Wildfire Prediction using LFMC Maps

Manvendra Kumar Nema
MT22038

Shambhavi Pathak
MT22067

Tarang Dineshbhai Viroja
MT22081

Abstract

As we see a rise in the temperature across the globe, its potential to burn has also increased. Over the past few decades, there have been many incidents of occurrence of wildfires in the world. An earlier prediction of forest fires can help in better wildfire risk management. To predict wildfires, one of the key indicator is live fuel moisture content (LFMC), which means the mass of water per unit dry biomass in vegetation. LFMC data is captured at a large scale based on microwave remote sensing and eventually represented as maps. Using LFMC maps and the ground truth wildfire data from fire.ca.gov, we predicted whether there will be a fire in the near future within the $8km * 8km$ grid using machine learning models such as SVM and random forest. Our results show that Random Forest and SVM gave an accuracy of 91.69% and 80% respectively. Concluding that *Random Forest* performs better.

1 Introduction

Since the 1980's, more than a quarter of planet's vegetated surface has been facing prolonged wildfire seasons as a result of the global warming, and in some places fire has become nearly a year-round risk. One of such place is California. Wildfire can cause huge devastation to humans and environment and forecasting it in advance, can help in improving the management of wildfire risk.

Among the many causes of wildfire, live fuel moisture content has been shown to be a key determinant of fire ignition and spread. Live fuel moisture content refers to the mass of plant water per unit dry biomass. To explain in simpler terms LFMC can be referred as forest dryness because it indicates the amount of water in fuels, like trees, shrubs and plants, relative to their dry biomass. Lower the live fuel moisture content, drier the fuels, greater the risk of wildfires, and vice-versa.

The estimates of LFMC from optical remote sensing or meteorological indices have been noticed as insufficient to accurately map LFMC at landscape scale. The dataset we are using is generated using *physics-assisted recurrent neural network model for mapping LFMC every 15 days at 250 m resolution over the western US using microwave backscatter (from Sentinel-1) and optical reflectance (from Landsat-8).*

Our analysis has been completely based on the dataset provided at `https://github.com/kkraoj/lfmc_from_sar`. These maps are available from January, 2016 to April, 2021.

2 Problem Definition and Algorithm

Problem Definition: Given the image dataset, determine whether there will be fire in the next 15 days within $8km * 8km$ grid

2.1 Data Collection

The dataset we have are LFMC maps that are downloaded using Google earth engine API and the ground truth wildfire data has been taken from California Department of Forestry and Fire Protection over the same duration.

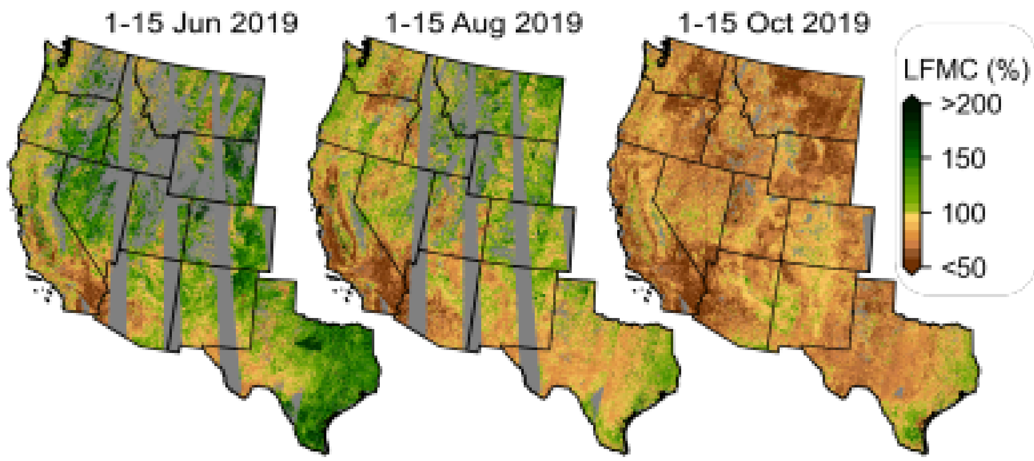

Figure 1: Sample images from LFMC dataset.

2.2 Data Preprocessing

To be able to achieve the task of predicting the wildfires using LFMC maps, the dataset required preprocessing. The LFMC maps collected were .tif files of 250m resolution i.e. 1 pixel = 250m, of California region and were generated using Google Earth.

The wildfire ground truth data included the date of fire, location coordinates generated using Google Maps and were created as .shp file.

