Maxwell Snodgrass

ECON 2824 SEC 1020

CA Housefire Report

Introduction

The recent wildfires in California have heightened the importance of fire prevention and the understanding of how specific structural vulnerabilities of buildings contribute to its likelihood of destruction during a fire. Some structures have features that often contribute to an increased risk of severe damage during a fire hazard. Additionally, other factors can contribute to how likely a structure is to be damaged by a fire. An example of this would be the response time of a fire department and the training status of such a department. Although there could be countless factors that would play into predicting the damage a fire would cause to a structure, this research will specifically focus on architectural aspects of the structure as well as fire department quality, response time, and capabilities.

Data

The data used here is a collection of over 100,000 observations of fires that have occurred all over the state of California. Each observation is ranked on a scale from 1-5 explaining the damage severity (figure 1). For the research, we are only trying to predict which features predict the destruction of the structure through a fire hazard. This is because it is quite uncommon for a structure to be partially damaged. It is more likely for it to be unscathed or totally destroyed.

Each observation includes architectural variables of importance, such as structure type, roof construction material, exterior siding material, fence status, and vent screen status. These five variables describe the materials of the structure as well as whether the structures are attached to a combustible fence and the quality of the vent screen. Each observation also declares which fire department dealt with the fire. This variable will capture whether response time/quality of fire department plays a role in determining the damage severity of a fire.

Method

The method that was used to predict the damage severity of a housefire was a Bernoulli Naïve Bayes model. This model attempts to predict whether a house is destroyed based on the architectural and fire department variables specified above. The model only operates on binary variables, so each observation type was changed to a binary variable. Therefore, each observation will be independent of each other in their attempt to predict the damage severity.

Results

The top 15 variables that were the most powerful in predicting damage severity can be found in figure 2. The architectural variables seem to perform much better than the fire department variables. Specifically, the architectural variables of fence status, roof construction material, exterior siding material, and vent screen status. The single most important variable in the model is the lack of fence. This variable contributes the most to predicting whether a structure is destroyed by a fire. Other powerful predictors of whether a structure is destroyed are when the roof is made of asphalt, when the exterior siding is made of wood, and if there are no vents.

The accuracy of the model was 74.00%, meaning that it predicted the test data correctly around three fourths of the time. The confusion matrix in figure 3 shows how the model performed. The sensitivity was calculated to be 84.5%. This means that the model was 85.4% likely to predict correctly structures that were damaged. The specificity was calculated to be 42.1%. This means that the model was 42.1% likely to predict correctly structures that were not damaged. Evidently, the model is better at predicting when a house is likely to be destroyed by a fire than when it is not.

Overall, while the model is not perfect, it offers a relatively accurate prediction of the damage severity of a housefire. The model relies heavily on architectural variables, and therefore it can be concluded that the way a house is built heavily contributes to its susceptibility to damage. Houses with flammable building materials like asphalt and wood, and lack of fire prevention structures like vents are more likely to be damaged than houses that are made of materials like metal, brick, or ignition resistant materials.

