Data Mining Meteorological Data to Predict Forest Fires

Christopher de Freitas, Pranav Rajan

THE UNIVERSITY OF UTAH

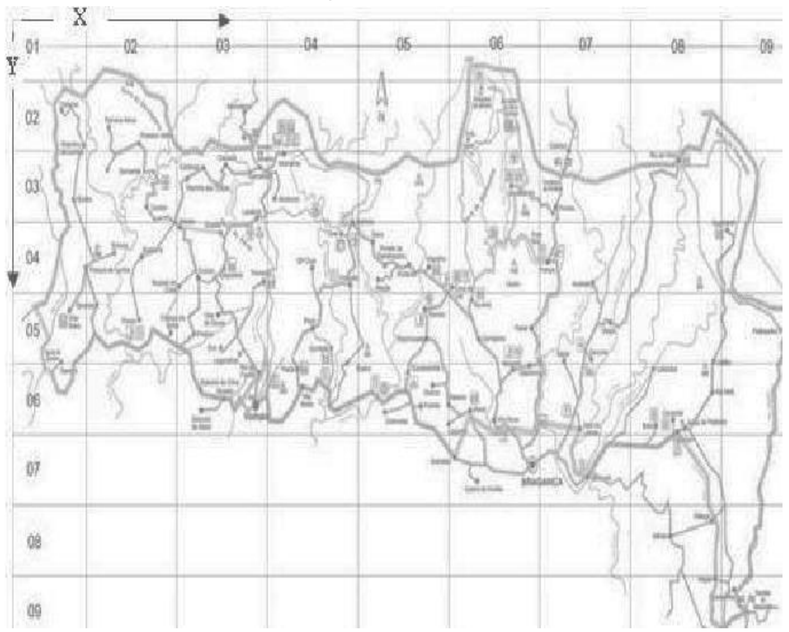
Background

- One major environmental cause of climate change is the rising frequency of forest fires
- Recent forest fires exacerbated by climate change include the 2019-2020 Australia Bushfires, the Amazon Fires of 2019-2020 and crop burning fires in Indonesia
- Goal: Determine which two factors are most important for predicting forest fires

Data

- 517 distinct forest fires occurring between January 2000-December 2003
- Data Features: Location, Time, Natural Weather Features, Burn Area, Canadian Fire Weather Index components
- Slice 1: Full normalized data using Cortez et. al preprocessing methods
- Slice 2: Full dataset without normalized burn area
- Slice 3: Normalized burn area on the days fires occurred
- Slice 4: Sub-slice of Slice 3 of fires between June-October

