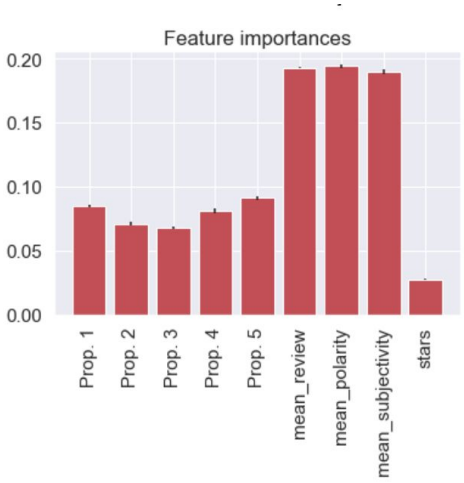
Predicting Business Closure: Visualization

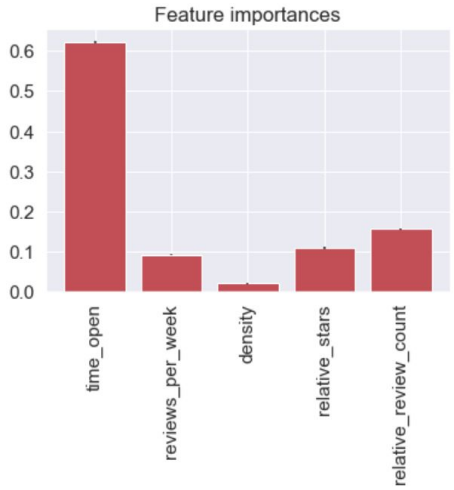
Nathan Timmerman and Micah Thompkins

Machine Learning Model #1



Our first model resulted in about 72% overall accuracy and correctly classified about 55% of closed businesses. Also, a little over 18% of open businesses were classified as closed. As can be seen from the feature importance graph, mean review, polarity, and subjectivity were the key features for the model in predicting whether a business is open or closed.

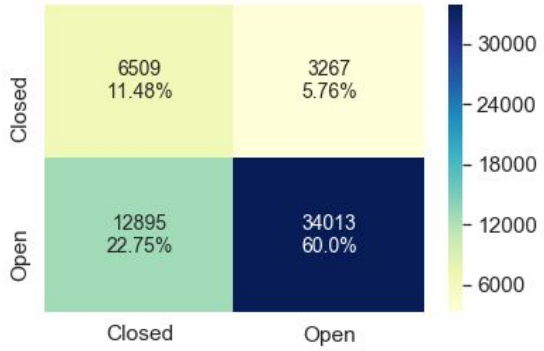
Machine Learning Model #2

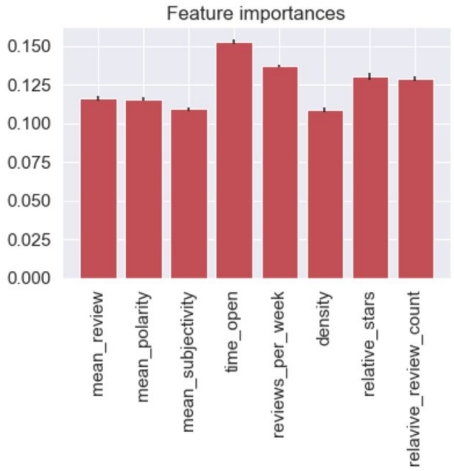


Our second model resulted in about 66% overall accuracy and correctly classified about 66% of closed businesses. A larger proportion of open businesses were classified as closed: a little over 28%. As can be seen from the feature importance graph, time open was far and away the

most important feature, while reviews per week, relative stars, and relative review count were somewhat important.

Machine Learning model #3





Our third model combined the best attributes of our first two: it resulted in an almost 72% overall accuracy while maintaining a 66% accuracy for closed businesses. This model was obtained using the most important features from the previous two, and so, as can be seen in the feature importances graph, each feature is of similar importance. Time open remains the most important.