To plot the Fire points on the LFMC maps using the wildfire ground truth data, we were requird to convert the Google Maps location points to Google Earth points as Google Maps is in a projected coordinate system that is based on the *wgs84 datum. (EPSG 3857)* and Google Earth is in a Geographic coordinate system with the *wgs84 datum. (EPSG: 4326)*

From the dataset, the nearest previous date from a fire occuring date, LFMC map was selected to plot the fire points and a grid of $8km * 8km$ with fire point as mid-point, a $32 * 32pixels$ image was cropped. Using this method we filtered out the fire points for the entire duration.

For no fire, we randomly sampled a data to extract an image from within the 8 months span of fire occurence date. The dimensions were kept same as fire data.

The final dataset contains 1928 fire samples and 1928 no fire samples and for machine learning models, the data was split into 80% training set and 20% testing set.

2.3 Machine Learning Models

We experimented with SVM and Random Forest machine learning algorithms for our ananlysis.

SVM: In machine learning, support vector machines are supervised learning models with associated learning algorithms that analyze data for classification. SVMs are one of the most robust prediction methods, and has the benefit of efficiently performing a non-linear classification by implicitly mapping their inputs into high-dimensional feature spaces.

Random Forest: Random forests or random decision forests is an ensemble learning method for classification and regression that operates by constructing a multitude of decision trees at training time. For classification tasks, the output of the random forest is the class selected by most trees.

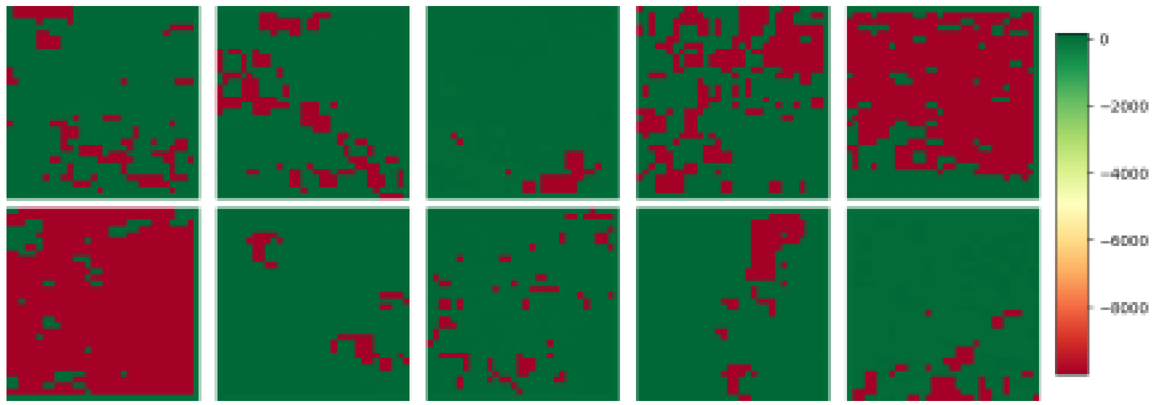

Figure 2: Sample $32 * 32$ pixels fire images.

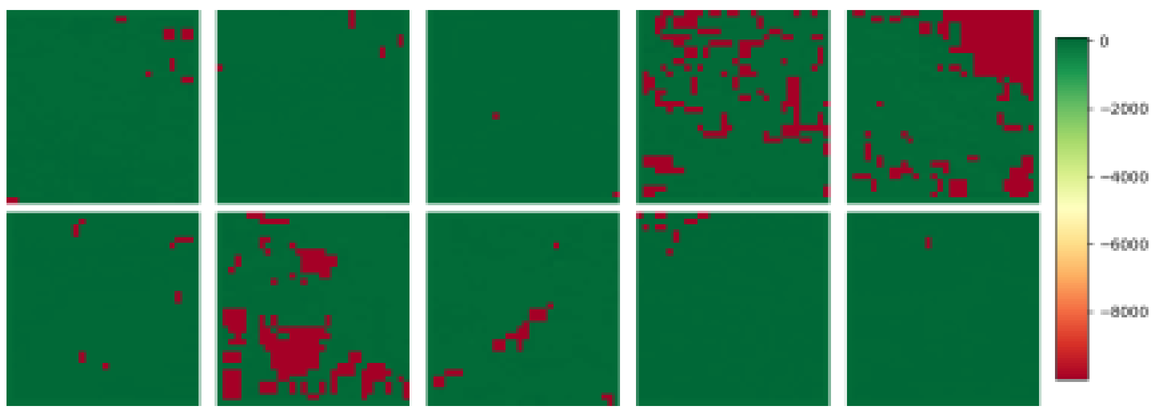

Figure 3: Sample $32 * 32$ pixels no fire images.

