnson and Chang [7] introduced the two-part verification to find the stiffness and strength of composite laminates. The first part represents the characterization of matrix crack which helps for damage progression. Second part calculated the amount of damage. The proposed technique has been implemented using computer code, PDcell. Saxena et al. [8] experimented how the delamination are influencing on the velocity of guided lamb wave. The density of matrix crack in a particular path and delamination was identified using local regression technique in [9]. Lorrosa et al. classified and predicted damage in [3] and [10]. These researchers clearly indicate the effective use of ML algorithms to classify the data generated from the piezoelectric actuators in the surface of composite materials. Nevertheless, there is no clear method to calculate the delamination size, which is the objective of this work.

ML and deep learning techniques are also used for infrastructure health monitoring. Isaac Osei Agyemang et al. [11] proposed multi-task architecture and ensembleDetNet technique to detect and classify infrastructure damage. This technique improved 5% accuracy than other state-of-art detection and classification technique. Niannian Wang et al. [12] represented faster R-CNN technique with RestNet101 architecture to detect and measure external damage in historic masonry buildings. This proposed methodology identified spalling and efflorescence damage with 0.950 and 0.999 respectively. Dongho Kong and Young-Jin cha [13] used ultrasonic becons instead of GPS in unmanned aerial vehicles (UAV) and CNN for damage identification. This method processed video data collected by UAV and produced 91.9% sensitivity and 97.7% specificity.

Now-a-days ML algorithms are used to analyse the relationship among the features available in a data set is used to predict the damage [14]. The delamination prediction problem is formulated as regression problem. However, much work not been carried out on expansion of delamination using ML method. To address this problem Liu et al. experimented to find the length of the path around the delamination instead of calculating the delamination area [15] by ML methods. This technique provided the solution to the overfitting problem in modelling phase. Though it provides the results in acceptable range, exact calculation of delamination size is remained unsolved. However, the prediction rate of delamination needs to be improved. This work focuses on this.
NASA performed experiments of fatigue aging on CFRP using following ASTM standards D3039 [16] and D3497 [17]. The test was done by using Torayca T700G. These materials are used in aircraft and sports goods which needs high property of composite materials. In composite

materials, weight of the surface is 600 g/m^2, fabric thickness is 0.90mm, density is 1.80 g/cm^3 and tensile strength is 4.900 MPa and it is called as coupon. Finally, it is fabricated and divided into 10-inch length and 6-inch-wide piece is presented in Figure 1 [5].
Huston narrated the effects of fatigue in unidirectional carbon fiber reinforced proxy using residual stiffness and strength model [18]. These results are compared with Chiachio et al. result. The authors taken the fatigue cycling test for each 50,000 cycles then collect the Piezoelectric Transducer (PZT) sensor data of 36 trajectories and 7 interrogation frequencies. The outcome of the fatigue test is 1) malfunction data collection for actuator-to-sensor system. 2) delamination size quantification 3) analyse the variation among coupons. All these outcomes are considered into account for this work. The sample coupon in Figure 1 has 6 actuators and 6 sensors. The lamb waves are disseminated from actuator and sensed by sensors. To calculate the delamination area X-ray images are used and to initiate the delamination at a point, notch with necking geometry was used.
To the best of my knowledge, existing research uses Machine Learning techniques to forecast delamination in aircraft structures. The proposed method predicts the size of the delamination using the ensemble algorithm, which combines one or more ML approaches. Furthermore, the computation of delamination size from X-ray is automated.

Materials & Methods

Dataset Description

The dataset used in this paper is downloaded from NASA Ames Research Center. It is a CFRP materials dataset. It clearly indicates that size of delamination is direct proportional to loading cycle, which was calculated against fatigue cycling. To improve the efficiency of experimental data the calculation was repeated number of times. Figure 2. Represents the X-ray image of composite coupon at 1, 50000 and 100000 loading cycles, sequentially.

The lamb waves are disseminated along the coupon surface to identify the delamination interrogation. The waves propagated through the delamination area has change in its strength while reaching the sensor. The delamination size is increased for increase in loading cycle. While the delamination is increase, the signal strength reaches through the delamination path is reduced. The changes in spectral amplitude of time and frequency domain intimates the delamination on the surface of coupon. To calculate the delamination size, sensor signal features loading cycle, interrogation frequency, Power Spectral Density (PSD) and Time of Flight (ToF) are considered in this paper. To characterize the property of materials the above features are mostly used by the researchers [3][6]. Figure 3 represents the raw sensor and actuator signal for data of CFRP coupon.

Loading cycle: To get the sensor signal for various load, fatigue test is done on composite coupon. The output of every 10000 cycle is recorded. The actuator and sensor signal for various loading cycle is given in the NASA dataset.
Interrogation Frequency: To decompose the actuator and sensor signal into various frequency spectrums, Fast Fourier Transform (FFT) is used. The input frequency correlate with high amplitude is considered as interrogation frequency.
Power Spectral Density: PSD for various frequencies is calculated using FFT by the function of time. The peak in the PSD values is reduced by increase in delamination size. However, the strength of the signal input is reduced due to wave scattering in delamination area.
Time of Flight: The time difference between the actuator signal peak and sensor signal peak is ToF.

Figure 4 represents the association between features in the dataset with scatter plots depends on the correlation matrix method is shown in equ. 1.

$$X_{mn} = \frac{\sum_{i=1}^{p}(m_i - \overline{m})(n_i - \overline{n})}{\sqrt{\sum_{i=1}^{p}(m_i - \overline{m})^2 \sum_{i=1}^{p}(n_i - \overline{n})^2}} \quad (1)$$

The x_{mn} is correlation coefficient, m and n are random variables and $\overline{m}$ and $\overline{n}$ are the means of m and n. The scattering of sensor signal feature is replicated on left axis and bottom axis and the diagonal represents the density plot of the feature.

Figure 5 represents the correlation between the pair of features. None of the correlation value exceeds 0.8, it clearly indicates that the features are not closely correlated with each other, and all the features are taken into account for further process. The figure also represents that there is a negative correlation between cycle and PSD.

MATLAB is used to process the raw data given by NASA to obtain the specified features. To calculate the ground truth (i.e., delamination size) Area property of region props method is used on X-ray images with delamination in MATLAB. Finally, the dataset has 150949 data points with 6 features like cycle, load, frequency, PSD, ToF and ground truth.

Delamination size prediction using machine learning

In this work, sensor signal features acquired from composite coupon is used to predict the delamination size. A deterministic technique is entrenched by regression investigation which permits the diagnostic values obtained by independent variable n specified the dependent variables m_x. Figure 6 shows the workflow of the prediction technique. The four sensor features are formed as the vector m_x= [m_1, m_2, m_3, m_4] = [cycle, frequency, PSD, ToF] as input to the prediction method. Delamination size is used as the ground truth n.

In recent years, ML algorithms are widely used to predict the delamination size regression problem and provided the best results [5]. Consequently, this work implemented the regression models like support vector machine (SVM), Extra tree, Gradient Boost, Ada Boost and Decision tree and finally, stack ensemble technique is used to improve the prediction accuracy.

Support vector machine: The SVM is basic and widely used prediction technique. Due to SVM's scalable capability, it can be well suited to small datasets. With the help of loss function SVM can be applied to prediction problems. In this work, SVR (support vector regression) with 'rbf' kernel is used. The degree of polynomial kernel method is set as 3, kernel coefficient for 'rbf' is set as scale, value for gamma is set as 1/ (n_features * X.var()) and stopping condition tolerance is set as 1e-3 by default. Kernal size used for this implementation is set as 200MB. SVM regression technique is presented in algorithm 1 [19].

Algorithm 1: SVM Model

Result: Prediction of delamination size

Input: Sensor features with ground truth $(m_p, n_p)_{p=1}^{x}$

1 clf = svr (m_p, n_p)
2 clf.fit (k='rbf', degree=3, g='scale', tol=0.001, C=1.0, c_size=200, m_it=- 1)

Output: SVM Prediction model

In the above algorithm k represents kernel, g represents gamma, c_size represents cache_size and m_it represents maximum iteration.

