Electrical Motor Temperature Prediction Using Ensemble Random Forest Regression Technique

KOLINBEN SUKHADIA
Electrical and Computer Engineering
Lakehead University
ThunderBay,Canada
ksukhadi@lakeheadu.ca

Abstract—To avoid increase in temperature, there should be accurate monitoring of temperature of stator winding. To predict better accuracy of permanent magnet synchronous motor, in this research, there will be ensemble regression technique is used. In this research, Ensemble Random Forest Regression technique is proposed for better accuracy of prediction and to reduce errors namely MAE(Mean Absolute Error), MSE(Mean Squared Error) and RMSE(Root Mean Square Error).

Keywords—Permanent magnet synchronous motor, machine learning technique, regression, Random Forest Regression technique, Ensemble method, Motor temperature prediction.

I. Introduction

Predicting motor temperature in permanent magnet synchronous motors is a challenging task for several decades. Due to high power, torque densities and high efficiency, permanent magnet synchronous motor is the best choice in various companies. There is not accurate temperature monitoring capabilities in most traction drive applications which ensures safe operation through expensive motor designs.Objective of this research is to predict the temperature of a motor with a permanent magnet using random forest regression technique. When the motor temperature is high, the results regarding motor performance gives warning. The input to algorithm contains one million data samples. Various parameters of datasets are u_q, coolant, stator_winding, u_d, stator_tooth, motor_speed, i_d, i_q, pm, stator_yoke, ambient, torque.

Random forest is ensemble technique. Random forest is made up of trees and more trees means more robust forest. It is an ensemble technique . It is better than a decision tree. bY averaging the result, this methodology reduces overfitting. Various evaluation metrics for motor temperature prediction are accuracy, mean absolute error, root mean squared error, relative absolute error. Various hyperparameters of random forest techniques are n_estimators, max_features, max_depth, min_samples_split, min_samples_leaf. N_estimators is number of trees in the forest, max_features parameter is maximum number of features considered for splitting a node, maximum number of levels in each decision tree depicts max_depth, min_samples_leaf parameter explains method for sampling data points. By tuning this parameter, it helps to predict better motor temperature. It helps to increase motor temperature accuracy and reduces errors such as mean absolute error, root mean square error and mean square error.

II. Literature Review

In this research, there are target variables such as torque, rotor temperature, and stator winding temperature. Authors applied random forest classification method. To improve accuracy in overall result, feature selection played a very important role.For rotor temperature, the ambient temperature, coolant temperature and stator yoke temperature were important parameters for prediction of electric motor temperature. Ambient temperature, coolant temperature, and stator yoke temperature parameter were also considered a important attributes for prediction of electric motor temperature. There was splitting of dataset in 80% of total data as train data and 20% of dataset as test data. Three algorithms were used namely support vector regressor, random forest regressor and polynomial regression technique.

In this research[2], For prediction of dynamic temperatures inside permanent magnet synchronous motor, linear regression with gradient descent and normal equations is used. Using regularization techniques namely L1 and L2 regularization, prediction results are improved. In this, there is comparision of linear regression with normal equation and linear regression with gradient descent.Mean square error for linear regression with gradient descent and linear regression with normal equation is equal. Authors describe that linear regression models are simple and it takes lesser processing time.

Dataset output parameters are pm, stator_yoke, stator_tooth, stator_winding temperature. Feature engineering is used to precisely predict motor temperature. Authors compared evaluation metrics for two different algorithms namely k-nearest neighbour and linear regression. Authors of this research depicts that k- nearest neighbour methodology performs better than linear regression methodology. There are two evaluation metrics such as RMSE(Root Mean Square Error) and MSE(Mean Squared Error). After experimental analysis, it is concluded that k-nearest neighbour algorithm performs much better than linear regression methodology.

In this paper[3], deep recurrent and convolutional neural network are evaluated to solve the issue of sequential learning of prediction high-dynamic temperatures. In order to assessthe consistency of model learning and topologies probabilistics search, bayesian optimization method is used.

In this[4], random foresat method is choosen for motor temperature prediction. Authors described various benifita and demerits of using random forsest algorithm.They described advantages such as ability to efficiently process data with numerous classes, high scalability of method, insensitivity to scaling. In this technique, discrete features and continuous features are treated equally. When there is lot of noise in dataset, this method relearn on some tasks. The proposed methodology used dataset collected from permanent motor synchronous motor placed on a test bench. Various parameters are ambient, coolant, voltage of d-component, Q-component voltage, motor speed, current of d-component, current of q-component and profile id, surface temperature of permanent magnet, stator_yoke temperature, stator_tooth temperature, stator_winding temperature. In this research, authors used python programming language with add-ins such as NumPy, Pandas, Scikit-learn, Matplotlib and seaborn.It achieves accuracy of 98.02%. They also measured accuracy on different PCs.

