Is a random forest machine learning model effective in predicting wildfires in Southern California?

1. Collect public datasets: fire history, weather, and vegetation
2. Python was used to read and process the data.
3. Random generation was made to test the model to training dataset for the machine learning model. A random position and time was chosen and the relevant data points were removed from the datasets.
4. Random forest was used for training and bagging and SHAP were used for hyperparameter tuning.

Figure 1: Precision is the number of true positives over the number of true positives and recall is the number of true positives over the number of the number of true positives and false negatives. The F1 score is a harmonic mean of the precision and recall values.

Figure 2: Predictor variables, called features, are listed in order of most predictive at the top and least predictive at the bottom.

Figure 3: The SHAP value represents the impact of each feature on the model's prediction. If the feature value becomes less red and more blue moving from left to right, it is inversely related to wildfire occurrence, since as the value becomes lower (bluer) the wildfire occurrence value will become higher, and vice versa.

Further work includes training a machine that considers weather conditions before the fire when creating the combined dataset is resistant to seasonal changes in weather.

This project was successful in producing a machine learning model that is able to predict the occurrence of a wildfire in Ventura County during a six-year period due to its accuracy of 97.0%.

Out of all predictor variables, used the normalized difference vegetation index, surface pressure, and volumetric soil water were the three most important predictor variables in the model's prediction-making.

Due to the seasonality of wildfires in Ventura County, seasonal shifts in the values of predictor variables may be associated with wildfires.

This project was successful in producing a machine learning model that is able to predict the occurrence of a wildfire in Ventura County during a six-year period due to its accuracy of 97.0%.

Out of all predictor variables, used the normalized difference vegetation index, surface pressure, and volumetric soil water were the three most important predictor variables in the model's prediction-making.

Due to the seasonality of wildfires in Ventura County, seasonal shifts in the values of predictor variables may be associated with wildfires.

Further work includes training a machine that considers weather conditions before the fire when creating the combined dataset is resistant to seasonal changes in weather.

This project was successful in producing a machine learning model that is able to predict the occurrence of a wildfire in Ventura County during a six-year period due to its accuracy of 97.0%.

Out of all predictor variables, used the normalized difference vegetation index, surface pressure, and volumetric soil water were the three most important predictor variables in the model's prediction-making.

Due to the seasonality of wildfires in Ventura County, seasonal shifts in the values of predictor variables may be associated with wildfires.

Further work includes training a machine that considers weather conditions before the fire when creating the combined dataset is resistant to seasonal changes in weather.

This project was successful in producing a machine learning model that is able to predict the occurrence of a wildfire in Ventura County during a six-year period due to its accuracy of 97.0%.

Out of all predictor variables, used the normalized difference vegetation index, surface pressure, and volumetric soil water were the three most important predictor variables in the model's prediction-making.

Due to the seasonality of wildfires in Ventura County, seasonal shifts in the values of predictor variables may be associated with wildfires.

Further work includes training a machine that considers weather conditions before the fire when creating the combined dataset is resistant to seasonal changes in weather.

This project was successful in producing a machine learning model that is able to predict the occurrence of a wildfire in Ventura County during a six-year period due to its accuracy of 97.0%.

Out of all predictor variables, used the normalized difference vegetation index, surface pressure, and volumetric soil water were the three most important predictor variables in the model's prediction-making.

Due to the seasonality of wildfires in Ventura County, seasonal shifts in the values of predictor variables may be associated with wildfires.

Further work includes training a machine that considers weather conditions before the fire when creating the combined dataset is resistant to seasonal changes in weather.

This project was successful in producing a machine learning model that is able to predict the occurrence of a wildfire in Ventura County during a six-year period due to its accuracy of 97.0%.

Out of all predictor variables, used the normalized difference vegetation index, surface pressure, and volumetric soil water were the three most important predictor variables in the model's prediction-making.

Due to the seasonality of wildfires in Ventura County, seasonal shifts in the values of predictor variables may be associated with wildfires.

Further work includes training a machine that considers weather conditions before the fire when creating the combined dataset is resistant to seasonal changes in weather.

This project was successful in producing a machine learning model that is able to predict the occurrence of a wildfire in Ventura County during a six-year period due to its accuracy of 97.0%.

Out of all predictor variables, used the normalized difference vegetation index, surface pressure, and volumetric soil water were the three most important predictor variables in the model's prediction-making.

Due to the seasonality of wildfires in Ventura County, seasonal shifts in the values of predictor variables may be associated with wildfires.

Further work includes training a machine that considers weather conditions before the fire when creating the combined dataset is resistant to seasonal changes in weather.

This project was successful in producing a machine learning model that is able to predict the occurrence of a wildfire in Ventura County during a six-year period due to its accuracy of 97.0%.

Out of all predictor variables, used the normalized difference vegetation index, surface pressure, and volumetric soil water were the three most important predictor variables in the model's prediction-making.

Due to the seasonality of wildfires in Ventura County, seasonal shifts in the values of predictor variables may be associated with wildfires.

Further work includes training a machine that considers weather conditions before the fire when creating the combined dataset is resistant to seasonal changes in weather.