Leanna Hue, Casey Poon
Professor Hu
3/13/2020
CS 396

Visualization

For our visualization component of our project we included 4 plots. The first three plots show the distribution of average sentiment throughout the three meal times (Breakfast, Lunch, and Dinner) for all Fast Food, American (Traditional), and American (New) restaurants.We used plotly to implement interactive histograms to show the distribution. From our visualization we can see that in terms of center for each cuisine, the modes for average sentiment in increasing order are consistently Fast Food, American (Traditional), and American (New) across all meal times. This makes sense in the context of American cuisine culture; typically Fast Food restaurants such as McDonalds are perceived more negatively as opposed to Traditional American Diners and upscale American New cuisine restaurants.

The ternary plot shows the “distance” to each cuisine from each business based on the features we chose and engineered (average tip sentiment, average star rating, price range, attire, romantic ambience, hipster ambience, classy ambience, and casual ambience). We used the Naive Bayes model of the business data to get the probabilities of each target for each business, and we used plotly to create an interactive ternary plot from this data. From the plot, there does not seem to be any clear distinction or clustering toward businesses of a certain cuisine. This result makes sense since the accuracy score for the Naive Bayes classifier was around 50%, and

there may be overlap or unclear semantic distinctions between the cuisines categories Fast Food, American (New), and American (Traditional).