Extra tree Model: The extra tree model contains number of prediction trees capitulated from various training data [20]. Every tree is considered as self-prediction method and average of every prediction tree's output gives the final regression. Extra tree regression technique is presented in algorithm 2. The increase in number of prediction trees yields to better performance. In this work, the amount of prediction tree available in forest is 100, mean squared error criterion, the amount of samples needed in leaf node is 1, amount of samples needed to divide in internal node is 2 are used.

Algorithm 2: Extra Tree Model

Result: Prediction of delamination size

Input: Sensor features with ground truth $(m_p, n_p)_{p=1}^{x}$

1 clf = ExtraTreeRegressor (m_p, n_p)
2 clf.fit (n_est=100, c='squared_error', m_s_s=2, m_s_l=1, max_features='auto')

Output: Extra Tree Prediction model

In the above algorithm n_est represents n_estimators, c represents criterion, m_s_s represents min_sample_split, m_s_l represents min_samples_leaf and m_ft represents max_features.

Gradient Boosting Model: Gradient Boosting is a supplement model in an onward step-wise technique. It permits for improvement of random differentiable loss method. At every epoch a prediction tree is fit on the negative gradient of the specified loss method [21]. Gradient boosting technique generate a regression technique in the structure of an ensemble of weak regression technique. Gradient boosting regression technique is presented in algorithm 3. Squared error loss function is used for regression. The contribution of prediction tree shrinks by learning rate and it was set as 0.1. The increase in boosting epoch provides good performance and it was set as 100.

Algorithm 3: Gradient Boosting model

Result: Prediction of delamination size
Input: Sensor features with ground truth $(m_p, n_p)_{p=1}^{x}$

1 clf = GradientBoostingRegressor (m_p, n_p)
2clf.fit (loss='squared_error', learn_r=0.1, n_est=100, subsample=1.0, c='friedman_mse', m_s_s=2, m_s_l=1, max_depth=3, alpha=0.9, valid_frac=0.1)
Output: Gradient Boosting Prediction model

In the above algorithm n_est represents n_estimators, c represents criterion, m_s_s represents min_sample_split, m_s_l represents min_samples_leaf, learn_r represents learning_rate, valid_frac represents the validation_fraction.

AdaBoost model: An AdaBoost regressor is a meta-estimator.It starts by fixing a prediction on the given dataset, after that fixes extra copy of the predictor to the coupled dataset [22]. The instance weights are modified depends on the error of present regression. In essense, the successive predictors concentrate on hard instances. AdaBoosting regression technique is presented in algorithm 4. The highest amount of estimates used till boosting is stopped is set as 50. The weight put into every predictor at every boosting epoch is called as learning rate. The increase in learning rate, improves the benefaction of every predictor. The learning rate is set as 1. After every boosting epoch, the weights are getting changed by loss function. The linear loss function is used.

Algorithm 4: AdaBoost model

Result: Prediction of delamination size
Input: Sensor features with ground truth $(m_p, n_p)_{p=1}^{x}$

1 clf = AdaBoostRegressor (m_p, n_p)
2clf.fit (n_estimators=50, learning_rate=1.0, loss='linear')
Output: AdaBoost Prediction model

In the above algorithm n_est represents n_estimators, learn_r represents learning_rate.

Decision Tree model: Decision tree is a non-criterion supervised learning technique. The main is to produce a technique that regress the estimate of a desired variable through studying effortless decision rules worked out from the data features [23]. The trees are known as piecewise constant imprecision. The decision trees study from data to imprecise a sine curve with group of if-then-else decision rules. Decision tree regression technique is presented in algorithm5. The method to calculate the standard of a split is known as criterion. Squared error criterion is used. The amount of samples needed to divide an internal node is set as 2 and amount of samples needed at leaf node is set as 1.

Algorithm 5: Decision Tree Model

Result: Prediction of delamination size

Input: Sensor features with ground truth $(m_p,n_p)_{p=1}^{x}$

clf = DecisionTreeRegressor (m_p, n_p)

clf.fit (criterion='squared_error', splitter='best', min_samples_split=2, min_samples_leaf=1)

Output: Decision Tree Prediction model

In the above algorithm c represents criterion, m_s_s represents min_sample_split, m_s_l represents min_samples_leaf.

Ensemble model: The ensemble technique takes on several base prediction techniques, whose regression accuracy is best than any other learning model. It is contrast from the ensemble technique in statistical devices, which is normally limitless. This ensemble-based ML technique increases the pliable structure of alternate technique who is finite [24].

This work used, stacking ensemble. It is an ambiguous loss-based ML framework. Stack ensemble comprises in stacking the output of separate regressor and utilize a predictor to calculate the end prediction. Stack ensemble permits to utilize the robustness of every separate predictor by utilizing their result as input to end predictor. The base regressor used for ensemble technique in this work is SVM, extra tree, gradient boosting, Adaboost and decision tree. Consequently, the base regressor techniques are implemented separately is presented in algorithm 6.

Algorithm 6: Stacking ensemble model

Result: Prediction of delamination size.

Input: Sensor features with ground truth $(m_p,n_p)_{p=1}^{x}$.

Output: Stacking ensemble E.

Step 1: Develop base-level models CLF on E

Perform n-fold cross-validation on base level models

clf_1 = svr (m_p, n_p)

clf_2 = ExtraTreeRegressor (m_p, n_p)

clf_3 = GradientBoostingRegressor (m_p, n_p)

clf_4 = AdaBoostRegressor (m_p, n_p)

clf_5 = DecisionTreeRegressor (m_p, n_p)

Step 2: Construct the level-on data M

$M = \{\hat{m}_p, n_p\}_{p=1}^{x}, where$

$\hat{m}_p = \{clf_1(m_p), clf_2(m_p), clf_3(m_p), clf_4(m_p), clf_5(m_p)\}$

Return E comprising of CLF models

Results

To prove the efficiency of the proposed work, raw sensor data collected from one composite coupon is considered. The database was constructed with needed features for predicting delamination are ToF, cycle, frequency, PSD, and ground truth, i.e., delamination size. At final, the data set has 150949 data points. From this, 75% data points are used to train the model and remaining 25% data points are used for testing. First, SVM regression technique was implemented with RBF kernel, but the prediction results were not preferable. Hence, further regression techniques like extra tree, gradient boost, ada boost, decision tree were used to predict the delamination size. Finally, stack ensemble technique was used to combine the above said regression techniques.

MATLAB is used to process raw sensor data before building the data collection. Six piezoelectric actuators and six piezoelectric transducer (PZT) sensors are included in the composite coupon to collect raw data. Python scikit learn runs machine learning algorithms on an i5 processor, 8 GB of RAM, and Windows 10.

Model Estimation:

To calculate the efficiency of machine learning model used for delamination prediction is estimated using Root Mean Square Error (RMSE) and coefficient of determination (R^2). The formulas are as follows:

$$RMSE = \sqrt{\frac{1}{x}\sum_{i=1}^{x}(p_i - \hat{p}_i)^2} \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{x}(p_i - \hat{p}_i)^2}{\sum_{i=1}^{x}(p_i - \bar{p}_i)^2} \quad (3)$$

Where RMSE is absolute estimate to fit, the less RMSE is best estimate to fit and R^2 is a relative measure to fit, it varies from 0 to 1, the high R^2 specify a better model.

Figure 7 illustrates the R2 and RMSE values when combining two ML techniques. Ensembling is the combination of one or more approaches that enhance the outcome of the SHM procedure. Ensembles are very good at preventing overfitting, improving generalization, and handling noisy or inconsistent data. They provide a robust solution to a wide range of datasets, as different models may thrive in different areas of feature space. Furthermore, ensemble approaches are less susceptible to hyperparameter tuning and outliers, making them more durable and adaptive to a variety of real-world circumstances. Overall, the diversity and aggregation of numerous models inside an ensemble framework result in more robust, accurate, and reliable predictions in machine learning applications. This stage involves evaluating the efficiency of combining two strategies. Combining Gradient Boost with Decision Tree surpasses all other models in terms of maximising R2 and decreasing RMSE.

Figure 8 illustrates the effectiveness of combining three strategies. The comparison of figures 7 and 8 illustrates the performance, which demonstrates that gradient boost combined with other approaches delivers superior results in comparison to other approaches. Similarly, the combination of three ML methods does not outperform the combination of two ML methods. This demonstrates that the combination of three ML approaches does not always yield positive results. This combination of ML techniques yields results dependent on the characteristics of the dataset.

