Predicting Wildfires Using Machine Learning Models

Anika Tahsin
School of Data and Sciences
Brac University
Dhaka, Bangladesh
anika.tahsin5@g.bracu.ac.bd

***Index Terms*—Machine Learning, Regression, Wildfire Prediction**

I. Introduction

Scholars and forest managers have been intrigued for many years by the difficulty of predicting the incidence of wildfires. Attempts to understand the conditions that lead to the ignition and spread of wildfires can be dated at least 150 years back [1] and [2].

If a wildfire continues to spread uncontrollably due to strong winds, and starts affecting nearby fields or farms, we might face unavoidable famines due to the destruction of crops and other food resources. Loss of property, and food resources also direct to economical mayhem which may lead to instability in local labor markets. Wildfires have a very negative impact on our climate and weather as it releases large quantities of greenhouse gasses such as carbon dioxide, carbon monoxide, and fine particulate matter into the atmosphere. Due to climate change, forest fires are likely to become more frequent, and more severe, and cause billions of dollars in damage yearly. For the prevention of disasters and environmental protection, wildfire forecasting is vital.

In recent years, Artificial Intelligence (AI) models have proven to be very effective for predicting natural hazards [3], [4]. The models applied to predict forest fires just forecast the parameter of the area impacted by fires because the study of forest fire prediction has mostly been concentrated on the prediction of forest fire frequency [5], [6].

In this wildfire prediction study, the data were quantitative, numerical and categorical data to predict forest wildres. I implemented a Regression model and Neural Network model to predict the scale of wildfires. The main objective of this study is to investigate the capability of regression (i.e., Linear, Ridge, Lasso, Decision Tree, Random Forest, XGBoost) and a neural network model to predict the scale of wildfires in forests using quantitative data.

II. Related Words

A rise in temperature and an increase in the duration and severity of droughts are recent climate changes that contribute to the occurrence of bigger wildfires. Since the 1990s, artificial intelligence has been used in wildfire science and management; early examples include neural networks and expert systems. Since then, the field has advanced quickly in line with the environmental sciences' widespread use of machine learning (ML) and Deep Learning (DL) techniques. The research in this area is briefly summarized in the following sections.

David Chaparro et al. [7] stated that roughly 3% of fires were removed from the Database because moisture and/or temperature data on the day of the incident to three days in the past were not available. After eliminating all of these features, they were left with a forest fires dataset from the European Forest Fires Information System (EFFIS). Once the data were prepared, moisture and temperature conditions for both burned and unburned cells were plotted using a decimal logarithm, which showed that both variables limit the extent of the fire. They tested out Pearson correlation coefficients which gave $r = -0.54$, indicating no strong redundancy between moisture and temperature so the two variables were combined in a single model. And they applied the Linear model for performing the regression analysis to predict potential extension of wildfires.

Chaparro et al. analyzed the variance of the model, its coefficients and terms in their paper. The equations Chaparro et al. used were validated and the results were identical to those from the main model. They [7] mentioned that the use of a longer fire record would be necessary to neutralize the shifting and to extend the final model. Even with limitations, the model suggested in the research described 68% of the potential burned area, a considerable portion of the variance as a function of moisture-temperature circumstances. A maximum inaccuracy of 40.5 ha and 83.3% agreement were found during validation. A potential burned area's fluctuation was 33% explained by soil moisture. I think the implementation of their model is feasible as the results showed how the suggested methodology provides a good approximation to the maximum expected fire spread. The results demonstrated that the suggested methodology provides a good approximation to the maximum projected fire spread, so I believe that their model can be implemented and could help to anticipate forest fires extension in the Iberian Peninsula.

The performance of the proposed method by Can Lai et

al. [8] is examined using data on forest fires gathered from Portugal's Montesinho Natural Park. Can Lai et al. mentioned that more than half of the sample outputs of the dataset were zeros which they considered as seriously imbalanced when visualized through a histogram. They used Sparse Autoencoder aimed towards extracting useful features from the K-dimensional input of N training samples. Deep Neural Networks (DNN) that aimed at constructing an empirical mapping function from a K-dimensional input space was used. In order to predict forest fires using numerical environmental factors, a new strategy based on a sparse autoencoder-based DNN and data balancing procedure was developed in the paper by Can Lai et al. I think with a larger amount of data DNN approach would be better and further improved, knowing the capability of DNN to extract useful information from a large amount of data.