Graph Appendix

Figure 1: Damage Severity Geospatial Scatterplot

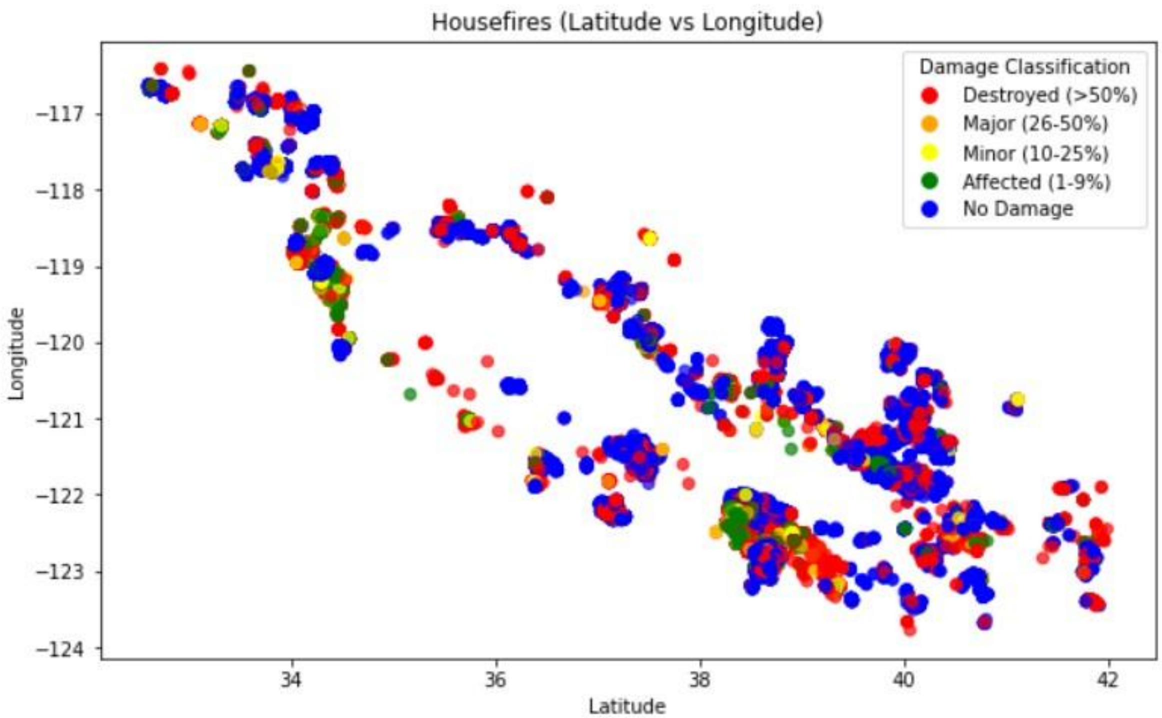

Figure 1.1: Enhanced Heat Map

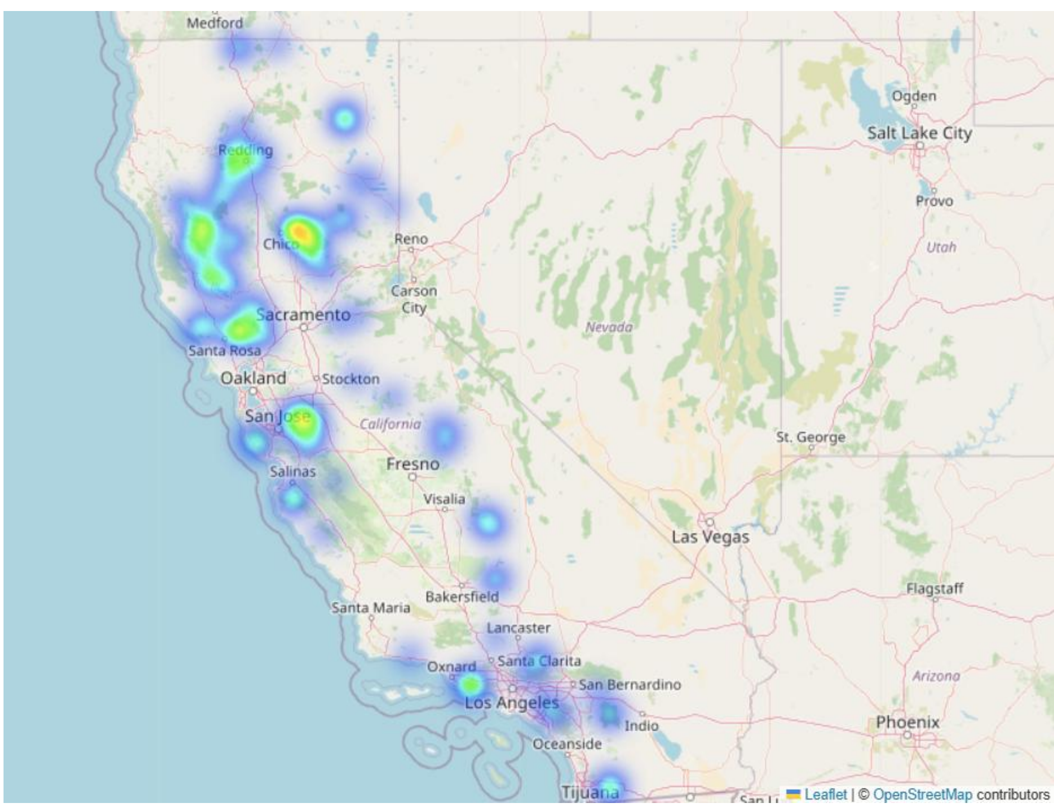

Figure 2: Feature Importance for Naïve Bayes Model

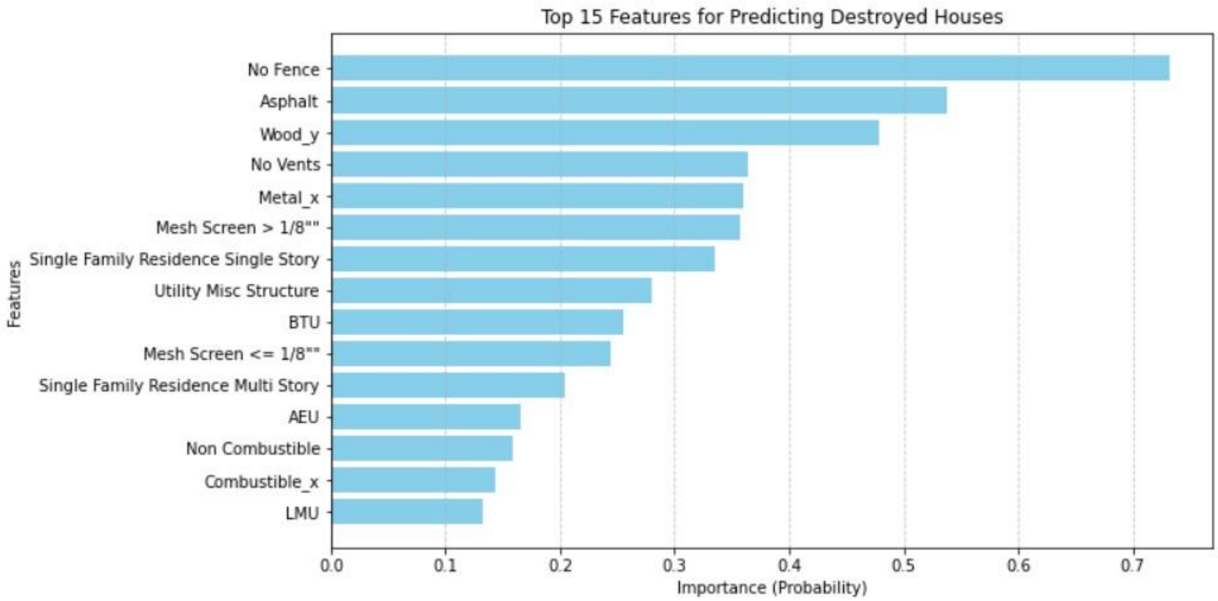

Figure 3: Confusion Matrix for Naïve Bayes Model