Figure 9 depicts the result of combining four methods. Combining three ML techniques and four ML techniques. Observing figures 8 and 9 demonstrates conclusively that integrating multiple ML algorithms does not produce optimal results for all datasets. Before utilising ensembling techniques, thoroughly examine the test data and then use the appropriate combination of ML algorithms.

After analysing each and every test data, combination of five ML approaches forms an ensembling approach. The Combination ML Methods (SVM, Extra Tree, Ad Boost, Gradient Boost and Decision Tree) outperforms the best result compared to individual Ml methods as well as combination two, three and four ML Methods. The evidence is provided in the Table 1.

Table 1 represents the evaluation result of each separate model and stacking (ensemble) model. The evaluation result shows that ensemble model outperforms all the single model with lowest RMSE and highest R^2 value. Table 2 represents the evaluation result of each model and stacking ensemble model for a new composite coupon which was not trained yet. The new composite coupon is made up of different materials and tries to check the performance of ensembling techniques.

Several cause for error in the accuracy of prediction are delamination area calculation (ground truth) in MATLAB. Sensed signal orientation, external noise affected the sensed signal and less amount of data. Also, there are some technical difficulties for constructing the data set which may cause some error in delamination size prediction.

Ensembling techniques are tested against the new coupon and analyse the performance metrics of R^2 and RMSE. The comparison of new coupon and old coupon are displayed in the Figure 10. Even though the prediction accuracy is less than old coupon, the ensemble model outperforms all the single model with lowest RMSE and highest R^2 value.

Discussion

The experimental assessment shows an efficient technique for delamination prediction using machine learning model. In this research work, ensemble technique produces better accuracy with error rate, because ensemble technique has the strength of each regression technique and acted better than each technique.

The SVM, in particular, wraps the perseverance of the variables for a given usefulness of the method, kernel variables, and kernel possibility. The SVM technique assures the difficulty of overfitting from variable enhancement to procedure choice. Nonetheless, kernel approaches will be entirely diplomatic in terms of overfitting the technique determination criteria. In a decision tree, it will be difficult to evaluate all possible attribute combinations in order to find unseen data with deprecatory failure. The decision tree focuses on discovering errors by distinguishing between success and error data. Extra trees are typically powerful for discrepancy. Anyway, due of its proclivity for overfitting, it is prone to sampling errors. When the testing data set differs significantly from the training data set, the extra tree cannot be fitted. Overfitting is possible with boosting approaches (gradient and Ada boost), and the maximum number of regression trees is not allowed for one.
Each regression technique has its own advantage and disadvantage when these features are interrelated. Accordingly, the stacking ensemble technique, take in from each regression technique's advantages to balance their disadvantages, accomplishing correctly in together or more than the best individual technique with reference to improving the prediction accuracy. The main strength of the stacking ensemble model is, considered each separate regression technique and taken their advantage and produced better accuracy for given data set.

Conclusions

The primary outcome of this research is to focusses on finding the suitable ML algorithm to predict the delamination size in the structure of the aircraft. The work represented in this paper focuses on construct a damage assessment technique for structural health monitoring of aircraft. In this paper, the damage assessment mainly aims in designate the increase of delamination in composite materials. This work shown a innovative approach to identify the damaged area through delamination size prediction with machine learning model. Five machine learning techniques with stacking ensemble approach were used to identify the size of delamination in a composite coupon. Analysed the results produced by SVM, Extra tree, Ada boost, Gradient boost, decision tree and stacking ensemble technique, the stacking ensemble method outperformed all the technique with 0.975 R^2 and 0.023 RMSE for old coupon and 0.928 R^2 and 0.053 RMSE for new coupon. It shown the increase in R^2 and decrease in root mean square error (RMSE).

The features frequency, cycle, ToF and PSD alone considered in this paper. Adding more features will increase the performance. Other than delamination, skin stiffener, matrix cracking, stain patterns, fiber failure also need to be considered while monitoring the structural heath of aircraft. These things will be concentrated in future work.

Funding Interest: The authors thank Natural Sciences and Engineering Research Council of Canada (NSERC) and New Brunswick Innovation Foundation (NBIF) for the financial support of the global project. These granting agencies did not contribute in the design of the study and collection, analysis, and interpretation of data.

References

1) Yue N, Aliabadi MH. A scalable data-driven approach to temperature baseline reconstruction for guided wave structural health monitoring of anisotropic carbon-fibre-reinforced polymer structures. Structural Health Monitoring. 2020;19(5):1487-1506. doi:10.1177/1475921719887109
2) W. Xu, N. Duan, S. Wang, Y. Guo, and J. Zhu, “Modeling and measurement of magnetic hysteresis of soft magnetic composite materials under different magnetizations,” IEEE Transactions on Industrial Electronics, vol. 64, no. 3, pp. 2459–2467, 2017.
3) C. Larrosa, K. Lonkar, and F.-K. Chang, “In situ damage classification for composite laminates using gaussian discriminant analysis,” Structural Health Monitoring, vol. 13, no. 2, pp. 190–204, 2014.
4) Seno AH, Aliabadi MF. Uncertainty quantification for impact location and force estimation in composite structures. Structural Health Monitoring. 2022;21(3):1061-1075. doi:10.1177/14759217211020255
5) H. Liu, S. Liu, Z. Liu, N. Mrad and A. S. Milani, "Data-Driven Approaches for Characterization of Delamination Damage in Composite Materials," in IEEE Transactions on Industrial Electronics, vol. 68, no. 3, pp. 2532-2542, March 2021, doi: 10.1109/TIE.2020.2973877.
6) N. Toyama, S. Yashiro, J. Takatsubo, and T. Okabe, “Stiffness evaluation and damage identification in composite beam under tension using lamb waves,” Acta materialia, vol. 53, no. 16, pp. 4389–4397, 2005.
7) P. Johnson and F.-K. Chang, “Characterization of matrix crack-induced laminate failure part ii: Analysis and verifications,” Journal of composite materials, vol. 35, no. 22, pp. 2037–2074, 2001.
8) A. Saxena, K. Goebel, C. C. Larrosa, V. Janapati, S. Roy, and F.- K. Chang, “Accelerated aging experiments for prognostics of damage growth in composite materials,” DTIC Document, Tech. Rep., 2011.
9) Z. Su, L. Ye, and Y. Lu, “Guided lamb waves for identification of damage in composite structures: A review,” Journal of sound and vibration, vol. 295, no. 3, pp. 753–780, 2006.
10) C. Larrosa, V. Janapati, S. Roy, and F.-K. Chang, “In-situ damage assessment of composite laminates via active sensor networks,” in Aircraft Airworthiness and Sustainment Conference 2011, 2011, pp. 1– 10.
11) Y. Lei, F. Jia, J. Lin, S. Xing, and S. X.Ding, “An intelligent fault diagnosis method using unsupervised feature learning towards mechanical big data,” IEEE Transactions on Industrial Electronics, vol. 63, no. 5, pp. 3137–3147, 2017.
12) H. Liu, S. Liu, Z. Liu, N. Mrad, and H. Dong, “Prognostics of damage growth in composite materials using machine learning techniques,” in Industrial Technology (ICIT), 2017 IEEE International Conference on. IEEE, 2017, pp. 1042–1047.
13) A. D3039, “Standard test method for tensile properties of polymer matrix composite materials,” American Society of Testing and Materials: West Conshohocken, PA, 2000.
14) A. D3479, “Standard test method for tension-tension fatigue of polymer matrix composite materials,” American Society of Testing and Materials: West Conshohocken, PA, 2000.
15) R. Huston, “Fatigue life prediction in composites,” International journal of pressure vessels and piping, vol. 59, no. 1-3, pp. 131–140, 1994.