Random forest and feed forward neural network algorithms are used on electric motor temperature datasets[5]. Evaluation metrics after implementing technique on motor temperature datasets are MAE(Mean Absolute Error) and MSE(Mean square error). Mean square error is 0.007303 and Mean Absolute error is 0.066300 for implementing random forest algorithm on motor temperature datasets. For feed forwarding neural network, mean squared error and mean absolute errors are 0.1381 and 0.2959 respectively. The dataset consists of multiple measurement sessions which can be distinguised by profile_id column. So after applying both methods on electric motor temperature datasets, random forest methodology outperforms the feed forward neural network[5].

In this paper[6], authors described direct and indirect temperature estimation for motors. There is reduction in weight, volume and cost of the engine due to maximizing the degree of thermal use of permanent magnet synchronous motors so real time temperature data is required for this. There are indirect temperature estimation methods which detect changes in temperature-semnsitive parameters. For increasing accuracy and reliability, fusion methodology can be used.

In this paper[7], authors described to predict motor temperature using deep recurrent and convolutional neural networks. In order to assess the consistency of model learning and probabilistic search of topologies, the search for the model hyperparameter is sequentially performed using bayesian technique on different cores of the random number generator.

In this paper[8], with residual connections deep recurrent and convolutional neural networks are evaluated for prediction of high dynamic temperature inside permanent magnet synchronous motors. Neural network or black-box method is independent of motor sheet information. Reccurent neural network and convolutional neural network's function approximation properties helps in predicting high estimation accuracy.

The temperature of permanent magnet asynchronous motor(PMSM) is affected by heat loss and cooling modes of temperature rise[9]. To the safety and reliability of PMSMs, accurate prediction of stator winding temperature is necessary. DNN is extension of an artificial neural network. Neural network which consists of two or more hidden layers can be identified as a DNN. In this paper DNN has nine layers. The first layer is the input layer., the number of nodes is equal to number of input variables. The ninth layer is the output layer. The layers between input and output layer is identified as hidden layers. The layers from 2^{nd} to 8^{th} is considered as hidden layers. Each of them has 14 nodes. The technique proposed in this research was compared with DNN models with different numbers of hidden layer nodes, various activation functions, different learning rates. Mean Absolute Error and Root Mean Square Error were 0.1515 and 0.2368 respectively. R2 evaluation metrics was 0.9439 , which is closest to 1. DNN methodology prposed in this paper depicts better performance than other data mining methods.

In [10], authors depicts that lack of temperature estimations leads to high material cost. In this research, for prediction of high dynamic temperature, various machine learning methodologies are applied.

III.Material and Methodology :

Random forest is particularly for trees.It is similar to bagging except that each model is a random tree rather than a single model and each tree is grown according to bootstrap sample of training set to N. Forest is ensemble of several decision trees. There are various advantages of using random forest such as it is easy to build and faster to predict. There is resistance to over training and over-fitting of data. Whenever using random forest, it is able to handle data without preprocessing or rescaling. It is resistance to outliers and it can handle missing values. Random forest algorithm works as folllows. It is a combination of two sources of randomness, bagging and random input vectors.

Bagging means there is creation of ensembles by bootstrap aggregation.

Bootstrap sample contain 63.2% of data while others are replicates.

In the next step, a decision tree is grown to its greatest depth using bootstrap sample. It helps to minimize the loss function.

At each node, there is best split of decision tree is choosen from random sample of input variables instead of all variables. For each tree, there is calculation of misclassification rate, which is also known as out of bag error rate. In the last stage, there is aggregation of error from all trees to obtain overall out of beg error rate for classification.