Shruti Lall and Bonolo Mathibela [9] proposed and developed a novel data driven intelligent system utilizing artificial neural networks (ANN) for wildfire risk assessment for the City of Cape Town. They first used the Neural Network (NN) technique to suggest that it would be useless to use NN to categorize the fire hazard level for the Renosterveld. Therefore they developed fynbos and strandveld vegetation. The data was balanced and normalized to be as input for the Artificial Neural Network with the developed vegetations. Accuracy, Precision, Recall, and F-score were the four performance criteria they examined, and the overall F-score rating was good at 0.981. The overall system provides satisfactory accuracy and precision of 0.97 and 0.87. The created system, which uses artificial neural networks, can accurately categorize the risk of fire ignition under diverse environmental situations, I believe.
Haijun Zhang et al. [10] experiments and stated that the most effective explanation for why fires occur is provided by NDVI. They applied the Binary Logistic Regression Model to describe the relationship of several independent variables. The AUC value of the model proposed by Haijun Zhang et al. is 0.842 (84.2%), which implies a good correlation between independent variables and dependent variables. I believe the generated fire probability maps can help fire managers find spatial potential fire danger areas so that they can respond appropriately to situations in Northeast China's fire protection activities.

David R. et al. [11] statistically evaluated FireCast for total prediction accuracy, total burn accuracy, and F-score in their paper. They applied accuracy, recall, and f-score prediction. They implied machine learning within FireCast allows the system to learn important correlations. FireCast is quicker, less complicated, and simpler to construct than Farsite because of the smaller quantity of necessary and variable data, as well as the lack of a requirement to generate both landscape and canopy cover files. With terms to overall prediction accuracy, recall, and F-score, FireCast performs better than Farsite. They thought about the limitation of FireCast which is the availability of appropriate training data. Due to access limitations, they disregarded the impact that firefighter action might have on the training data while doing research for their work.

Ashima Malik et al. [12] used six models which gave the following accuracies; Random Forest: 92, AdaBoost: 91.5, Gradient Boosting Tree: 90.5, Weighted Decision Tree: 89.1, Mlp: 86.1, LSTM: 91.6. Then they continued using an ensemble random forest model and combined model for their research based on the accuracy. The best threshold for ensemble on the ROC curve is 0.499112, and the G-Means is 0.866. Whereas, the classification report for the combined model shows balanced results with 0.95 precision and 0.92 f1-score with 1695 support for positive and 0.89 precision and 0.92 f1-score with 1694 support for negative test data. The combined model performed better in terms of accuracy when compared to the ensemble model. However, drawbacks were observed, the first one being that our machine learning model depends heavily on the data availability for the terrain.

Assaf Shmuel et al. [13] applied multiple Machine Learning models which were random forest, XGBoost, multilayer perceptron, and logistic and linear regressions. Once trained, the best models offered promising predictive accuracy and could foresee the occurrence of wildfires with over 90% accuracy and the size of the burned area with an MAE score of 3.13 km2. Results were most accurately produced by the XGBoost model, then by the random forest model. In all of the regions that were studied, the XGBoost classification model performed better than the logistic regression at predicting the likelihood of wildfires.

Sibo Cheng et al. [14] have developed a scheme which combines different reduced-order modeling and data-driven prediction models for efficient parameter-flexible burned area forecasting. They used Reduced-Order Modeling which were Principal Component Analysis, Convolutional Autoencoding, Singular Value Decomposition autoencoding. Also, Machine Learning predictions as Random Forest Regression, K-Nearest Neighbours (KNN) Regression, Multi Layer Perceptron. Moreover, Inverse Problems as Four Dimensional Variational Approach, Generalised Latent Assimilation. When comparing the outcomes, SVD AE outperforms PCA in terms of both online effectiveness and reconstruction/prediction accuracy. Therefore, for parameter estimation, Sibo Cheng et al. simply paid attention to PCA and CAE. Sibo Cheng et al. mentioned that their model would face difficulties in predicting real-time satellite images even though the reduced-order modeling and the forward prediction methods achieved a precise reconstruction on the test dataset.