16) Rohit, R.V.S., Chandrawat, D., Rajeswari, D. (2021). Smart Farming Techniques for New Farmers Using Machine Learning. In: Mahapatra, R.P., Panigrahi, B.K., Kaushik, B.K., Roy, S. (eds) Proceedings of 6th International Conference on Recent Trends in Computing. Lecture Notes in Networks and Systems, vol 177. Springer, Singapore.
17) D. Sikka, Shivansh, R. D and P. M, "Prediction of Delamination Size in Composite Material Using Machine Learning," 2022 International Conference on Electronics and Renewable Systems (ICEARS), 2022, pp. 1228-1232, doi: 10.1109/ICEARS53579.2022.9752123.
18) X. Ma, C. Ding, S. Luan, Y. Wang, and Y. Wang, "Prioritizing influential factors for freeway incident clearance time prediction using the gradient boosting decision trees method," IEEE Transactions on Intelligent Transportation Systems, 2017.
19) G. Rätsch, T. Onoda, and K.-R. Müller, "Soft margins for adaboost," Machine learning, vol. 42, no. 3, pp. 287–320, 2001.
20) H. Soni, P. Arora and D. Rajeswari, "Malicious Application Detection in Android using Machine Learning," 2020 International Conference on Communication and Signal Processing (ICCSP), 2020, pp. 0846-0848, doi: 10.1109/ICCSP48568.2020.9182170.
21) D. Rajeswari, S. R, R. S and P. M, "Intelligent Refrigerator using Machine Learning and IoT," 2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI), 2022, pp. 1-9, doi: 10.1109/ACCAI53970.2022.9752587.
22) Isaac Osei Agyemang, Xiaoling Zhang, Daniel Acheampong, Isaac Adjei-Mensah, Goodlet Akwasi Kusi, Bernard Cobbinah Mawuli, Bless Lord Y. Agbley, Autonomous health assessment of civil infrastructure using deep learning and smart devices, Automation in Construction,Volume 141, 2022, 104396, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2022.104396.
23) Niannian Wang, Xuefeng Zhao, Peng Zhao, Yang Zhang, Zheng Zou, Jinping Ou, Automatic damage detection of historic masonry buildings based on mobile deep learning, Automation in Construction, Volume 103, 2019, Pages 53-66, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2019.03.003.
24) Dongho Kang,Young-Jin Cha, "Autonomous UAVs for Structural Health Monitoring Using Deep Learning and an Ultrasonic Beacon System with Geo-Tagging", computer aided civil and infrastructure engineering, volume 33, Issue 10, 2018

