Capstone Project Proposal 2: Dota 2 Sentiment Analysis

For a company, one of the most important things is understanding their customer base. Without that information, decisions can be made that negatively impact customer satisfaction and end up driving away people from your business. Sentiment analysis is one way that companies learn more about what their customers like or dislike and use that knowledge to fix problems, make better products, reach more customers, and in general, improve their business.

Dota 2 is a well known video game that has been around since 2013 by Valve. The game has had changes constantly over the years, making improvements and implementing new features. In order to stay relevant in such a competitive market, updates and changes like that either need to be carefully considered before, or analyzed after to ensure that your customer base is happy with them. I found a recent survey on subreddit forums specifically for Dota 2. Normally for survey data, the company would just directly survey their clients but I don't have that opportunity which is why I'm using the survey data from the forums. Although there is some bias (data representing only a group of players that also use Reddit), it still gives some insight into at least one group of Valve's customer base.

Using the survey data (hosted on Kaggle and collected in Jan 2019), I will analyze the results and use unsupervised learning to group and label the users based on their responses and other information they provided. This could provide new information about customers that isn't inherently obvious from the survey.

Simple data analysis will provide the basic results from the survey to give actionable information on customer responses. For example, one survey question asks if people play more frequently solo or in a party (with others). Using responses to this and clustering, could provide better information as to why people chose one vs the other. Once finished, this analysis will be provided with a paper and presentation, just as it if was to be delivered as a report to Valve.