Mitchell Lee

Coursera – Applied Data Science Capstone

08/13/20

**Problem Description and Background**

For this capstone project, I am pretending that I have been approached by representatives of CAVA—a popular chain of Mediterranean restaurants with 105 locations in the north east, California, Colorado, and Texas—to create a ML model that can predict profitable new locations for the chain. They inform me that they have capital to invest in new locations, but want to be confident that their investments will likely succeed so that they can increase their profitability with minimal failures and losses. They then tell me that they are confident that they have already expanded to every viable zip code in states where they have locations (meaning that zip codes without a CAVA in those states are not viable), so they want to expand to new states and are open to any area of the country as long as the viabilities of proposed locations are supported by a strong model. After discussing my uncertainty as to whether ML can provide strong predictions and what the predictors should be, they agree to fund a preliminary investigation of whether the success of a CAVA in a new location can be predicted based on the types of venues that are nearby. I agree to attempt to develop a ML model to predict which of the most populous zip codes in states across the US that do not already have CAVA would be profitable locations for new CAVA restaurants. Hopefully, this model will serve as a fruitful starting point for the development of better models that will enable CAVA to maximize its return on investment during expansion to new areas in the US.