Grid Based Map of fire data collected

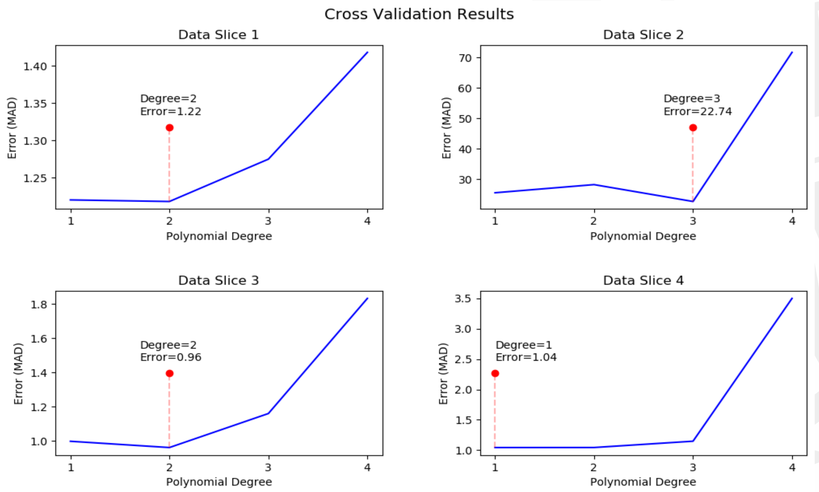


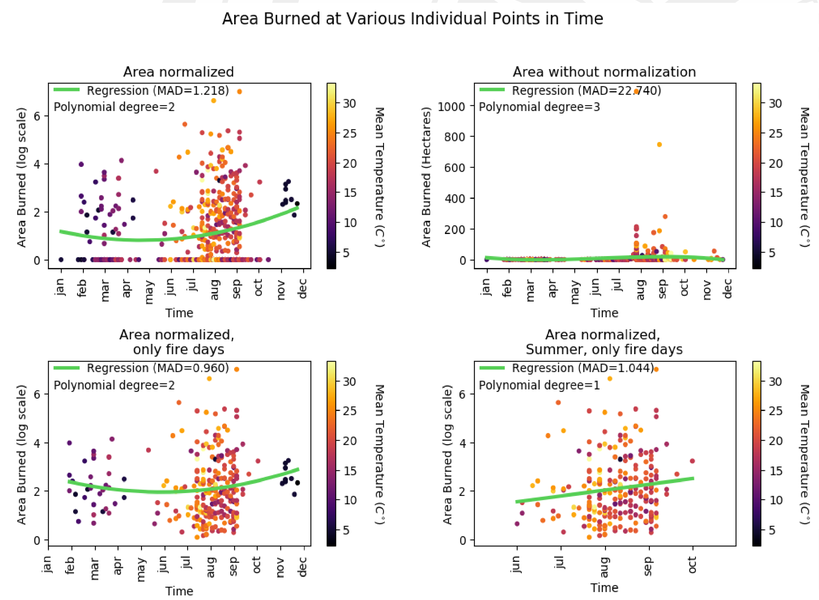
Preprocessing and Modeling

- Natural Weather Features were used for preprocessing
- Burn area normalized with $ln(y + 1)$
- Removed days with no fires for training
- Time attribute standardized into scalar values between 0 and 1

Technique I: Regression

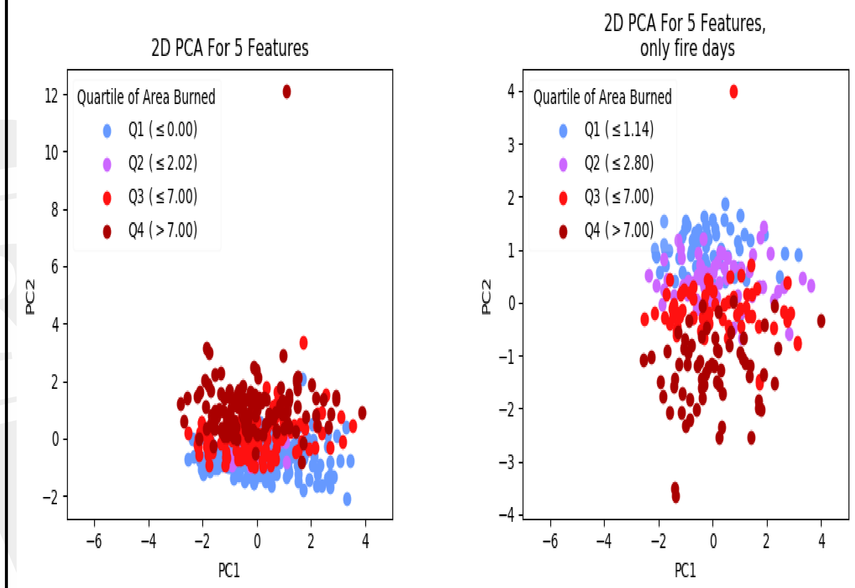
- Polynomial Regression
- Cross-Validation with 70-30 random split
- Error Metric: MAD(Mean Average Deviation)





Technique II: Dimensionality Reduction

- Project 5 features: Temperature, Humidity, Rain and Burn Area to 2 and 3 dimensions
- PCA used on all the data and the days fires occurred
- Revealed a negative correlation between PC 2 and Burn Area Size



CONCLUSIONS

- Collinearity and non-linear relationships in data affected model
- Difficult to determine the importance of different weather features
- Normalizing data for PCA affected relationships between features
- Other techniques such as PC Regression may produce a better model

REFERENCES

[Cortez and Morais, 2007] P. Cortez and A. Morais. A Data Mining Approach to Predict Forest Fires using Meteorological Data. In J. Neves, M. F. Santos and J. Machado Eds., New Trends in Artificial Intelligence, Proceedings of the 13th EPIA 2007 - Portuguese Conference on Artificial Intelligence, December, Guimarães, Portugal, pp. 512-523, 2007. APPIA, ISBN-13 978-989-95618-0-9