NCAA March Madness Project Proposal

Problem Statement: Can I build a machine learning model that can accurately predict game winners, or even the NCAA Tournament winner using historical NCAA Basketball data sets?

Client: Baylor University, a top 10 college basketball program, seeks to leverage historical NCAA Basketball data to enhance game strategy and recruiting efforts.

Data: Utilizing a dataset sourced from Google Cloud containing game information from the 2013/2014 season to the 2017/2018 season, the project aims to analyze game outcomes and trends.

https://docs.google.com/spreadsheets/d/1yE8X8tZFfKT9jaowTQiEQ6N154N5b7BaATXEqV3rWoQ/edit#gid=1937406620

Criteria for Success: Success will be determined by identifying new insights from the data that contribute to predicting game winners with sufficient accuracy. While acknowledging inherent limitations, achieving a model that provides actionable insights will be considered successful.

Scope: The project encompasses all college basketball data, with a focus on regular season games to inform predictions for tournament games. Training the model with data from the aforementioned seasons will be followed by validation using current season data to assess predictive capabilities.

Constraints: Limited availability of online data poses a challenge, with only four seasons of well-defined game data accessible. While final score data beyond the selected seasons is available, the emphasis remains on utilizing more detailed game-specific data for model development.

Stakeholders: The project stakeholders include Baylor basketball head coach Scott Drew and Athletic Director Mack Rhoades, who will utilize the findings to inform strategic decisions and player recruitment.

Deliverables: The primary deliverable will be a presentation deck summarizing key findings, supported by documentation detailing the methodology and code used. Additionally, a deployable model capable of predicting game outcomes for future seasons will be provided.