Figure 1

Composite coupon

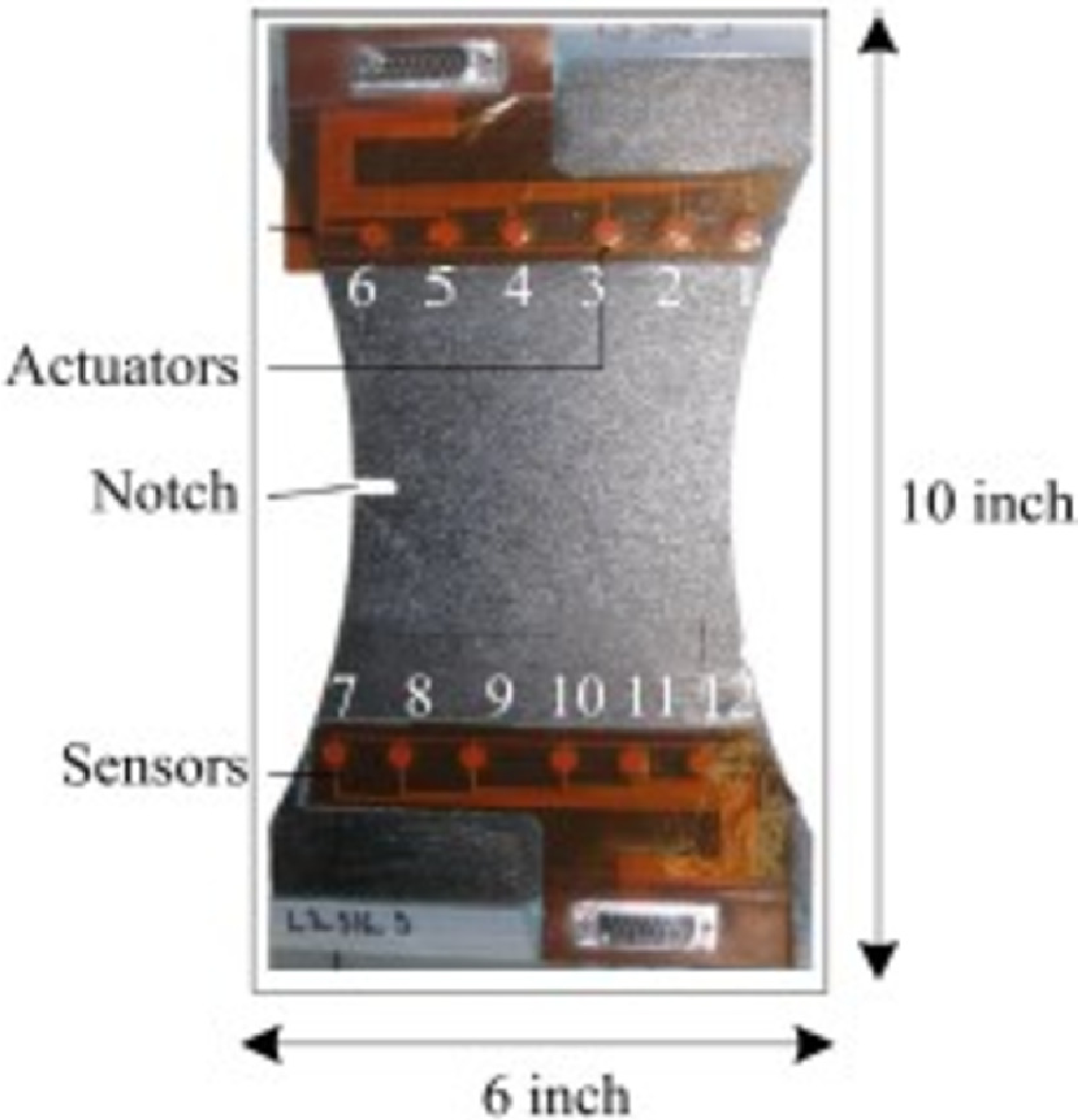

Figure 2

X-ray images for various loading cycles

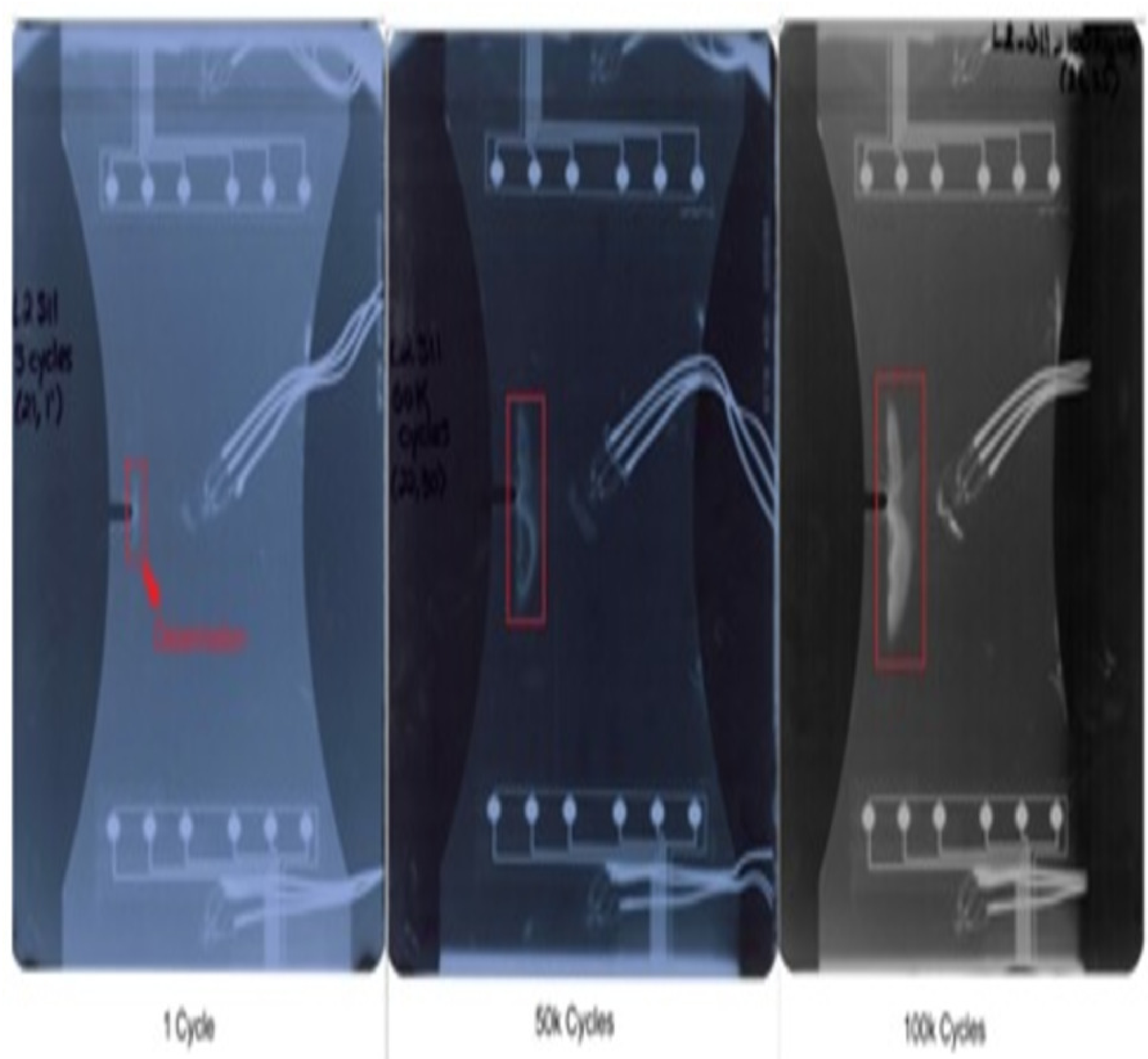

Figure 3

Sensor and actuator signal for delamination data of CFRP coupon

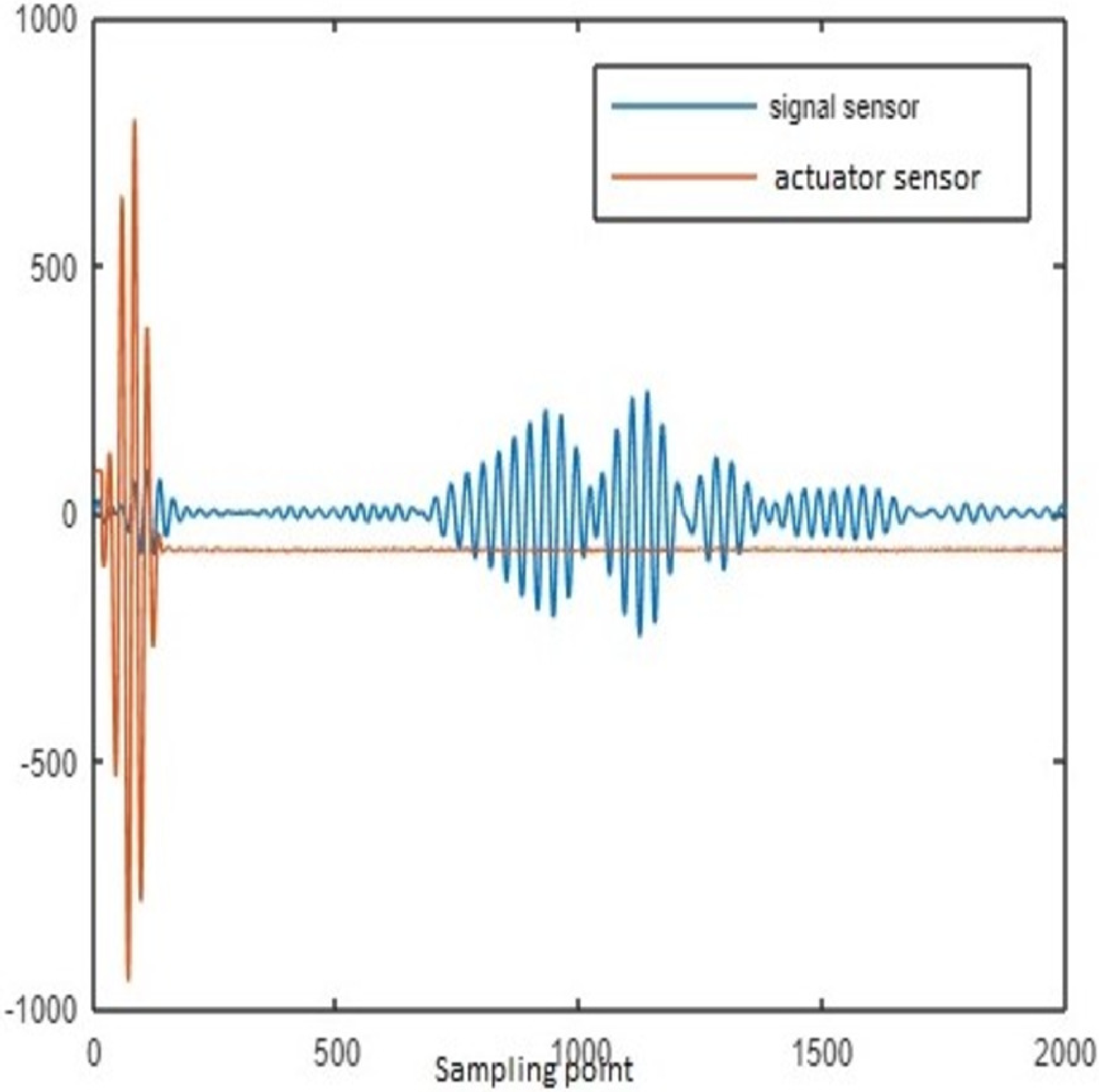

Figure 4

Correlation between features

