Using Analytics for Wildfire Prediction and Management

15.072 Advanced Analytics Edge

Kiran Gite, Pei-Pei Kuo, Jorge Quintanilla, Rebecca Schubertrügmer

11/25/2020

Executive Summary

We use predictive analytics to help decision-makers in California counties with the challenge of allocating resources for wildfire management. We consider data on all wildfires on California from 1993-2015, and daily weather data in that same period. We provide a model that predicts at the time of discovery whether a fire will be big or small so that decision-makers can allocate appropriate resources. We define a "big" fire as burning >2 acres, and a "small" fire as burning ≤2 acres. We produce a boosted trees model that gives a test AUC of 0.6442 and misclassifies only 23% of big fires.

To complement the reactive classification model, we provide a proactive survival analysis. We use survival regression to predict the expected number of days until the next fire in a county and the probability of fire on any given day based on past weather and fire data. Our Cox Proportional Hazard regression gives survival predictions with a 0.64 concordance index. With this model, decision-makers can allocate resources and advise county residents on their behavior based on daily fire risk.

We estimate costs to quantify the benefit of our classification model. Predicting a small fire and getting a big fire will result in a wildfire expanding and causing extra damage, which we estimate will cost $4 million per fire on average. Predicting a big fire and getting a small fire results in wasted resources, which we estimate costs $1 million per fire. Comparing our model with a baseline that predicts that every fire is small, we find that roughly $300 million in fire damage could be saved annually by predicting fire size with some degree of accuracy.

There is substantial evidence that predictive analytics can help to improve wildfire responses and reduce the negative financial, ecological, and health impacts that wildfires have.