Figure 1 depicts Random Forest Flow diagram

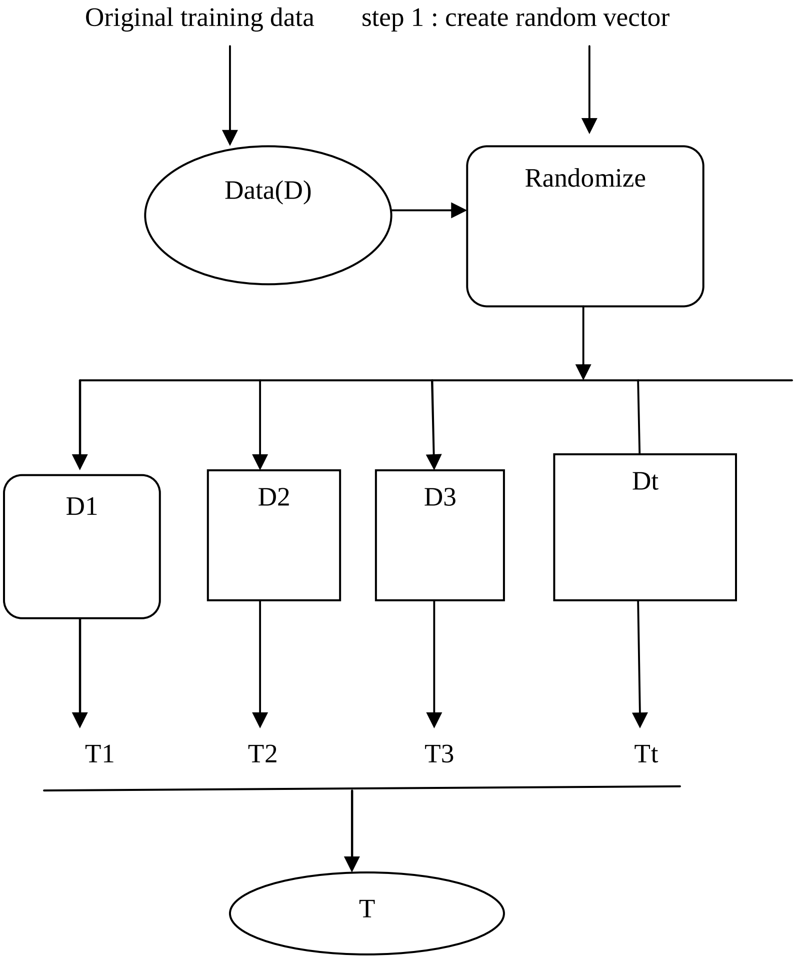


Figure1 : Random forest Flow Diagram

Step2 : use the random vector to build multiple decision trees.
Step3 : combine decision trees means there is average of single tree predictions

Proposed Methodology Flowchart :

Random forest builds a number K of regression trees making them grow from different training data subsets, there is resampling the original dataset with replacement. In different models, most data will be used multiple times.When random forest makes a tree grow, it uses the best predictor within a subset of predictors(m). This characteristic of random forest creates a greater prediction accuracy and stability. At same time, it also avoids the correlation of different regression trees. It also increases the diversity of patterns that can be learnt from data.

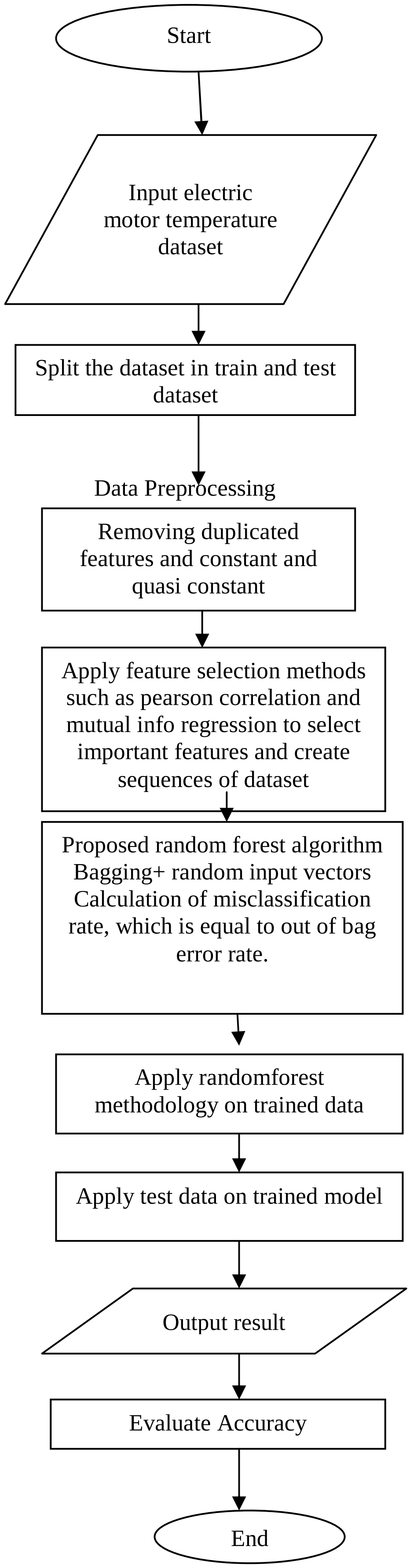


Figure 2 : Proposed methodology flowchart

In proposed methodology, after splitting dataset in train and test , there is removal of constant, quasi constant and removing duplicated features from dataset. Removing constant, quasi constant and removing duplicated features makes dataset better and this helps to predict electric motor temperature accurately.After that , there is creation of sequences of dataset and then apply proposed random forest algorithm on trained model. Evaluation metric is r2_score which measures prediction accuracy.

Dataset: In this research, dataset is divided based on profile id. If profile id<81 then those data of profile id goes to train dataset and profile id = 81 data is set as test dataset.

Feature Selection techniques

Pearson corelation method : In this, pearson correlation is number betweemn -1 and 1 that indicates the extent to which twop variables are linearly related. It is also known as the product moment correlation coefficient.

The correlation coefficient has values between -1 to 1. A value closer to 0 implies weaker correlation. A value closer to 1 indicates stronger positive correlation. Stronger negative correlation value is closer to -1.

While applying this feature selection technique on electric motor temperature dataset, there is accuracy of 0.99998070 so in percentage, it is 99.998070.

Mutual information technique for feature selection :

Mutual information is calculated between two variables and measures reduction in uncertainty for one variable . There is Known value of other variable.

While applying this feature selection technique on datset, there is accuracy around 97%.