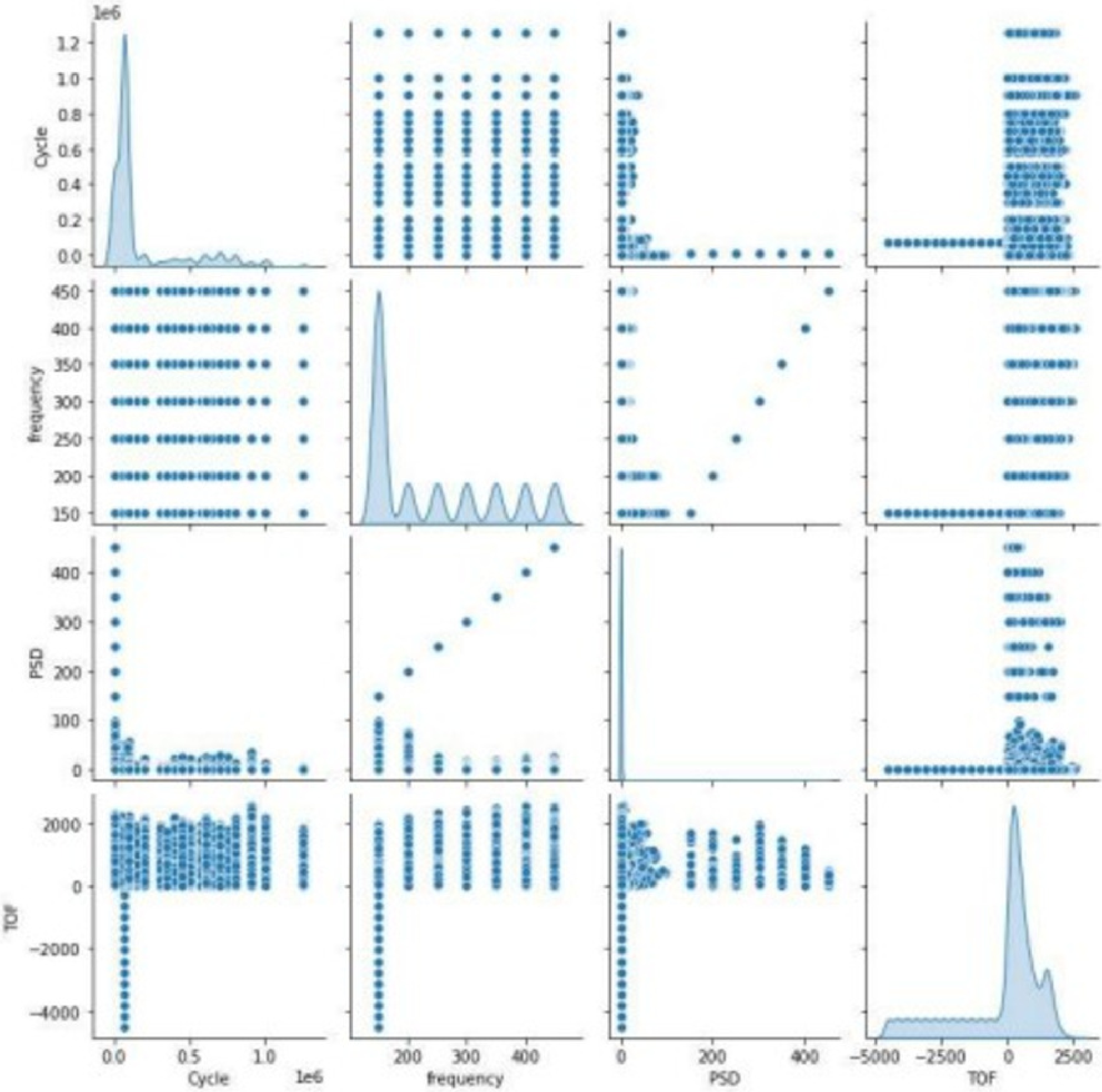

Figure 5

Relationship between pair of features

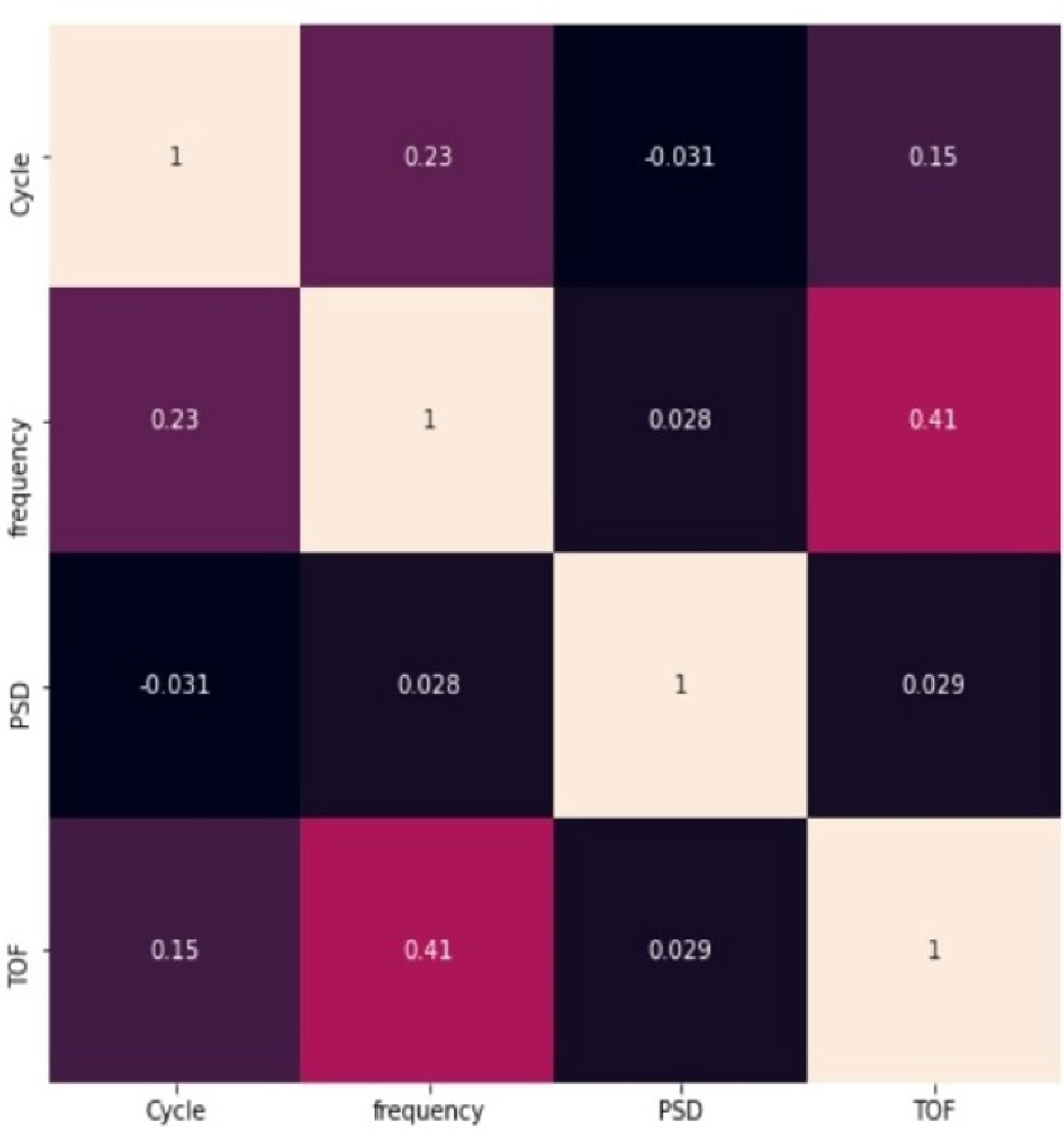

Figure 6

Architecture of the proposed model

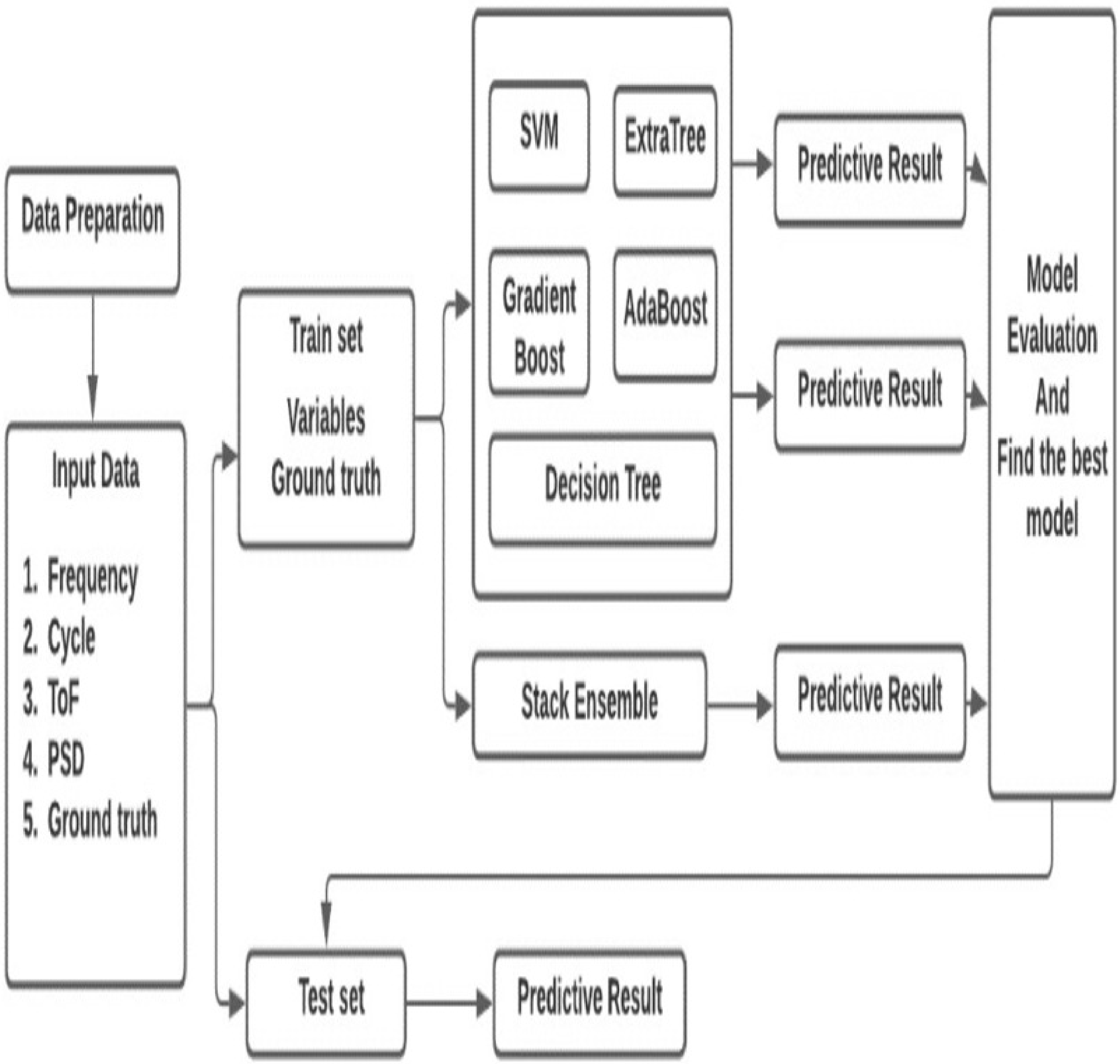

Figure 7

Ensembling using 2 Methods

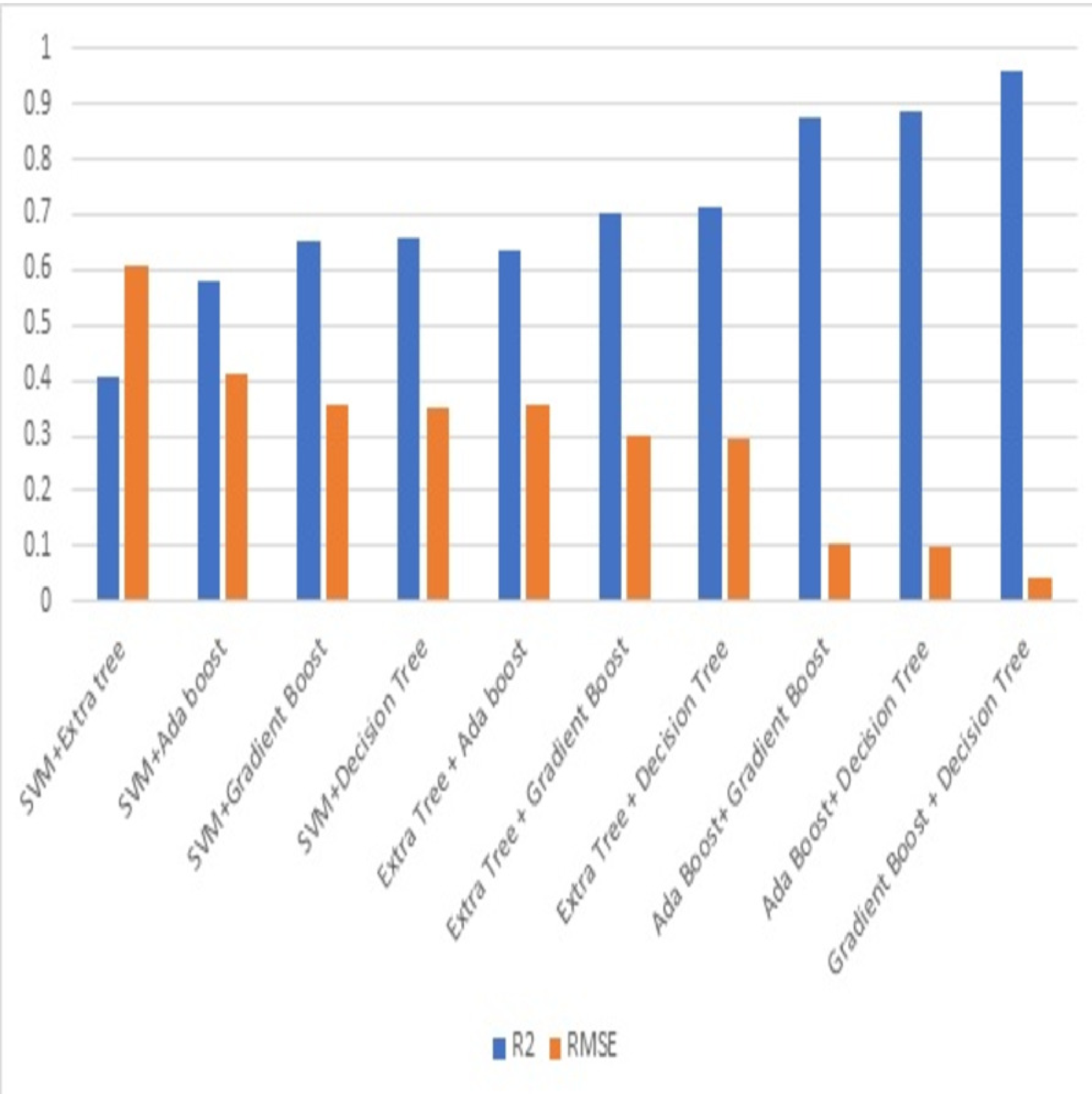

Figure 8

Ensembling Using 3 Methods

