New York University Machine Learning and Computational Statistics

Spring 2018 Project Report

**PREDICTING RESULT OF NCAA FOOTBALL GAMES**

|  |  |
| --- | --- |
| TEJAL LOTLIKAR  New York University  Courant Institute of Mathematical Science  [tl2482@nyu.edu](mailto:tl2482@nyu.edu) | DIWAKAR PALIWAL  New York University  Courant Institute of Mathematical Science  [dp2757@nyu.edu](mailto:dp2757@nyu.edu) |
|  |

KUMAR MEHTA

New York University

Courant Institute of Mathematical Science

[kjm627@nyu.edu](mailto:kjm627@nyu.edu)

**Advisor:**

DAVID FROHARDT-LANE

**Abstract**

We present a sports prediction model that predicts the winning team and average margin of victory in NCAA football games. Our goal is to analyse the performance all games played by both the teams playing current match and predict the winning probability of each team in the current game using those statistics. We consider various data features (like points scored by a team against all teams, goal attempts, penalty etc) of each team that primarily influence the result of a game and use them in various data models to improve the accuracy of prediction. Along with the winning probability, we also predict the average margin victory in the current game by using different models of the data using appropriate features. We found which of the features from entire dataset are important as instruments using different feature selection algorithms.

We use cbstats.com dataset containing information on College Football Statistics from 2005 to 2013 for this analysis and prediction model.

Dataset Link:

<https://old.datahub.io/dataset/college-football-statistics-2005-2013>

Github Link:

<https://github.com/dwkr/mlcs_project_2018>