IV.RESULTS AND DISCUSSION

Random Forest regression technique parameters are given as below:

n_estimators, criterion, max_depth, min_samples_split, min_samples_leaf, min_weight_fraction_leaf, max_features, max_leaf_nodes, min_impurity_decrease, bootstrap, oob_score, n_jobs, random_state, verbose, warm_state, ccp_alpha, and max_samples.

N_estimator is the number of trees in the forest. Criterion is function to measure the quality of split. Its values are squared error, absolute error and poisson. Its default value is squared error. Parameter max_depth determines maximum depth of the tree.Its default value is None. Parameter min_samples_split has default value 2. Its value is either in integer or float. Min_samples_leaf default vaule is 1. It is either in integer or float. another parameter min_weight_fraction_leaf default value is 0.0. max_features parameter has its values namely, auto, sqrt and log2. Max_leaf _nodes has default value None. In this parameter, Growing trees with `max_leaf_nodes` in best-first fashion. Best nodes are defined as relative reduction in impurity. If None then unlimited number of leaf nodes. Min_impurity decrease parameter indicates that node will be split if split induces a decrease of the impurity greater than or equal to this value.

Hyperparameter Tuning of Random Forest regression algorithm :

In this research, there is tuning of hyperparameters such as setting min_samples_leaf equal to 2. There is also tuning of parameter max_depth. Its value is 8. After tuning n_estimators value to 282 , this proposed methodology randomforest regression provides low error, mean squared error and good r2_score value.

parameter	value
Min_samples_leaf	2
Max_depth	8
n_estimators	282

Table 1: hyperparameter tuning of RandomForest regression algorithm to predict electric motor temperature

After hyperparameter tuning of random forest regrssion methodology, r2_score value is around 0.9725.

Feature Selection techniques

Pearson corelation method : In this, pearson correlation is number betweemn -1 and 1 that indicates the extent to which twop variables are linearly related. It is also known as the product moment correlation coefficient.

The correlation coefficient has values between -1 to 1. A value closer to 0 implies weaker correlation. A value closer to 1 indicates stronger positive correlation. Stronger negative correlation value is closer to -1.

Stator Winding temperature prediction

Table of Result and Discussion

Evaluation metrics	Random Forest without hyperparameter tuning and feature selection	Random Forest with hyperparameter tuning and feature selection	Decision tree regressio n
R2_score	0.8655	0.9725	0.6629
Mean Squared Error	51.2979	10.4791	81

Mean Absolute Error	5.3590	2.5459	9.3598

Table 2: evaluation metrics values between different methods while predicting stator_winding_temperature.

Mean Absolute Error	5.3590	0.9830	4.4390

Table 2: evaluation metrics values between different methods while predicting stator_tooth_temperature.

Stator_yoke_temperature prediction :

Evaluation metrics	Random Forest without hyperparameter tuning and feature selection	Random Forest with hyperparameter tuning and feature selection	Decision tree regressio n
R2_score	0.8655	0.9921	0.8445
Mean Squared Error	51.2979	2.2937	45.54
Mean Absolute Error	5.3590	1.0369	6.0238

Table 3: evaluation metrics values between different methods while predicting stator_yoke_temperature.

Evaluation metrics	Random Forest without hyperparameter tuning and feature selection	Random Forest with hyperparameter tuning and feature selection	Decision tree regressio n
R2_score	0.8655	0.9939	0.8884
Mean Squared Error	51.2979	1.8937	35.095

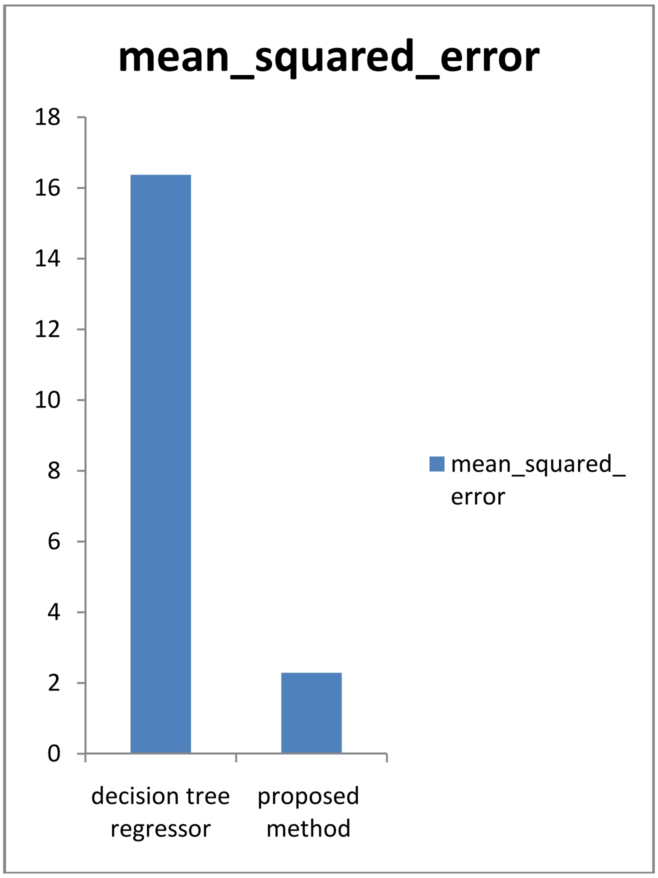


Figure 3 : Comparison of Mean Absolute Error between decision tree regressor and proposed Random forest method with feature selection and hyperparameter tuning while predicting stator_winding_temperature