Aman Preet Singh Gulati, in his repository, [15] analyzed data from UCI Machine Learning Repository to predict the credence of the forest fire based on some attributes. While

data cleaning, they used a scan method on the data to find the irrelevant features which are not valuable enough to affect the target feature. Afterwards, they used the scanned columns to make them categorical by using a binning method. The author tested their model with the Random Forest Regressor model. However the model did not do well in accuracy including the model was overfitting. So the author furthermore tuned the model using RandomizedSearchCV from scikit learn model selection. However the test set accuracy was improved by around 2%.

Ashishkumar Rana [16] studied the dataset from UCI Machine Learning Repository. The problem the author chose was the dataset on Algerian Forest Fires. This dataset contains a culmination of forest fire observations and data in two regions of Algeria: the Bejaia region and the Sidi Bel-Abbes region. The goal of Ashishkumar's [16] project was to identify the certain weather features that could predict forest fires in these regions using few Classification algorithms. For checking the multicollinearity the author uses the Variance Inflation Factor (VIF) technique. Greater the value of R-squared, greater is the VIF. Hence, greater VIF denotes greater correlation. They used Standard Scaler to remove the mean and scale the data to the unit variance before Using Regression model and Classification model. For the Regression model, Random Forest Regressor gave the best accuracy of 97.98%, and for the Classification model XGboost Model for classification gave the best accuracy of 97.26%.

Philip Sohn [17] used data from UCI Machine Learning Repository to predict the credence of the forest fire based on some attributes. As no feature seems to have ranges outside of the ordinary they did some data cleaning by looking into date time features. For modeling they used ANN, Linear Regression, SVM, Random Forest. But no models were able to give a good prediction. Their ANN model was positive but had an r^2 value of 0.0122 which is extremely low and not useful for any analysis. In order to improve they mentioned to have more parameter optimization, try other models like xgboost. I think the absence of some external data is what caused their models to be so weak.
The forest fire has become a threat not only to the forest wealth but also flora and fauna and ecology of the environment. All the authors more or less aimed to not just determine if a forest fire will take place or not, they predicted the confidence of the forest fire based on some attributes in different cases and areas of forest fire.

III. Software

Data collection and preprocessing, and the establishment of the regression and neural network models, were developed and implemented in the TensorFlow framework of the Anaconda 3 software (Anaconda, Inc.).

IV. Abbreviation

Names	Abbreviation
Machine Learning	ML
Neural Network	NN
Artificial Neural Network	ANN
Mean Square Error	MSE
Root Mean Square Error	RMSE
Ordinary Least Squares	OLS
Decision Tree	DT
Support Vector Machine	SVM
Variance Inflation Factor	VIF
Cross-validation	CV
K-Nearest Neighbours	KNN
Exploratory Data Analysis	EDA

V. Research Methodology

A. Dataset Description

The data was collected from the United States Department of Agriculture Forest Service (USDA) [18], [19].

The data consists the Monitoring Trends in Burn Severity (MTBS) program which assesses the frequency, extent, and magnitude (size and severity) of all large wildland fires (including wildfires and prescribed fires) in the conterminous United States (CONUS), Alaska, Hawaii, and Puerto Rico from the beginning of the Landsat Thematic Mapper archive to the present. All fires reported as greater than 1,000 acres in the western U.S. and greater than 500 acres in the eastern U.S. are mapped across all ownerships.

