

LigaMX Forecaster

Final project for TC3041 - Advanced Databases, where I try to forecast the winner of Liga MX championship by using a soccer feed and analyzing sport experts tweets.

Author

Juan Carlos Sanchez Cruz - A01631462

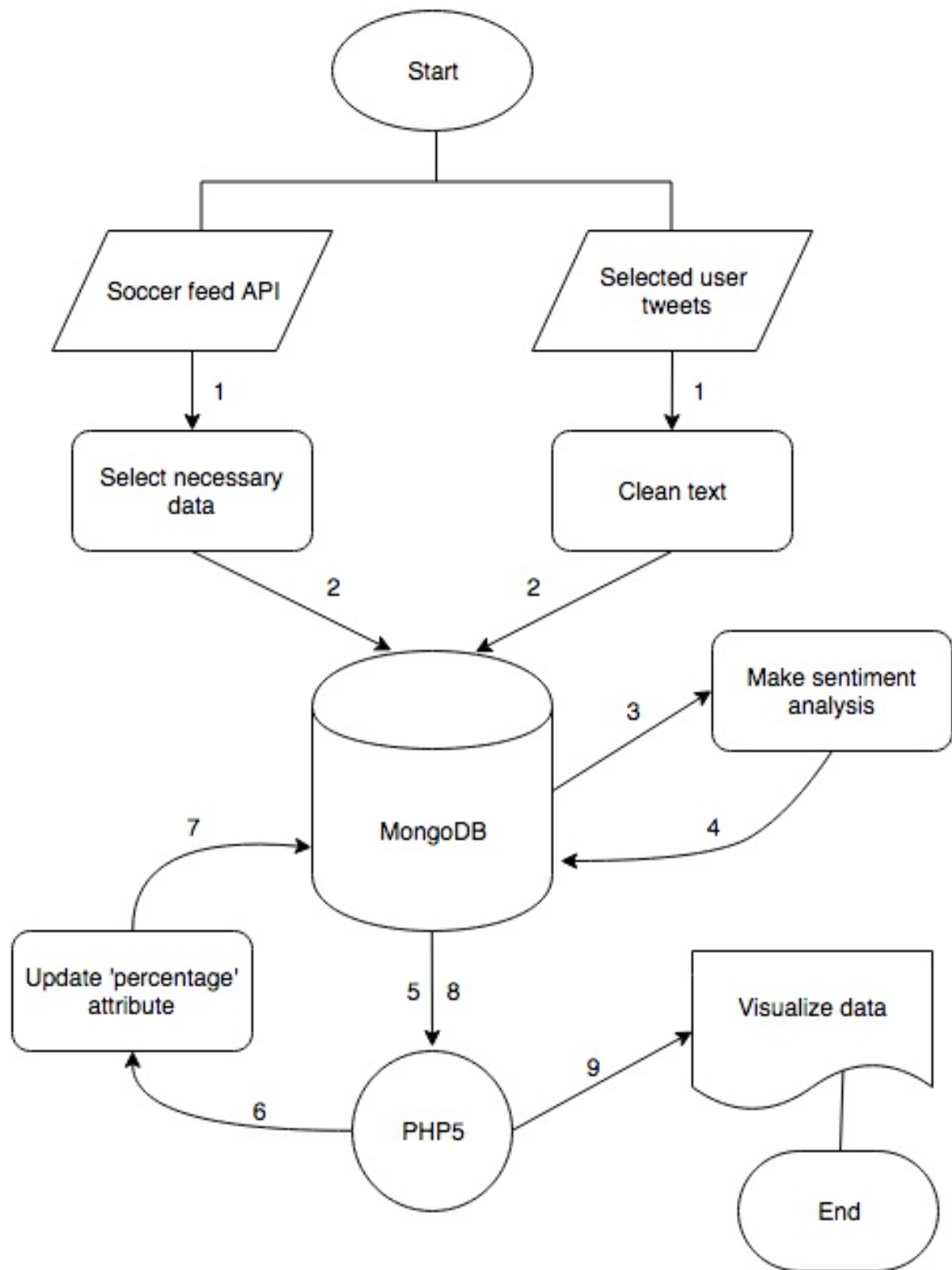
Motivation

This project seeks to solve a simple question, who will be the winning the liga mx season 2018? In Mexico football is almost treated with political language and it is no coincidence that there are so many sports analysis programs. This sport is a round-the-media business. The live broadcasts of the matches are very expensive due to the number of people summoned. After every football game, the conversation turns to interpret the events and plays. But one of the things that draws my attention is the speculation that is created before, and that the Mexican League is very difficult to predict. Bets never fully favor one side, and there are many factors that can affect a team's performance.

What I'm looking for in this project is to make a prediction that uses historical data about the teams in the playoff and complement it with the positive or negative opinions on tweets of a selected group of sports analysts.

System Architecture

The data flows in the system as follows:



The numbers correspond to the chronological order in which the processes are developed. In the case of the two sources of data mining, this could be done in parallel.

Data Mining

Soccer Feed

Sport Experts Tweets

Tangible Data

Validity

Accuracy

Completeness

Consistency

Uniformity

Data Visualization

DIKW Pyramid

Conclusions

Constraints

Run this Project

Folder Path

Built With

- MongoDB
- PHP5
- Tweepy

- Fantasy Data
- Bootstrap Material Design

Acknowledgments

- To all sports analysts whose twitter accounts were used for this project.
- Rodolfo Rubén Alvarez, teacher of TC3041 course at ITESM.