	Decision Tree regressor	Proposed method Random Forest Regression
Mean Absolute Error	9.3598	2.5459

Table 3 : Comparison of Mean Absolute Error between decision tree regressor and proposed Random forest method with feature selection and hyperparameter tuning while predicting stator_winding temperature

Figure 4 : Comparison of R2_score between decision tree regressor and proposed Random forest method with feature selection and hyperparameter tuning while predicting stator_winding_temperature

	Decision Tree regressor	Proposed method Random Forest Regression
R2_score	0.662944	0.972524

Table 4 : Comparison of R2_score between decision tree regressor and proposed Random forest method with feature selection and hyperparameter tuning while predicting stator_winding_temperature

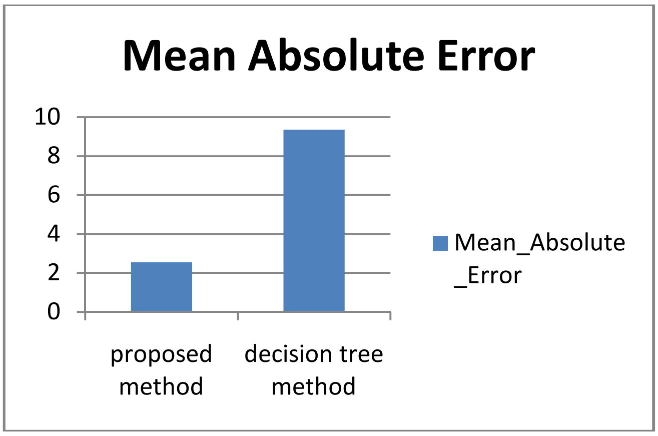


Figure 5: Comparison of Mean Squared Error between decision tree regressor and proposed Random forest method with feature selection and hyperparameter tuning

Stator_yoke_temperature prediction

Evaluation of various evaluation metrics namely, r2_score, Mean Absolute Error and Mean Squared Etrror.

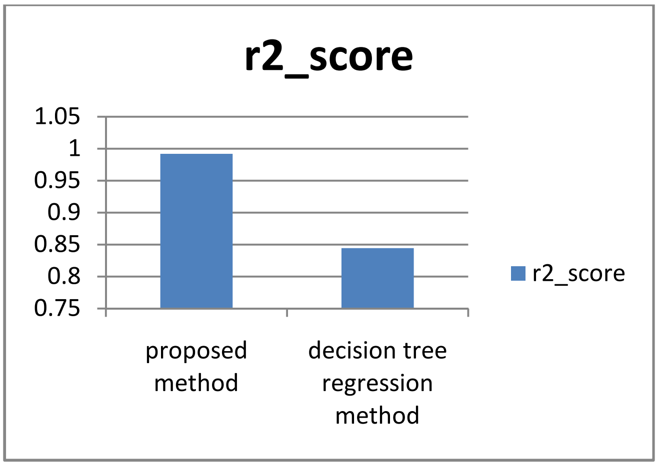


Figure 6: Comparison of r2_score between decision tree regressor and proposed Random forest method with feature selection and hyperparameter tuning while predicting stator yoke temperature

	Proposed method	Decision tree regression method
R2_SCORE	0.9921	0.8445

Table 5 : Comparison of r2_score between decision tree regessor and proposed methodology while predicting stator_yoke temperature

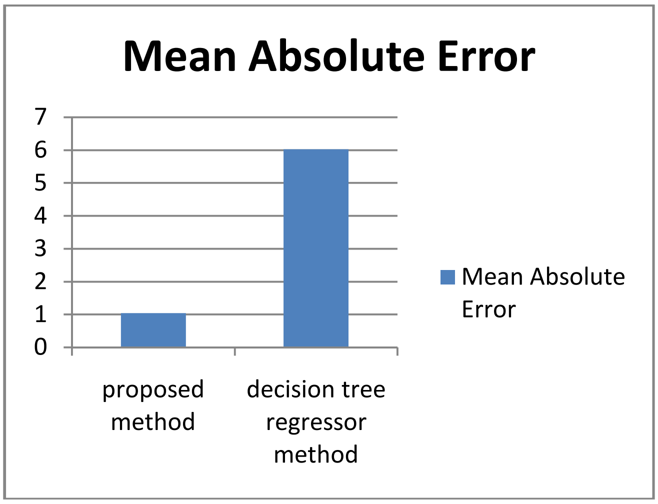


Figure 7: Comparison of Mean Absolute Error between decision tree regessor and proposed methodology while predicting stator_yoke temperature