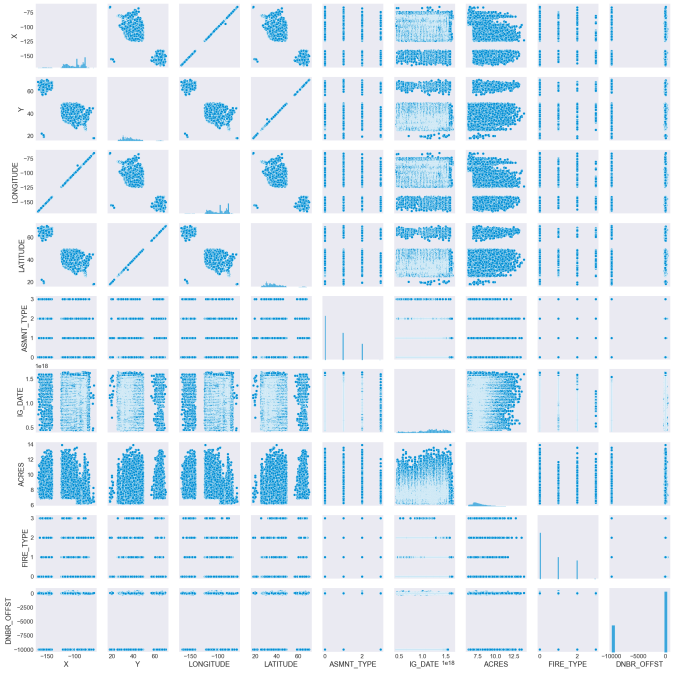

Fig. 1. Pairplot For Dataset

B. Data Filtering and Handling

A total of 29,533 fire data with 25 feature records were obtained from the year 1984 to 2021. The data is arranged by

Acres, Threshold, Latitude, Longitude, Fire Type, Coordinates. It also consists of comments and ID series of Object, Fire, Area, Map. Since the ID features will not affect the target feature, these were dropped using numpy.
The data also had several missing values along with some redundant features for Wildfire prediction. To handle missing data and Na/NaN values of the features, Pandas was used to deal with numeric missing values. For the categorical features, there were 4 unique values which were encoded into indicator variables using Pandas. Exploratory Data Analysis (EDA) was used to analyze data, correlation of the target feature was shown with correlation matrix using heatmap.
The ID features were dropped and stored for future prediction.

C. Finding the Target Variable

To start with, I selected the ACRES feature for prediction at random to run the models, but this gave a huge root mean squared error. Since RMSE is better near zero, I tested out Pearson and Spearman correlations to determine what feature to choose as the target variable.

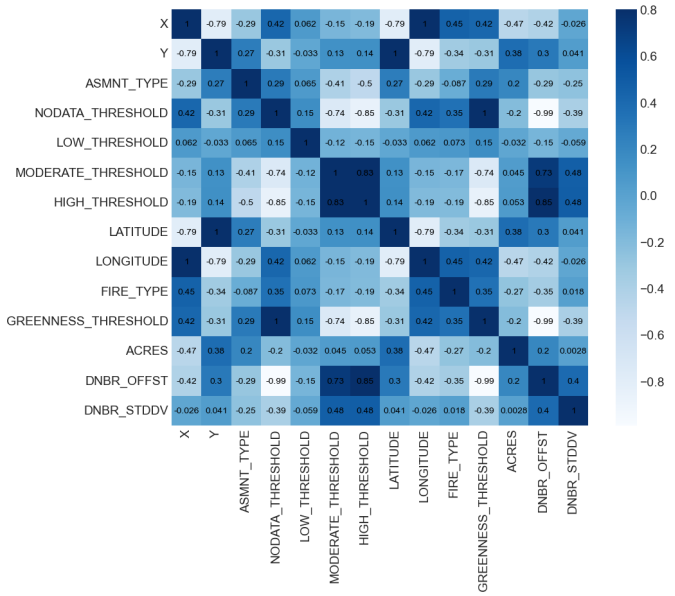

Fig. 2. Correlation Matrix

From the correlation matrix, there were 6 pairs of features that were either 1 or were very close to 1. The pair of features and their respective values were,

- X, LONGITUDE = 1
- Y, LATITUDE = 1
- GREENNESS_THRESHOLD, NODATA_THRESHOLD = 1
- HIGH_THRESHHOLD, DNBR_OFFST = 0.85
- HIGH_THRESHHOLD, MODERATE_THRESHOLD = 0.83
- MODERATE_THRESHOLD, DNBR_OFFST = 0.73

After testing the Pearson and Spearman correlation, the following results were found,

Test	X, Y	Nodata with High	Moderate with DNBR	Moderate with High	High with DNBR
Pearson	-0.792	-0.85	0.73	0.83	0.84
Spearman	-0.707	-0.83	0.642	0.88	0.66

Since, in each pair the Pearson test had better outcomes than Spearman so I leaned towards the Pearson results. As for DNBR_OFFST the Pearson test result were better than the rest, so I chose the feature DNBR_OFFST to be target. This feature did give a better RMSE than ACRES but it still needed some modification.
So I went back to the feature 'ACRES' again. But this time, before multicollinearity test, I standardized the data and plotted it to find the value density to be between 0 to 1. Finally, standardization helped in finding the target feature, 'ACRES'.