3 Evaluation

For the evaluation of our Machine learning models, we chose accuracy as we had a balanced data of fire and no-fire and accuracy works best for calculating the percentage of correct predictions made by our classification model.

$$Accuracy = \frac{no.of correct predictions}{total data}$$

Other than Accuracy, we also considered *F1 score* as it gives a trade-off between false negatives and false positives and for the current data, avoiding both is equally inportant.

$$\textit{False positive} = \textit{An area has been wrongly classified as fire}$$

$$\textit{False negative} = \textit{An area has been wrongly classified as no-fire}$$

$$\textit{True Positive} = \textit{An area has been correctly classified as fire}$$

$$\textit{True Negative} = \textit{An area has been correctly classified as no-fire}$$

$$Precision = \frac{TP}{TP + FP}$$

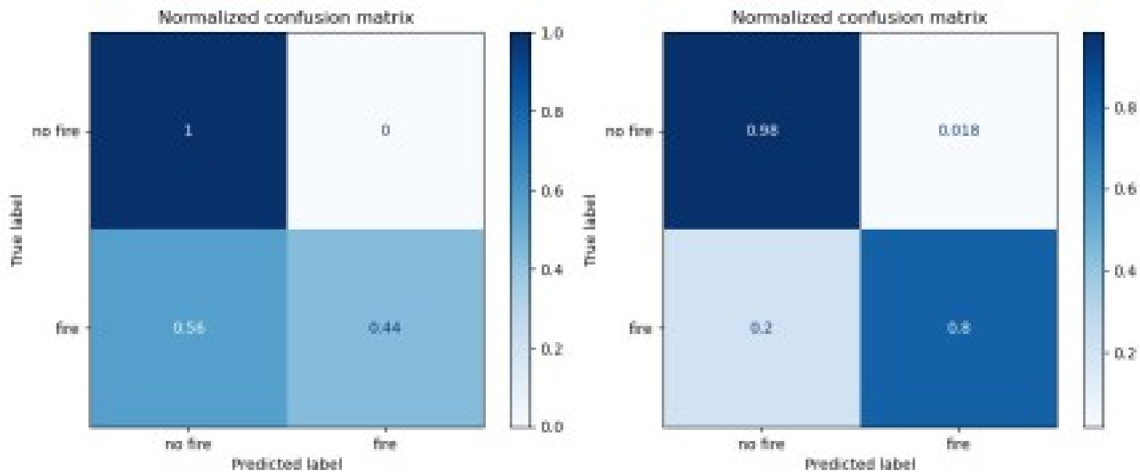

Figure 4: Confusion matrix of SVM result (left) and Confusion matrix of Random Forest (right)

$$Recall = \frac{TP}{TP + FN}$$

$$F1\ Score = \frac{1}{\frac{1}{Precision} + \frac{1}{Recall}}$$

The reason we chose f1 score was because a false positive will affect the state's economy by adding on the expenses of migrating people from there in vain and then bringing them back as the prediction of there being a fire was false. Similarly, a false negative will impact the human lives, as no one was aware about there being a fire.

In our analysis f1 score and accuracy of SVM model is 77% and 80% respectively. F1 score and accuracy of Random Forest is 92% and 91.69% respectively. Thus, implying that Random forest worked better for our dataset.

4 Results and Discussion

We used the LFMC image corresponding to the actual date of fire as the input to the model. The project was implemented in Python using Sklearn. First, we experimented by using SVM as the classifier. We performed hyperparameter tuning to find the best parameters using grid search for our RBF kernel. The optimal SVM parameters is $gamma = 0.001$ and the regularization $C = 1$. Similarly, we performed hyperparameter tuning to find the best parameters of random forest, the optimal parameters of random forests are $max\_depth = 30$, $min\_samples\_leaf = 1$, $min\_samples\_split = 5$ and $n\_estimators = 100$.

5 Conclusion

In this project our goal was to predict the occurrence of a fire in the near future given current fuel moisture levels? LFMC Maps and actual wildfire data from the *CAL* served as the input data. The result is a binary classification forecast of whether fire or no fire will break out soon in the area. For forecasting, we utilised a SVM and random forest. The Random Forest algorithm fared best in the first trial, with an accuracy of 91.69%.

Looking at such good results, and maybe after further analysis on the future LFMC maps that will be generated we believe, LFMC as an indicator can be used in India for mitigating wildfire risk.

6 References

- Rao, K., Williams, A.P., Fortin, J. & Konings, A.G. (2020). SAR-enhanced mapping of live fuel moisture content. Remote Sens. Environ., 245.
- Towards Accurate Fire Predictions Using AI, Krishna Rao, Medium