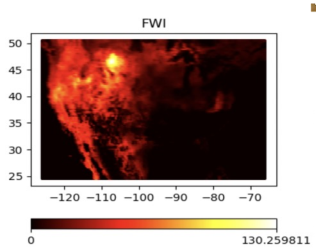
A Machine Learning based Wildfire Emission Forecasting Workflow

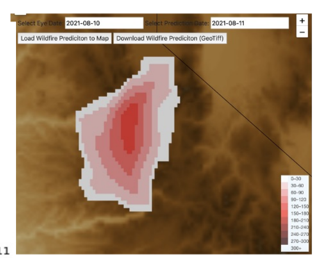
Ziheng Sun, Yunyao Li, Daniel Tong, Kate Zhang, Shan Sun
George Mason University; UC Boulder; NOAA

Abstract

The increasing frequency and severity of wildfires globally necessitate advanced forecasting models to mitigate their environmental and societal impacts effectively. This presentation focuses on leveraging machine learning, particularly LightGBM, to predict Fire Radiative Power (FRP), a critical indicator of wildfire intensity and emissions. Addressing current limitations in sub-seasonal to seasonal (S2S) forecasting, the presentation integrates a comprehensive dataset comprising meteorological inputs, geographical coordinates, historical FRP data spanning five days, NASA vegetation indices, and land use classifications. Through rigorous comparison, LightGBM emerges as optimal for its robust handling of complex data and superior predictive capabilities. The presentation delves into the detailed AI model development process, starting with preprocessing techniques such as missing data handling and feature engineering tailored for wildfire prediction. It emphasizes parameter tuning, model training strategies specific to LightGBM, and validation methodologies including cross-validation techniques to ensure model robustness across varied datasets and temporal scales. Furthermore, interpretability tools for AI models in wildfire forecasting, encompassing feature importance and output visualization, will be explored. This presentation aims to demonstrate advancements in the NOAA emission forecasting system by bridging AI capabilities with environmental science, thereby advancing early warning systems, supporting firefighting operations, and safeguarding public health through improved S2S wildfire prediction and mitigation strategies.

	Model	MAE	MSE	RMSE	R2	RMSLE	MAPE	TT (Sec)
lightgbm	Light Gradient Boosting Machine	1.2403	4.0905	2.0224	0.5853	0.4438	3.0546	0.3840
rf	Random Forest Regressor	1.2720	4.2205	2.0543	0.5721	0.4454	3.0234	27.1290
gbr	Gradient Boosting Regressor	1.2911	4.2201	2.0542	0.5721	0.4487	3.0747	6.5430
et	Extra Trees Regressor	1.5319	4.7066	2.1694	0.5228	0.4766	3.1198	8.4200
ada	AdaBoost Regressor	2.1254	5.8255	2.4132	0.4095	0.4926	2.9053	1.1760



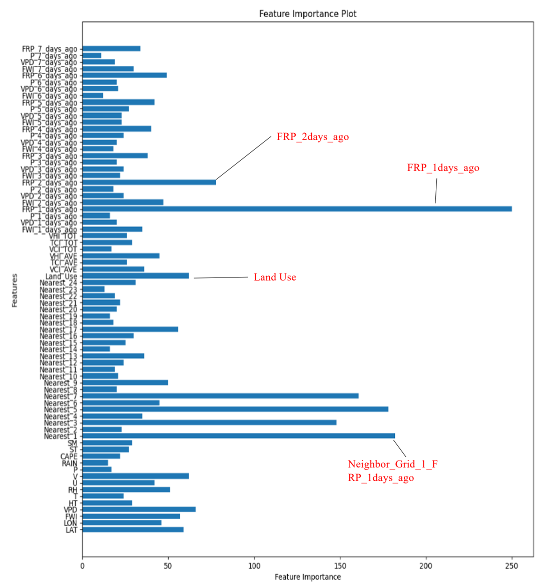


Methods and Materials

LightGBM, a powerful gradient boosting framework, is used to predict fire emission by using diverse environmental datasets. We begin by preprocessing and integrating data from various sources, including satellite imagery (e.g., MODIS), meteorological stations, and historical fire records. Key input variables for our LightGBM model include temperature, humidity, wind speed, vegetation type, and topographical features. These features are used to train the model to predict the Fire Radiative Power (FRP), an indicator of fire intensity and energy output. The data, sourced from NASA Earth observation systems and local meteorological agencies, provide a comprehensive view of fire risk factors and support accurate, real-time predictions for effective wildfire management and response. Geoweaver is used to support the full-stack workflow management of the entire processing pipeline.

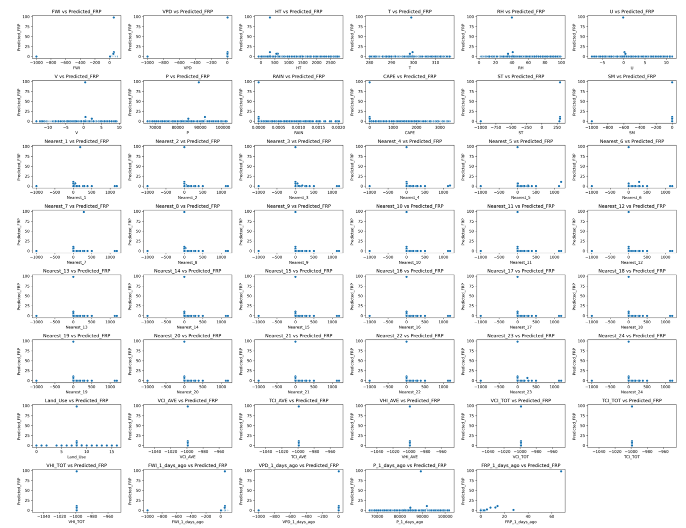
$$F_t(x) = F_{t-1}(x) + \eta \cdot h_t(x)$$

the previous model — Learning rate — New fit model on current gradient



Results

The model has been trained on various datasets, each representing different dates. Overall, the performance metrics vary, showing a range of Mean Squared Error (MSE) from approximately 0.014 to 0.810, with Root Mean Squared Error (RMSE) ranging from 0.119 to 0.900. The Mean Absolute Error (MAE) fluctuates between 0.086 and 0.591. The model is consistently saved to the same file path (`fc_lightgbm_single.pkl`), indicating updates with each new dataset processed. The variations in error metrics suggest that the model's performance is highly dependent on the specific dataset used, with some cases showing significantly better fit than others.



Contact

Ziheng Sun zsun@gmu.edu
703-993-6124

Open Source Geoweaver Workflow GitHub Repository:
https://github.com/GMU-SESS-AQ/firecasting

Wildfire AI Interactive Website:
http://geobrain.csiss.gmu.edu/firecasting_site/

References

Zhang, Li, Georg Grell, Partha Bhattacharjee, Shan Sun, Anders Jensen, Jordan Schnell, Haiqin Li et al. *Predicting Fire Aerosols and their Impact on Subseasonal to Seasonal Weather Forecasts in NOAA's Global Aerosol Forecast Systems*. No. EGU24-6827. Copernicus Meetings, 2024.

Sun, Ziheng. "Actionable Science for Wildfire." In *Actionable Science of Global Environment Change: From Big Data to Practical Research*, pp. 149-183. Cham: Springer International Publishing, 2023.

Li, Yunyao, Daniel Tong, Li Kate Zhang, Shan Sun, Ziheng Sun, Sophia Tang, and Arushi Desai. "Global Ensemble Fire Emission Dataset and Subseasonal Wildfire Emission Forecast." In *104th AMS Annual Meeting*. AMS, 2024.