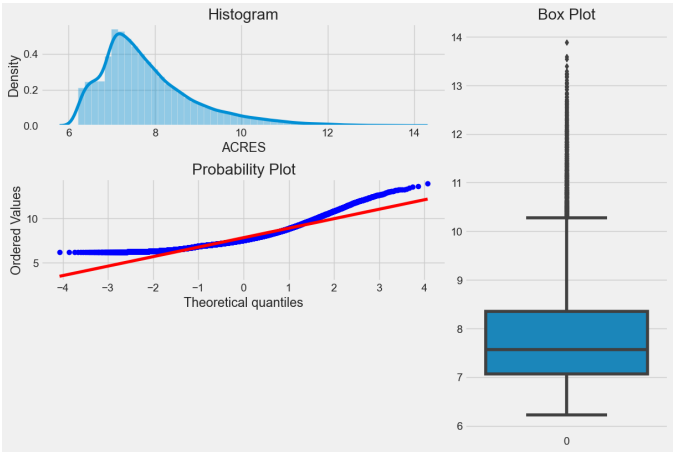

Fig. 3. Standardized Target feature 'ACRES' data

D. Multicollinearity Test

Strong correlations between the explanatory variables in a regression model are referred to as multicollinearity, and they can lead to departures from the actual data and distort the model's estimates [20].
The accuracy of the wildfire scale prediction model should not be adversely affected by factors with significant collinearity. So, I applied a correlation matrix and selected those independent variables with high correlation with dependent variables. I set the high correlation to be >0.1.

E. Models

In this study, two types of models were employed to evaluate the likelihood of wildfire occurrence. The first group of ML models were the regression models which predict the ACRES for wildfire. The objective function for wildfire scale prediction modeling was the root mean square error (RMSE) measuring the magnitude of the error between the observations and predictions.

Application of Regression Models

1) Linear Regression: To start with, I used Linear Regression on our preprocessed data, which makes the assumption that the predictors and target variable have a linear relationship, here 'ACRES'. The linear regression can be expressed in the following form:

$$Y \approx \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

Fig. 4. Linear Regression Equation

Where a = coefficients, x = features, b is the parameter of the model. The model's parameters a and b are chosen using the Ordinary Least Squares (OLS) method. It operates by reducing the sum of residuals' squares (actual value - predicted value).

2) Ridge Regression: An approach to regularization is called Ridge Regression. Ridge regression is a linear regression extension in which the loss function is changed to reduce the model's complexity. The modification is accomplished by including a penalty parameter that is equal to the square of the coefficients' magnitude. A high alpha can result in underfitting, so I set the parameter alpha with a value of 0.001 for the loss function computation.

$$\sum_{i=1}^{M} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{M} \left(y_i - \sum_{j=0}^{p} w_j \times x_{ij} \right)^2$$

Fig. 5. Ridge Regression Equation

3) Lasso Regression: Lasso Regression is a regularization technique. Another adaptation of the linear regression is lasso regression, commonly known as the Least Absolute Shrinkage and Selection Operator. By limiting the sum of the absolute values of the model coefficients, Lasso modifies the loss function to reduce the model's complexity. In this case, the parameter alpha is the same as used in the Ridge regression model.

Hyper Parameter Tuning: The ridge, lasso models are tuned by various hyperparameters. I performed hyperparameter optimization for both models to achieve optimal predictions. The Grid Search Cross-validation method present in the sklearn library was used to find the best parameters of ridge, lasso. The following hyperparameters were examined: tuned alpha, cv = 10, n_jobs = -1, verbose = 1.

The accuracies presented in the Results section are of the optimal hyperparameters.

4) Decision Tree Regression: Decision Tree [21] is a Supervised learning technique that can be used for both classification and Regression problems. The decision node is where the tree divides into different branches, with each branch denoting the specific decision the algorithm is making and leaf nodes denoting the results of the model. The objective of DTs is to accurately capture the relationships between input and outputs using the smallest possible tree that avoids overfitting. In a decision tree, the algorithm starts with X_train and y_train as parameters with a cross validation of 5. This is done to predict the class of the given dataset.