	Proposed method	Decision tree regression method
Mean Absolute Error	1.03695	6.02386

Table : Comparison of Mean Absolute Error between decision tree regessor and proposed methodology while predicting stator_yoke temperature

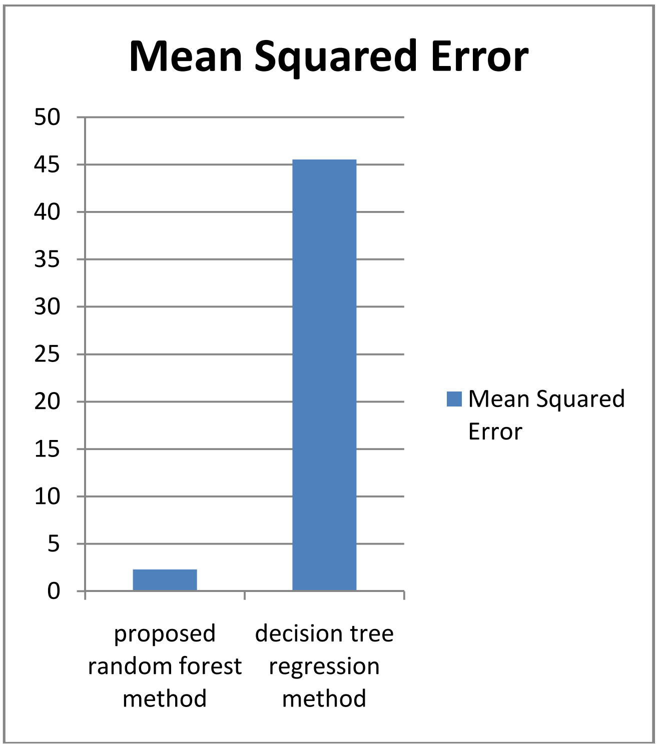


Figure 8: Comparison of Mean Squared Error between decision tree regressor and proposed Random forest regressor method with feature selection and hyperparameter tuning while predicting stator yoke temperature

Stator Tooth Temperature :

Comparison of R2_score, Mean Absolute Error and Mean Squared Error

	Proposed random forest regression method	Decision tree regressor method
R2_score	0.993981	0.8884
Mean Squared Error	1.893702	35.095
Mean Absolute Error	0.9830094	4.4390

Table : For stator_tooth_temperature prediction, Comparison of R2_score, Mean absolute error and mean squared error between proposed random forest regression and decision tree regression method.

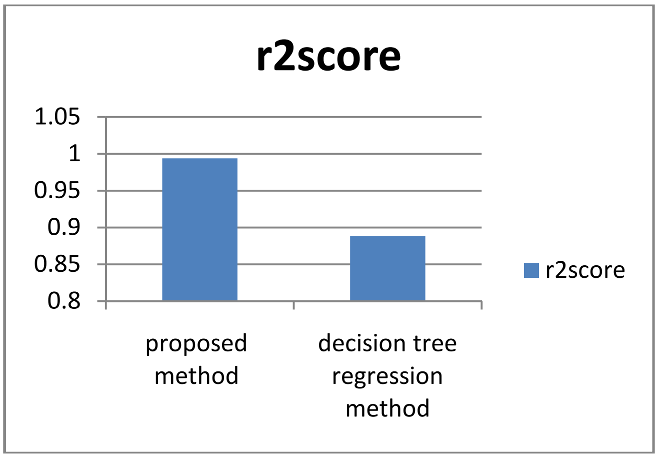


Figure 9: Comparison of r2score value between proposed random forest method and decision tree regression method while predicting stator tooth temperature

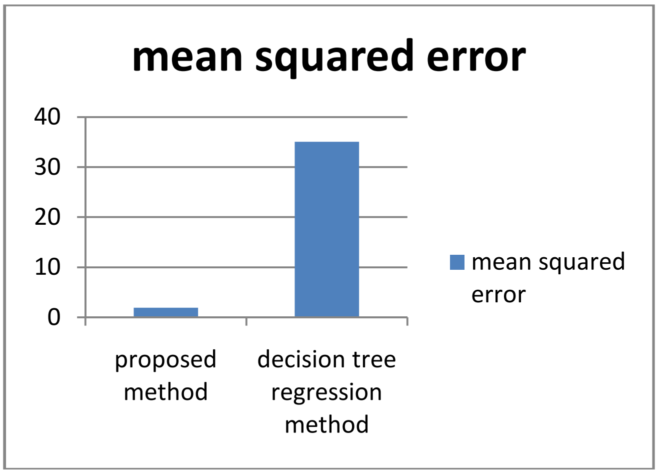


Figure : Comparison of mean squared error value between proposed random forest method and decision tree regression method while predicting stator tooth temperature

