Ulab Final Report

Wildfire Detection

Mentors: Hunter McCoy & Nandita Radhakrishnan

Team: Vaibhav Mohata, Pratyush Das, Shani Lyubomirsky, Neel Datta, Peter Grinde-Hollevik, Sifath Mannan, Aashritha Srirambatla

Content Overview

Introduction

Goals

Data

Methods

Models

Potential Issues

Final Model

Evaluation

Summary

Introduction

Wildfires are a growing natural hazard in most regions of the United States, posing a threat to life and property, particularly where native ecosystems meet developed areas. **(**US Geological Survey, 2006). The societal consequences are clear: Just in the United States alone, over 13 million acres of land burned, $3.2 trillion were spent in wildfire suppression costs, and over 14.700 structures were lost in the 2020 wildfire season. (National Large Incident Year-to-Date Report, 2020).

While the current research is good at generating models for current or immediately-close fires, there is still a gap in the use of data in considering the effects of climate change on the intensity of wildfires.

We look to close that gap with our research and through the development of our prediction engine.

Goals

Identify the instantaneous and future likelihoods of fires in a certain location.

Achieve a higher prediction accuracy using a temporal data-based approach in comparison to current measurements of wildfire prediction.

Train the model from scratch to achieve up to 92% accuracy

Data

- Our data was sourced from the MODIS LSAT(Land Surface Air Temperature), Vegetation, and Fire 8-Day V3 datasets.
- We limited our data to the time period between May 2018 to June 2020 and focused on Eastern Australia (Queensland, New South Wales, Victoria)

Initial Approaches and Methods

Data Cleaning and Aggregation

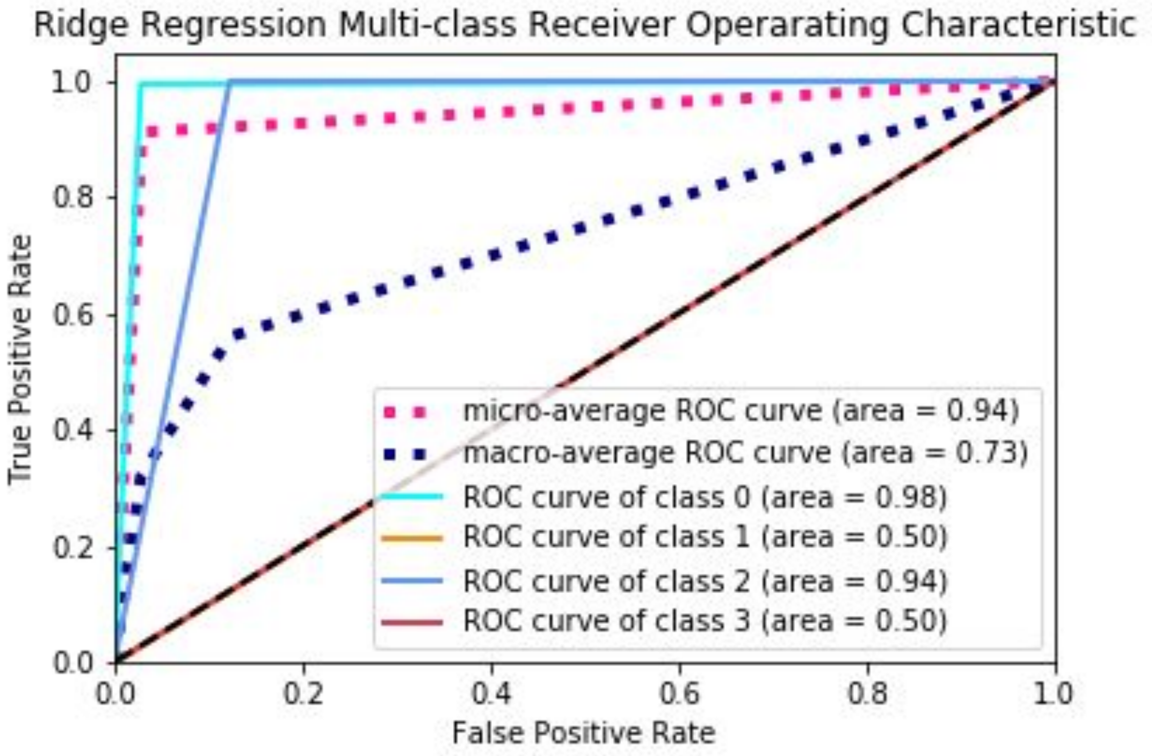
- To prepare the data, we extracted it from the hdf files and applied a reverse transform to get each pixel in terms of its latitude and longitude.
- We then performed an averaging pooling to reduce the size of the problem, and aggregated the three datasets along latitude, longitude, and date.
- We then removed samples with incorrect or partial data, and removed any samples that weren't given a high accuracy by MODIS.
- The remaining data was split into training, validation, and testing sets, and boosted to ensure equal numbers of all samples.

