A Logistic Regression/Markov Chain Model for the NFL

Aryan Mittal and Joel Sokol

Abstract. We apply a modified version of the logistic regression/Markov chain (LRMC) model developed by Kvam and Sokol to generate improved National Football League (NFL) team power rankings. We calibrate the logistic regression model using several different metrics of evaluating team performance, including simple point difference, Pythagorean point ratio, and multiple variants of "game control" (area under the team's in-game win probability curve). Through analysis identical to that by Kvam and Sokol, we determine that home field advantage in the NFL is worth about 5 points. We then regress the steady state values of the Markov chain with a boolean variable representing home field advantage against the final point spreads of the regular season. Our final model outperforms both the consensus spread and other top computer ratings in predicting the outcomes of NFL playoff games (63% accuracy between the 2005 and 2022 seasons).

School of Industrial and Systems Engineering, Georgia Institute of Technology