5) Random Forest Regression: Random Forest [22], a well-known machine learning algorithm that falls under the category of supervised learning, is an ensemble model made up of several individually trained Decision Trees (DTs). This approach achieves excellent performance by minimizing correlation across trees and reducing model variance, resulting in a huge number of diverse trees that are more accurate than individual trees.

6) XGBoost Regression: XGBoost [23] is a powerful ensemble learning approach that can be used directly for building supervised regression models. By knowing about its objective function that contains loss function and a regularization term, and also base learners, the validity of this statement can be inferred. It tells about the model results from the real values, which is mainly the difference between actual and predicted values.

$$\tilde{\mathcal{L}}^{(t)} = \sum_{i=1}^{n} [g_i f_t(\mathbf{x}_i) + \frac{1}{2} h_i f_t^2(\mathbf{x}_i)] + \Omega(f_t)$$

Fig. 6. XGBoost Simplified Objective

The second model I employed was the Artificial Neural Network (ANN).

Application of Neural Network Model

Artificial Neural Networks are functions that attempt to mimic the way a human brain makes decisions. They are particularly useful when dealing with complex non-linear classification problems [24].
In this study, the ANN architecture used was that of the Multilayer Perceptron. In the algorithm, a hidden layer node j gets the values of a group of independent variables (x_i to x_p) from an input layer. Then each value is multiplied by a weight w_h and then added together to produce a value u_j. This value is transferred through a non-linear sigmoid transfer function, f(x), to create the value h_j, which is then subjected to additional weighting before being sent to the output layer. Finally, the weighted v_j values are added together and passed into another sigmoid transfer function f(x), which outputs the final y_k values of the model [25].

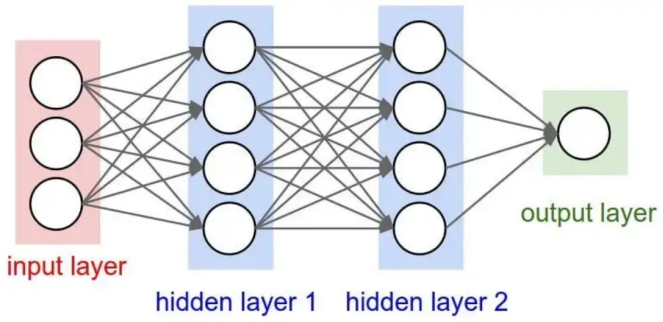

Fig. 7. Artificial Neural Network Layers

The ANN model was also tuned by various hyperparameters to optimize the model to achieve optimal predictions. Grid Search Cross-validation method present in the sklearn library to find the best parameters of ANN. The following hyperparameters were examined: a regression ANN model as estimator, batch size = [10,20,30], epochs between 10 and 20, Optimizer_trial as adam and rmsprop, cross validation = 5, and a customized scoring with greater accuracy score.

F. Model Accuracy

I evaluated the accuracy of the aforementioned models by measuring the mean-square error (MSE) and the root-mean-square error (RMSE) for each model. For Ridge and Lasso I did hyper parameter tuning and measured the Root Mean Squared Error (RMSE) with the new parameters. The R-squared value was also calculated in the study which is a measure that provides information about the goodness of fit of a model.

R-squared was calculated using [26],

$$R^2 = 1 - \frac{\text{sum squared regression (SSR)}}{\text{total sum of squares (SST)}}, \\ = 1 - \frac{\sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{y})^2}.$$

Fig. 8. R-squared Equation

RMSE was calculated using the equation [27],

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(f_i - \bar{f}_i)^2}{N}}$$

Fig. 9. RMSE Equation

VI. Results and Comparison

In this section, I am presenting the results of various ML regression and NN models which estimate the ratio of burned areas for each observation. From the results, I found that

Models	R-squared	MSE (km^4)	RMSE (km^2)	Parameter Tested	Best Parameter
Linear	0.098	1.178	1.096	-	-
Ridge	0.248	1.049	1.001	alpha	'alpha': 10.0
Lasso	0.247	0.997	1.002	alpha	'alpha': 0.0
Decision Tree	-0.258	1.67	1.295	-	-
Random Forest	0.343	0.867	0.936	-	-
XGBoost	0.319	0.897	0.95	-	-
ANN	-0.0005	1.334	1.155	-	-

XGBoost outperformed all the other models, and from all models, XGBoost and Random Forest had a very close RMSE. The proportion of the variance, R-squared value of Random Forest was slightly higher than XGBoost. The model XGBoost regression gave a RMSE of 0.95 (95%), and Random Forest regression gave a RMSE of 0.936 (93.6%).

