Parlays for Days: Project 3 Proposal

Jacob Barkovitch
Binghamton University
Vestal, New York, USA
jbarkov1@binghamton.edu

Guy Ben-Yishai
Binghamton University
Vestal, New York, USA
gbenyis1@binghamton.edu

Alan Bixby
Binghamton University
Vestal, New York, USA
abixby1@binghamton.edu

Jacob Coddington
Binghamton University
Vestal, New York, USA
jcoddin1@binghamton.edu

Ryan Geary
Binghamton University
Vestal, New York, USA
rgeary1@binghamton.edu

Joseph Lieberman
Binghamton University
Vestal, New York, USA
jliebe12@binghamton.edu

Figure 1: Amazing Meme by Ryan Geary

1 INTRODUCTION

For our final project, our group will continue to investigate whether fan activity prior to an NFL game has a correlation to a game's outcome. In project 2, we created a time series of each NFL game, tracking the polarity and post frequency, but our analysis was qualitative.

2 RESEARCH QUESTION

Our chosen research question will be the "Wisdom of the Crowd" question from Project 2 (see proposal/report). We will investigate whether user sentiment prior to a game has a correlation to a game's outcome. To do this, we will use a combination of moving average and change point analysis to validate our findings. We can also compare sentiment to the odds to see whether users beat the odds.

3 PREVIOUS ANALYSES

We will incorporate Reddit and Twitter Sentiment over a time period (Figures 3/4 in Project 2 Proposal) as well as by NFL team over a time period. We will also incorporate a way to get the odds for a specific game or for a team over a time range as well.

4 TOOLS FOR RESEARCH

The tools we intend to use for research are:

- Python
- Flask
- Numpy/Pandas/Matplotlib/Seaborn
- REST API in ExpressJS
- NextJS + React + D3.js
- Mongoose (MongoDB connection)
- Maybe Docker for easy setup

5 OUR TOOL

Our interactive tool will be accessible on localhost, and we will provide instructions for users to set up and run the tool on their own machines. To ensure the success of our project, we will need to carefully plan and execute each stage of the process. This will include:

5.1 Data Collection

We will need to gather data on user sentiment and the frequency of posts on Reddit for each game we are analyzing. This will involve using web scraping and natural language processing techniques to extract and process the data.

5.2 Data Analysis

once we have collected the data, we will need to perform a variety of analyses to identify trends and relationships in the data. This will include using moving average and change point analysis to identify correlations between user sentiment and game outcomes, as well as using 2-vector analysis to identify key nouns being discussed in relation to the game. Tool Development: based on the results of our analyses, we will need to develop an interactive tool that allows users to explore and visualize our findings. This will involve using web development tools such as React and D3.js to create a dynamic and user-friendly interface.

5.3 Testing and Validation

Before we release our tool to the public, we will need to carefully test and validate our findings to ensure that they are accurate and reliable. This will involve conducting additional analyses and collecting feedback from users to identify potential issues and make any necessary improvements.

6 CONCLUSION

Overall, we believe that our project has the potential to provide valuable insights into the relationship between user sentiment and game outcomes. By using advanced data analysis techniques and interactive visualizations, we hope to make our findings accessible and understandable to a wide audience. In conclusion, our project aims to answer the research question: does user sentiment prior to a game have a correlation to a game's outcome? To do this, we will use a combination of moving average and change point analysis, as well as 2-vector analysis, to identify trends and relationships in the data. We will also develop an interactive tool using web development tools such as React and D3.js, which will allow users to explore and visualize our findings. By conducting thorough data collection and analysis, and creating an engaging and user-friendly tool, we hope to provide valuable insights into the relationship between user sentiment and game outcomes.