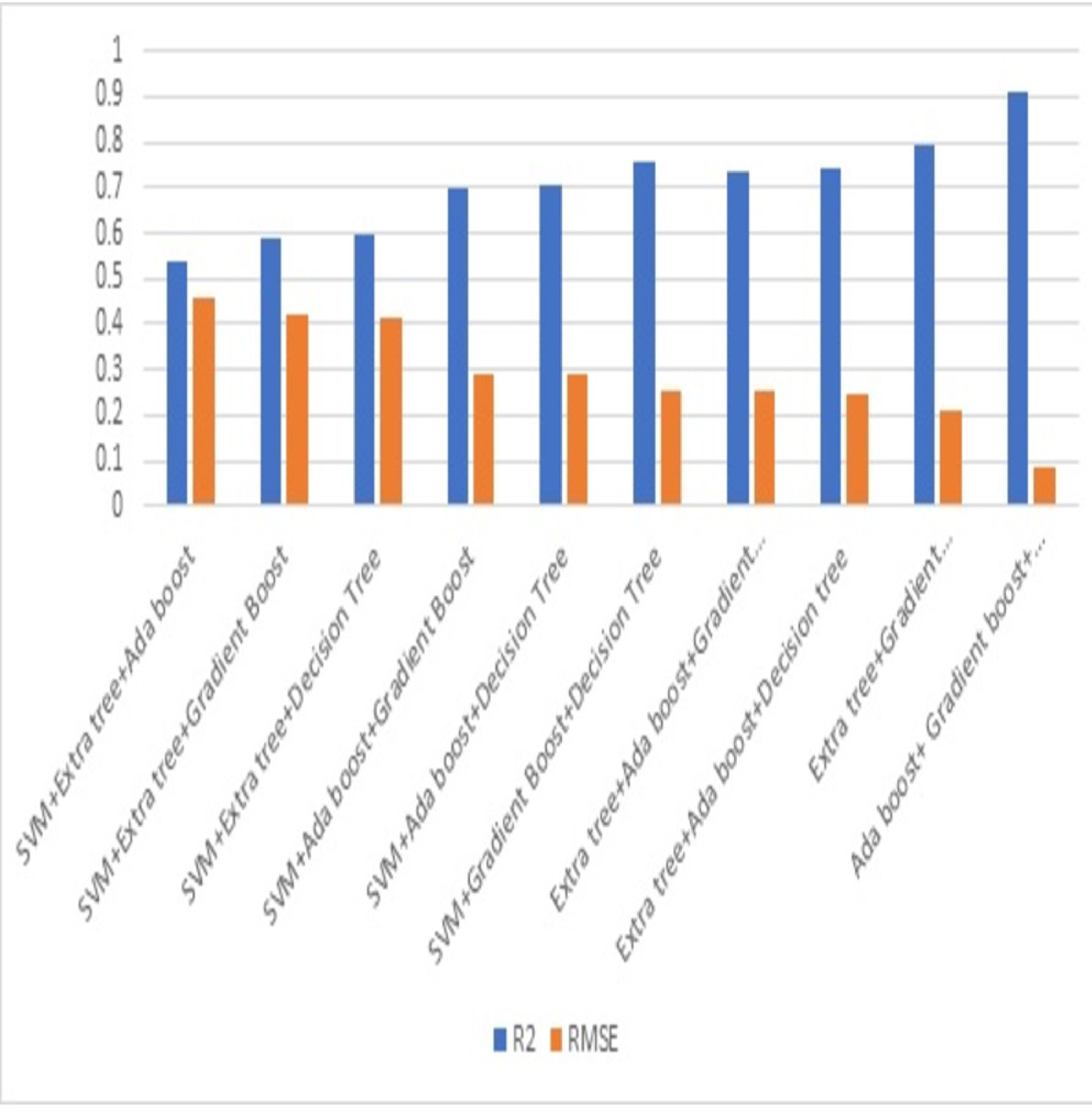

Figure 9

Ensemble Using 4 Methods

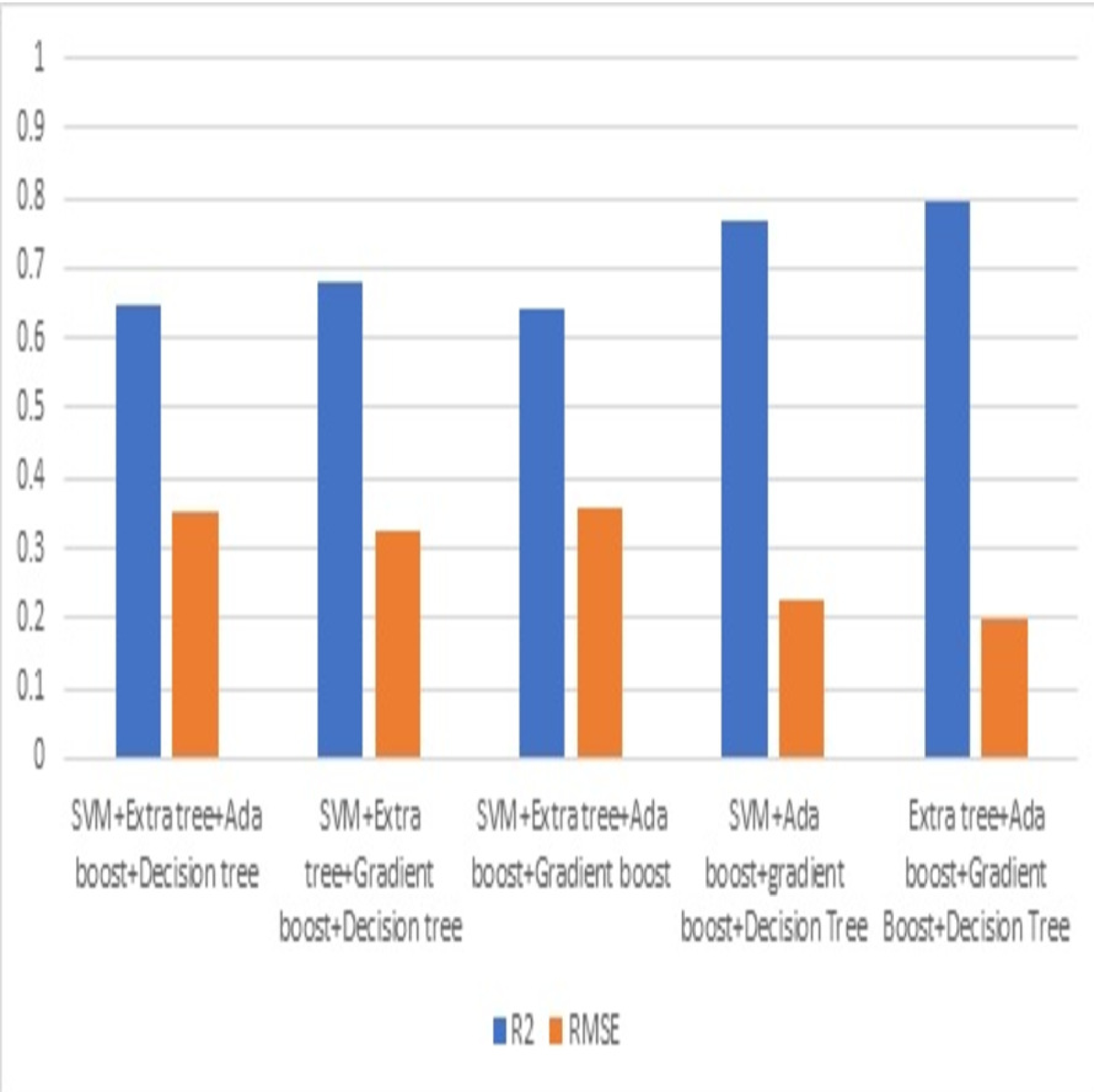

Figure 10

Comparison of R^2 value and RMSE value for old and new coupon

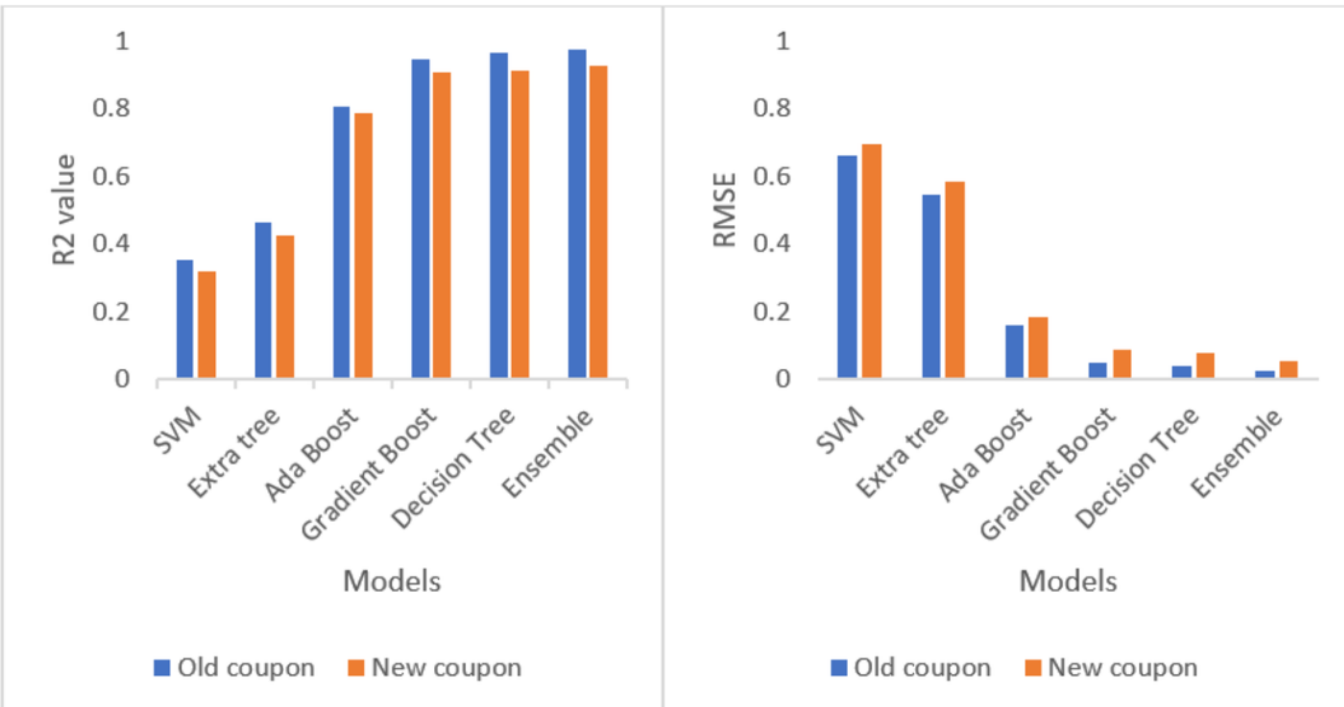

Table 1(on next page)

Model Evaluation

Model	R^2	RMSE
SVM	0.352	0.662
Extra tree	0.462	0.548
Ada Boost	0.806	0.159
Gradient Boost	0.948	0.050
Decision Tree	0.967	0.040
Ensemble	0.975	0.023

Table 2(on next page)

Model Evaluation for new coupon

Model	R^2	RMSE
SVM	0.319	0.694
Extra tree	0.423	0.586
Ada Boost	0.785	0.181
Gradient Boost	0.907	0.087
Decision Tree	0.913	0.078
Ensemble	0.928	0.053