Models

1. **Ridge Regression**

   Accuracy: 91.5 %
   AUC:

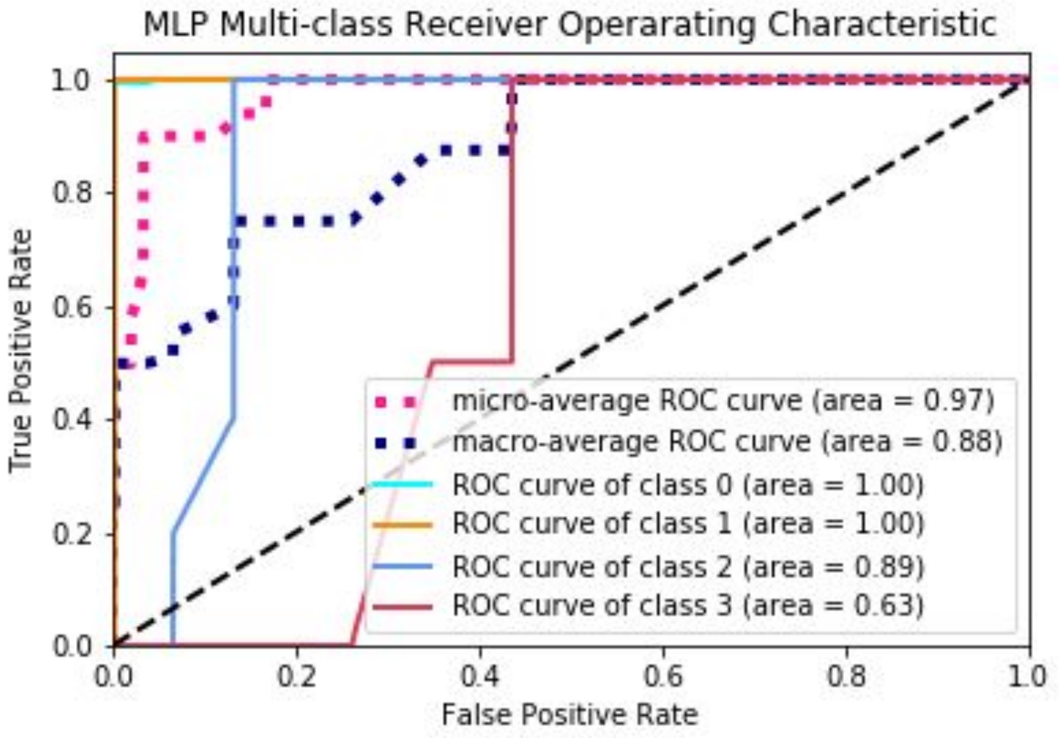
Class 0: No fire
Class 1: Low
Class 2: Moderate
Class 3: High



2. MLP

Accuracy: 91.33
AUC:

Class 0: No fire
Class 1: Low
Class 2: Moderate
Class 3: High



Potential issues with models

Main Issue: The nature of wildfire data makes it difficult to accurately train a model.

- The model has difficulty giving high fire rating. Looking at the AUC curve, classes No Fire, low and Moderate Fire were all at approximately 90% AUC. Contrast that with High Fire, which had a much worse 63% AUC.

  A potential reason for the above result is that majority of the real-life reads come from areas with no fire, meaning that the relative portion of fire samples is low. We fixed this in our dataset with boosting, but it wasn't enough to completely correct the issue.

Summary - Takeaways

The final model achieved a test accuracy of 91.33%. Due to limited time and resources, it is difficult to train a model from scratch to reach optimal performance. Even with a complex architecture, the model had difficulty surpassing the threshold in its predictions.

With more time and budget for training, we are confident we can increase accuracy to a near-optimal level.

We learned how to use Amazon AWS, MODIS, and pyspark build a large scale model for wildfire detection.

References

1. U.S. Geological Survey, U. S. D. of the I. (2006). Wildfire Hazards—A National Threat https://pubs.usgs.gov/fs/2006/3015/2006-3015.pdf
2. National Interagency Fire Center. (2020). National Large Incident Year-to-Date Report.

   https://gacc.nifc.gov/sacc/predictive/intelligence/NationalLargeIncidentYTDReport.pdf

3. Walsh, C. (2020, January 31). *The long-term effects of wildfires*. Harvard Gazette. https://news.harvard.edu/gazette/story/2020/01/the-long-term-effects-of-wildfires/
4. Y.O. Sayad, H. Mousannif, H. Al Moatassime, Predictive modeling of wildfires:A new dataset and machine learning approach, Fire Safety Journal (2019), doi: https://doi.org/10.1016/j.firesaf.2019.01.006
5. Coen, J. Douglas, C. (2020) Computational Modeling of Large Wildfires: A Roadmap https://opensky.ucar.edu/islandora/object/articles%3A18989/datastream/PDF/download/citation.pdf
6. P. Cortez and A. Morais, “A Data Mining Approach to Predict wildfires using Meteorological Data,” Proc. 13th Port. Conf. Artif. Intell., pp. 512–523, 2007https://dergipark.org.tr/tr/pub/dumf/issue/56929/661925
7. O. Rios, E. Pastor, M.G. Valero, E. Planas (January 2016) “Short-term fire front spread prediction using inverse modelling and airborne infrared images” https://www.researchgate.net/publication/309531321_Short-term_fire_front_spread_prediction_using_inverse_modelling_and_airborne_infrared_images