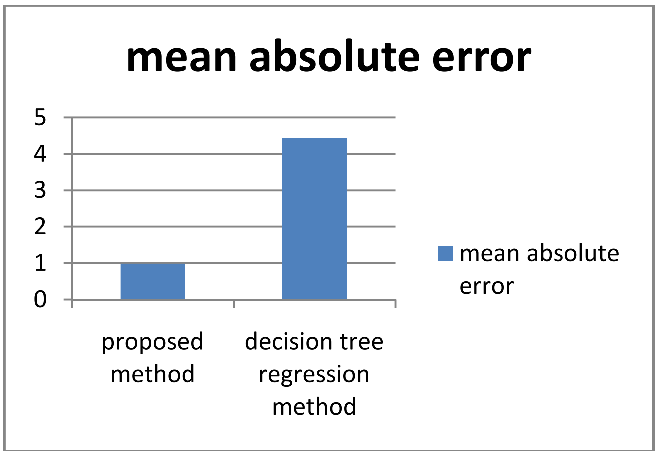
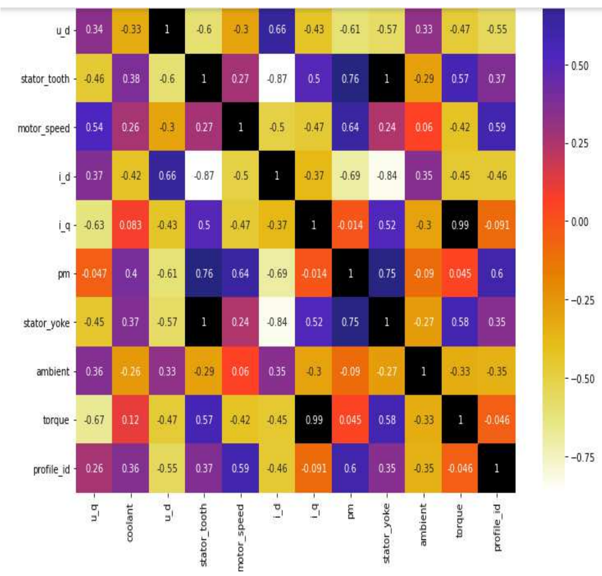


Figure : Comparison of mean absolute error value between proposed random forest method and decision tree regression method while predicting stator tooth temperature.

Pearson Corelation method for feature selection:



In this feature selection technique, there is a graph , which depicts corelation between features. This helps to select most influencing features and helps in best prediction of temperature.

Data Visualization:
For data visualization, feature selection technique is applied on electric motor temperature dataset. Pearson corelation feature selection technique is used for this. To visualize data, feature importance is necessary.

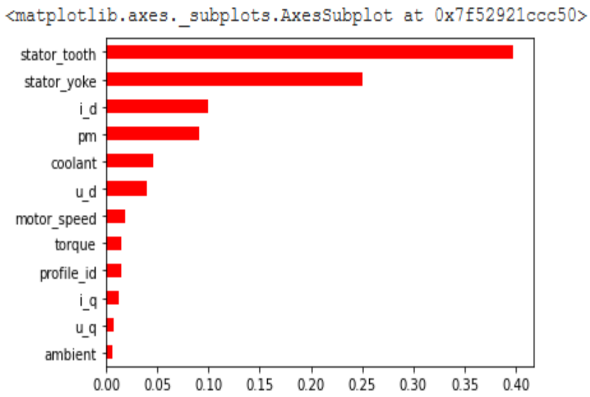


Figure 1: Feature importance of various columns of electric motor temperature dataset

Figure depicts values on x-axis and y axis determine

```
u_q             0.007244
coolant         0.046490
u_d             0.040335
stator_tooth    0.397159
motor_speed     0.018899
i_d             0.099838
i_q             0.012509
pm              0.090503
stator_yoke     0.250409
ambient         0.006582
torque          0.015178
profile_id      0.014854
dtype: float64
```

Figure : Column wise feature importance

Electric motor temperature has its own feature importance. For example, coolant parameter has feature importance value equal to 0.046490. u_d has 0.040335 feature importance value. It determines which features are important to predict electric motor temperature.

In this research, gridsearchcv method is applied on random forest parameter..

Evaluation Metrics :
These evaluation metrics depicts whether regression models are accurate or misleading.
R2_score : R2 score is known as coefficient of determination. R2_score varies between 0 and 100%. It is a statistical measure of fit that indicates how much variation of dependent variable is explained by bthe independent bvariable.

Mean Squared Errror : Mean Squared error is the average of the square of the errors.The larger the number , the larger the error. Error means the difference between the observed values and predicted values.

Root Mean Square Error : It is computed from mean squared error. RMSE is square root of value obtained from mean square error function.

In this research , there is implementatiomn of Recurrent neural network, Lstm based network. In this research, I generate data sequences of dataset and then fit the model of lstm based recurrent neural network on sequentisal dataset. I implement this model to compare its prediction error and accuracy rate with my proposed methodology of random forest regression method with feature selection and hyperparameter tuning. Moreover, in this research , I also applied gridsearchcv method for parameter tuning of random forest algorithm.

