Machine Learning Project Proposal

Match Outcome Predictor

Michael Gagliardi

With sports betting recently becoming legal in Massachusetts, the demand for models that can accurately predict the outcome of sports events is at an all-high in this area.
This project plans on using a logistical regression model to help predict whether or not a team will win its upcoming game using various independent variables of both the team's and the oppositions current performance. In doing so, the model will be able to accurately predict which team may end up winning, and in turn give some lucky bettors an upper hand in deciding a winner. Due to recent advancements in tracking technology in the world of sports, there are a large number of datasets out there that accurately record a wide variety of statistics in the majority of major sports. This can potentially allow this model to be used across multiple sports with some minor adjustments.

This model will most likely be using a supervised learning approach, implementing both logistical regression and support vector machines. There are a large number of datasets available on Kaggle and an English Premier League dataset appears to be the most viable to use at this time. The model will use python with the scikit-learn library. Additional background research will be required to aid in the implementation of SVMs and logistical regression. In addition to that, more knowledge is required in order to accurately tune the parameters.

References:
[1] https://www.kaggle.com/code/saife245/football-match-prediction/notebook#Applying-the-Logistic-Regression