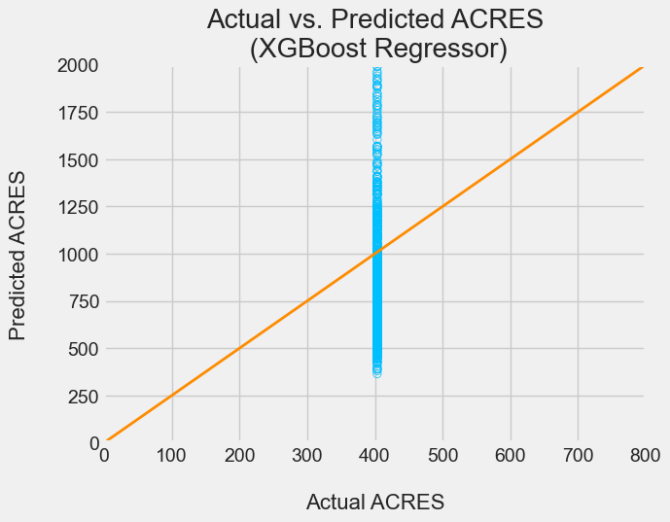

Fig. 10. XGBoost Regressor Prediction Plot

VII. Discussion

In this paper I applied multiple ML models and a NN model to predict the occurrence and size of Wildfires from the collected dataset.

The models included linear, ridge, lasso, decision tree, random forest, XGBoost, and ANN regressions. The models were trained using a large dataset which includes wildfire observations from the year 1984 to 2021. Since I had to select the target, it took a good amount of time just to find the target feature. I predicted using the ACRES at first, but that gave a large RMSE. After exploring the correlation matrix, I applied Pearson correlation test on a few features and found a better suited target variable, DNBR_OFFST. But then I standardized the ACRES feature and finally used it as the target variable for prediction. Upon training, the best models showed promising prediction accuracies and could predict the likelihood of wildfires with burned area ratios with RMSE scores of 0.95 km^2.
One of the most troubling issues I faced was hyper parameter tuning Artificial Neural Network. A promising prospect to enhance wildfire alerts and give forest managers tools to analyze regional wildfire risk is accurate wildfire hazard estimation by Machine Learning models.

VIII. Conclusion

The ACRES provided a flexible, robust, analytically simple approach that could be applied anywhere within any continents.

The study's findings show that it is possible to forecast the size of wildfires using quantitative, numerical, and categorical data. This information will be useful for forest fire prevention and rescue, particularly for wildfires that start in forests. The results of this research demonstrate the superiority of machine learning (ML) models over conventional fire weather indices for estimating wildfire hazard.

Acknowledgment

I greatly appreciate Lecturer Meem Arafat Manab, for his constructive comments to help improve the paper.