So in this research, there is good r2_score value and low mean squared value after doing hyperparameter tuning of randomforest algorithm . In this research, pearson corelation feature selection technique is applied. So there is mean squared error value of 0.05; however existing method has value of 0.0607. Error is decreased by applying proposed method on dataset to predict electric motor temperature. In this research, stator winding temperature of motor is predicted. It is taken as target variable.

V.CONCLUSION

To predict better motor temperature, Random Forest technique is used for better prediction. By tuning parameters of random forest regression technique, there is better accuracy of motor temperature prediction. Low error is achieved by setting N_estimators value to 150, Min_samples_split to 6, Max_features equal to Sqrt and Bootstrap value to False.Tuning parameters help to predict better stator winding temperature. There is also prediction of stator_tooth_temperature, stator_yoke_temperature of electric motor. By applying Proposed random forest regression methodology on dataset, it achieves r2score value 0.9939, 0.9921 and 0.9725 while predicting electric motor stator_tooth, stator_yoke and stator_winding temperature respectively. It achieves mean absolute error values 0.9830, 1.03695 and 2.5459 while predicting electric motor stator_tooth, stator_yoke and stator_winding temperature respectively. After that, there is comparison of proposed random forest regression methodology with decision tree regression method. By comparing, Mean absolute error, mean_squared_error values of proposed regression method with decision tree regression method, it is concluded that, proposed methodology with hyperparameter tuning and feature selection method achieves low errors than decion tree regression method. Moreover, There is also comparison of evaluation metrics of proposed random forest regression method containing hyperparameter tuning and feature selection with proposed methodology without hyperparameter tuning and feature selection. There are various evaluation metrics namely, R2Score, mean_absolute_error, mean_squared _error. So in this research, this proposed methodology of random forest regression is the best to estimate various temperature of elecric motor namely, stator_winding, stator_tooth and stator_yoke temperature.

Limitations of Random Forest : The range of predictions a Random Forest can make is bound by highest and lowest labels in training data. This is a regression problem.When the training and prediction inputs differ in distributions and range, this make more problamatic. Additionally, it takes a more time for fitting on train dataset. So it takes more time for execution. Additionally, Depending on the trend, when training data is missing timre periods, Random Forest under or over predict.

Future Direction : To get better electric motor temperature accuracy and low errors, there is hyperparameter tuning of proposed method random forest will done. There are various hyperparameters namely n_estimators, criterion, max_depth, min_samples_leaf, min_weight_fraction_leaf, max_features, max_leagf_nodes, bootstrap, oob_score, n_jobs, verbose, warm_start, ccp_alpha and max_samples. There will also data visualization techniques applied for finding more influencing attributes on result prediction and use only those more influencing attribuites for electric motor temperature prediction.

VI.REFERENCES

1. Rushank Savant, Abhiram Ajith Kumar, Aditya Ghatak, “Prediction and Analysis of Permanent Magnet Synchronous Motor parameters using Machine Learning Algorithms”, International Conference on Advances in Electronics, Computers and Communications, 2020.
2. Poorva Thosar, Jyotika Patil, Mishail Singh, Swaraj Thamke, Sudhansu Gonge, “Prediction of Motor Temperature using Linear Regression”, IEEE, 2020, PP.7-12.
3. Ran Le, Kaijun He, Ace Hu, “Motor Temperature Prediction with K-Nearest Neighbours and Convolutional Neural Network”, pp.1-7.
4. Oleh Soprun, Myroslava Bublyk, Yurii Matseliukh, Vasyl Andrunyk, Lyubomyr Chyrun, Ivan Dyyak, “Forecasting Temperatures of a Synchronous Motor with Permanent Magnets Using Machine Learning”.
5. Aanchan Narayanan, ”Prediction of stator Winding Temperature of PMSM Motor Using Random Forest And Feed Forward Neural Network”, International Research Journal of Engineering and Technology,May 2021, pp. 623-625.
6. Wallscheid, O. Bocker, J.,”Fusion of direct and indirect temperature estimation techniques for permanent magnet synchronous motors”,IEEE International Electric Machines and Drives Conference(IEMDC), 2017, pp. 1-8.

7. Wilhelm Kirchgassner, Oliver Wallscheid, Joachim Bocker, “Deep Residual Convolutional and Recurrent Neural Networks for Temperature Estimation in Permanent Magnet Synchronous Motors”,IEEE, 2019
8. Wilhelm Kirchgassner, Oliver Wallscheid, “Estimating Electric Motor Temperatures With Deep Residual Machine Learning”,IEEE transactions on power electronics,July 2021,pp.7480-7488.
9. Hai Guo, Qun Ding, Yifan Song, Haoran Tang, Likun Wang, Jingying Zhao, ”Predicting Temperature of Permanent Magnet Synchronous Motor Based on Deep Neural Network”, mdpi journal, 30 July 2020, pp. 1-14.
10. Wilhelm Kirchgassner, Oliver Wallscheid, ”Data-Driven Permanent Magnet Temperature Estimation in Synchronous Motors with Supervised Machine Learning”,IEEE, 17 Jan 2020, pp.1-9.