

FOREST FIRE PREDICTION

Data set and task

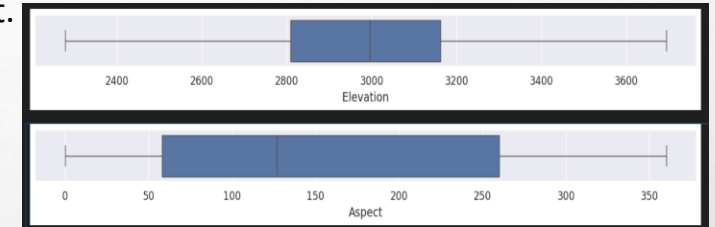
- ❖ Data set: The training size has (581012) samples and the number of features:55
- ❖ Goal: To Predict the Forest Fire and Save the Environment
- ❖ Classifier: Random Forest

Elevation	Aspect	Slope	Horizontal_Distance_To_Hydrology	Vertical_Distance_To_Hydrology	Horizontal_Distance_To_Roadways	Hillshade_9am	
0	2596	51	3	258	0	510	221



ML Challenges

- ❖ Class Imbalance- The distribution of classes in a classification dataset is not uniform
- ❖ Feature Selection- To standardize or normalize the range of independent variables or features in the dataset .
- ❖ Outlier- The data points that significantly deviate from the rest of the observations in a dataset.



Key Experimental Result

Outliers

Random Forest	Do Nothing	Trimming	Interquartile Range (IQR)
Accuracy	94.82	94.9302	94.8544
Mean Squared Error	0.26884	0.26559	0.26533
Mean Absolute Error	0.09243	0.09116	0.09171
R-Squared Error	0.86163	0.8633	0.86343

Future Work

- ❖ Recursive Feature Elimination (RFE)
- ❖ Principal Component Analysis (PCA)
- ❖ Also, by adding more feature in dataset like Temperature , Humidity, Windspeed , Vegetation type and more.

By Implementing the above method can Improve my Model prediction for project evaluation method.