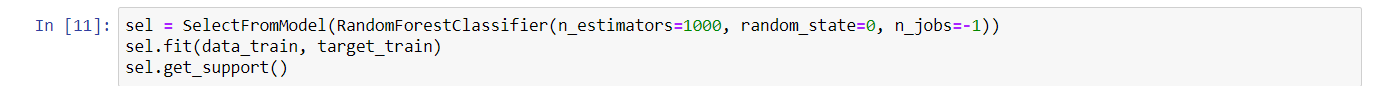
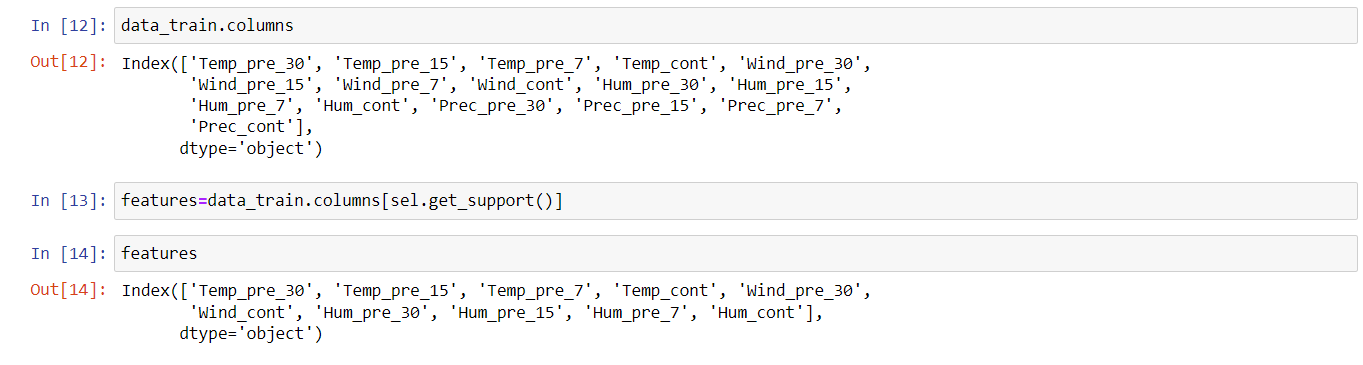
Background:

Over the last few centuries forest landscapes across America have suffered from dramatic climate change. As result, many changes in landscape conditions have resulted in an increase in wildfires across the United States. Unquestionably many of the fires were caused by humans. Using data provided by the U.S. National Interagency Fire Center, we investigate the possibility of predicting future wildfire events based on fires from within the past 20 years. Specifically, Suppression cost which includes DOI Agencies and Forest Services varies from $239,943,000 to $2,274,000,000 from 1985 to 2020 respectively [1]. Data include the size of fire, topography, terrain, and cause would be used to help predict an effective evacuation and mitigation strategy. This could become an essential tool when planning evacuation strategies and wildfire prevention.   
This research will focus on training a machine learning model in order to predict the magnitude of the fire intensity based on natural weather condition by using random forest classifier and recursive feature elimination to remove the feature.

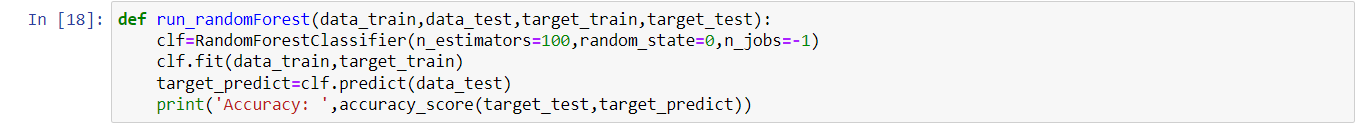
Methods & Dataset

This data [2-3] consists of wildfire reports across America from 1992-2014 with 1.88 million that were randomly sampled to produce a set of approximately 55,000 data points. Irrelevant features were eliminated using the Feature Selection-filter method and then were further paired down to final set of 10 using Recursive Feature Elimination.





NumPy and Pandas were some of the Python libraries used to manipulate the data. The Scikit-learn library was also used to run the algorithms. For this particular issue and dataset, we selected the Random Forest Algorithm due to the nature of dataset and the algorithm’s demonstrated accuracy in similar studies.



The original dataset consists of fire size classified using a A to F scale. However, the scaling factor between each classification was not identified. In our original calculations this appeared to throw off the precision of the predictions. We decided to adjust this range to only three categories; small, medium, and large. 70% of the original data set was allocated to train the model with the remaining data reserved for testing.





Result & Discussion

The Random Forest Classifier proved 87.2% accurate in predicting the magnitude of a fire’s intensity based on the set of 10 features in our dataset. For a more accurate prediction, it would be essential to acquire data points.

Direction For Future Research

This dataset contains information about amenities around wildfire areas, latitude, longitude and cause of fire. Future research can investigate how these factors play a role in the severity and frequency of wildfires.

Additionally, an improved model may allow emergency response teams to be better aware of the potential hazards a newly reported fire would present. This information could lead to better resource allocation and safety improvements.

References

[1] *Suppression costs*. National Interagency Fire Center. (n.d.). Retrieved February 3, 2022, from https://www.nifc.gov/fire-information/statistics/suppression-costs

[2] Short, K. C. (n.d.). *Spatial wildfire occurrence data for the United States, 1992-2015 [FPA\_FOD\_20170508] (4th edition)*. Home. Retrieved February 3, 2022, from https://www.fs.usda.gov/rds/archive/Catalog/RDS-2013-0009.4/

[3] Capcloudcoder. (2020, October 6). *U.S. wildfire data (plus other attributes)*. Kaggle. Retrieved February 3, 2022, from <https://www.kaggle.com/capcloudcoder/us-wildfire-data-plus-other-attributes/version/4?select=Wildfire_att_description.txt>