**Michael Telahun**

**CSE 632 – Data Mining**

**Dr. Kantardzic**

**9/17/2020**

**Temporal Analysis of Election Related Tweets**

Twitter data has become a hot topic for data mining tasks in several fields. The unstructured information in Twitter data streams makes mining processes and analysis difficult to carry out but often results in a rich creation of knowledge. In this work we manipulate, prepare, and ascertain information from two different days, one in June and one in September, in topics related to the 2020 presidential campaign. As such, several data mining pre-processing techniques were used to digest the information deliberately and thoroughly from both days, creating a temporal analysis that could be applied day by day or over any period in the future. The analysis was performed based on three hypothesis that wanted to be addressed with time gap of three months in mind: (1) *regions or states will be more in favor, in September, of one candidate over the other*, (2) *more popular people will be tweeting about both candidates*, (3) *more conflicting sentiment will exist for both candidates*. The development of these hypothesis and the implementation of munging this information from the tweets is explained in detail within this work. The results show that this information to some degree is not developed enough to make through conclusions; however, given a larger period the picture about these questions would be painted clearer. The result (1) shows that. The result (2) shows that. The result (3) shows that.

Part 2.4 (3.4)

Online Games was removed because it is to broad for the hypothesis in this discussion. The topic itself addresses several fields such as psychology, marketing, finance, politics, among others, but it will not address the aspects of the games specified in the hypothesis section.

**Hypothesis**:

People are discussing the upcoming Expansion of World of Warcraft much more

* World of Warcraft is an MMORPG that is widely among varying age groups around the world. It has been just recently announced that an expansion “Shadowlands” will be releasing in October. There were several forums and discussions online about when it would be released and now that it is confirmed it could suggest more volume.

People are discussing Dota2 more than League of Legends right now

* Dota2 and League of Legends are two of the most popular MOBAs right now. This is just an assumption based on the competitiveness of the industry. At all times one of these two games should be more popular. This might be related to events, updates, feature enhancements, “major news”, etc.

People are discussing different characters which are “overpowered” or “broken” (WOW/DOTA)

* Character updates leads to unbalanced strategies and abilities many games played online. An increase to a character’s potential is often called a “buff” whereas a decrease is referred to as a “nerf”. Buffing and Nerfing occurs frequently and often as the game developers aim to bring as much balance to the game as possible.

1. *Cover page*: Course, title of the project, student name, and date.
2. *Abstract*: Short description of the problem and the solution given in the project applied techniques and obtained results (one page max).

## **Project:** *TITLE*

3.1 **Problem description**. State what topic you have chosen and why you have chosen it.

3.2 **Data Gathering.** Explain how you gathered you own data. Follow instruction in part 1.2

3.3 **Data Exploration**. Exploring provided historical data. Discuss aspects of the dataset following instructions in Part 1.3

3.4 **Hypothesis.** Based on your data exploration result, clearly stated your three hypothesis. Follow instruction in part 1.4.

3.5 **Preprocessing**.

**Software tools**. Explain what tools you used for preprocessing.

**At Least Five preprocessing Techniques**. Explain in detail what techniques is applied and the results. Follow instruction in part 1.5

3.5 **Visualizing patterns:**

**Software tools**. Explain what tools you used for visualization

**Visualizations**. Explain in detail what each visualization mean. Follow instruction in part 3

3.6 **Conclusion**: Do your findings reflect what news says or what your real-life experiences tells you? What can you learn from your findings?