References

[1] U. A. S. Service, "Report on the michigan forest fires of 1881." Office of the Chief Signal Officer:Washington, DC, USA, 1881, p. 37.
[2] C. D. T. Craig C. Chandler, Theodore G. Storey, "Report on the michigan forest fires of 1881." Res. Paper PSW-RP-5. Berkeley, CA: Pacific Southwest Forest Range Experiment Station, Forest Service, U.S. Department of Agriculture, 1963, p. 110.
[3] A. A. Alencar, L. A. Solórzano, and D. C. Nepstad, "Modeling forest understory fires in an eastern amazonian landscape," *Ecological Applications*, vol. 14, no. sp4, pp. 139–149, 2004.
[4] M. G. Rollins, P. Morgan, and T. Swetnam, "Landscape-scale controls over 20th century fire occurrence in two large rocky mountain (usa) wilderness areas," *Landscape Ecology*, vol. 17, no. 6, pp. 539–557, 2002.
[5] D. T. Bui, B. Pradhan, H. Nampak, Q.-T. Bui, Q.-A. Tran, and Q.-P. Nguyen, "Hybrid artificial intelligence approach based on neural fuzzy inference model and metaheuristic optimization for flood susceptibilitgy modeling in a high-frequency tropical cyclone area using gis," *Journal of Hydrology*, vol. 540, pp. 317–330, 2016.
[6] H. Hong, M. Panahi, A. Shirzadi, T. Ma, J. Liu, A.-X. Zhu, W. Chen, I. Kougias, and N. Kazakis, "Flood susceptibility assessment in hengfeng area coupling adaptive neuro-fuzzy inference system with genetic algorithm and differential evolution," *Science of the total Environment*, vol. 621, pp. 1124–1141, 2018.
[7] D. Chaparro, M. Vall-Llossera, M. Piles, A. Camps, C. Rüdiger, and R. Riera-Tatché, "Predicting the extent of wildfires using remotely sensed soil moisture and temperature trends," *IEEE journal of selected topics in applied earth observations and remote sensing*, vol. 9, no. 6, pp. 2818–2829, 2016.
[8] C. Lai, S. Zeng, W. Guo, X. Liu, Y. Li, and B. Liao, "Forest fire prediction with imbalanced data using a deep neural network method," *Forests*, vol. 13, no. 7, p. 1129, 2022.
[9] S. Lall and B. Mathibela, "The application of artificial neural networks for wildfire risk prediction," in *2016 International Conference on Robotics and Automation for Humanitarian Applications (RAHA)*. IEEE, 2016, pp. 1–6.
[10] H. Zhang, X. Han, and S. Dai, "Fire occurrence probability mapping of northeast china with binary logistic regression model," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 1, pp. 121–127, 2013.
[11] ——, "Fire occurrence probability mapping of northeast china with binary logistic regression model," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 1, pp. 121–127, 2013.
[12] D. Radke, A. Hessler, and D. Ellsworth, "Firecast: Leveraging deep learning to predict wildfire spread." in *IJCAI*, 2019, pp. 4575–4581.
[13] A. Shmuel and E. Heifetz, "Global wildfire susceptibility mapping based on machine learning models," *Forests*, vol. 13, no. 7, p. 1050, 2022.
[14] S. Cheng, Y. Jin, S. P. Harrison, C. Quilodrán-Casas, I. C. Prentice, Y.-K. Guo, and R. Arcucci, "Parameter flexible wildfire prediction using machine learning techniques: Forward and inverse modelling," *Remote Sensing*, vol. 14, no. 13, p. 3228, 2022.
[15] A. P. S. Gulati, "https://github.com/aman-preet-singh-gulati/forest-fire-predicition."
[16] A. Rana, "https://github.com/ashishrana1501/forest-fire-prediction."
[17] P. Sohn, "https://github.com/psohn/only-you-can-prevent-forest-fires."
[18] fs.usda, "https://data.fs.usda.gov/geodata/edw/datasets.php."
[19] USDA, "https://catalog.data.gov/organization/usda-gov."
[20] H. Zainodin, A. Noraini, and S. Yap, "An alternative multicollinearity approach in solving multiple regression problem," *Trends in Applied Sciences Research*, vol. 6, no. 11, p. 1241, 2011.
[21] D. H. Moore, "Classification and regression trees, by leo breiman," 1984.
[22] B. Leo, "Statistical modeling: The two cultures," *Statistical science*, vol. 16, no. 3, pp. 199–231, 2001.
[23] J. Brownlee, *XGBoost With python: Gradient boosted trees with XGBoost and scikit-learn*. Machine Learning Mastery, 2016.

[24] A. Alonso-Betanzos, O. Fontenla-Romero, B. Guijarro-Berdiñas, E. Hernández-Pereira, J. Canda, E. Jimenez, J. L. Legido, S. Muñiz, C. Paz-Andrade, and M. I. Paz-Andrade, "A neural network approach for forestal fire risk estimation," in *Proceedings of the 15th European Conference on Artificial Intelligence*, 2002, pp. 643–647.
[25] S. Haykin, *Neural networks and learning machines, 3/E*. Pearson Education India, 2009.
[26] NCL, "https://www.ncl.ac.uk/webtemplate/ask-assets/external/maths-resources/statistics/regression-and-correlation/coefficient-of-determination-r-squared.html."
[27] statisticshowto, "https://www.statisticshowto.com/probability-and-statistics/regression-analysis/rmse-root-mean-square-error/."