ISYE 6501

Introduction to Analytics Modeling

Course Project

Introduction

As a fan of tennis, I always wonder how the tennis world adapts the latest technology to make their decision making more efficient, in terms of how the players know their strengths and weakness in every match, how the coach makes better strategies, how the sponsors decide which player to support, and how each championship host improve their profits, by having to consume so much versatile information for better result. When I am reading these case studies to search for a topic for my project, I came across this article about how the Wimbledon collaborated with IMB team, I decide to dig into this topic.

Wimbledon is one of the most prestigious tennis championships in the world. Coping with the latest technology, Wimbledon partner up with IBM iX to develop a unique approach to make the decision making easier for the advertising team. The two key questions the Wimbledon and IBM focus on are how to engage fans more deeply with Wimbledon and tennis on an ongoing basis, and how to encourage fans to follow up-and-coming players who weren't famous yet. From these, the teams conceptualized an index that shows how players perform and how they interact with their fans. The players performance and other factors are analyzed to generate player rankings, prediction for likely winners, potential matches combinations, and alerts for the matches.

Assumptions

To maximize the fan engagement, we could make assumptions based on the project:

1. The more precise the fan groups are divided based on their similarity, the better the company could target them for what they really want.
2. The more the fan react to players, the better sells the company could make.

Models would be used to test these hypotheses to determine whether they are correct or not.

Project Goals

1. Be able to create performance profile, calculate rankings, and predict potential matches of the players.
2. Analyze fan's behavior in order to encourage fan engagement based on the player's performance.
3. Analyze fan's behaviors, combine with player's performance, and together, come up with strategies to encourage interaction between fan and rising players

Analytics Process

Here, seven analytics models will be used:

1. Model 1: ANOVA model for creating performance profile

   Given{Data of all the matches of all the players collected in the past 10 years in Wimbledon}

   Use{ANOVA model}

   To{Analyze the match performance of tennis players in Wimbledon}

   The aim of this model is to use the past result of each match at the Wimbledon in the past ten years to build performance profile of each player, in order to make better decision in the prediction in the later analysis.
2. Model 2: Markov Chain Model for Rankings

   Given{List of results of all the matches between a set of players(Result from model 1)}

   Use{Markov Chain model}

   To{Output a matrix with transitional probability based on the match results}

   The goal of this model is to create a matrix to calculate the players performance in order to make the ranking.
3. Model 3:Markov Chain model to predict potential matches

   Given{List of the players and their individual performance at each match(Result from model 1)}

   Use{Markov Chain}

   To{Predict potential match}

The goal of this model is to create a mapping from a player and player tuple to a vector for potential matches prediction.

4. Model 4: Logistic Regression for predicting rising players

   Given{List of the performance profile of those that aren't famous yet}

   Use{Logistic Regression}

   To{Predict the increase of performance of the players that was not famous}

5. Model 5: Clustering prediction for fan engagement

   Given{List of fans, and their similarity and closeness to each other, their reaction to the players at each season, and their purchasing behaviors(Result from model 1-3)}

   Use{Clustering prediction}

   To{Divide audience based on their behaviors}

   The goal of this model is to apply clustering analysis to find similarities of fan behavior and thus divide fan in groups based on how likely they will engage more with the player and match, even their purchasing behavior such as (ways of buying tickets, merchandising, social media, etc). The clustering is a good way to categorize types of clients based on the similarity of their behaviors. Clustering makes it clear to see how a certain type of customer react with certain type of advertisement. Therefore, the advertising team would know how to advertise to the audience based on their interests.

6. Model 6: Optimization to predict how likely the fan's engagement would be with each factors(Players, rising players, and purchasing behaviors)

   Given{Clusters of fans, player's performance, rising player's performance(Result from model 4 & 5)}

   Use{Optimization}

   To{Predict fan's engagement with factors}

   The goal of this model is to use result from model 4(rising players) and model 5(clusters of fans) to maximize the fan's engagement with the rising players. Here, for each question(engagement based on the player's performance could be encourage from advertising certain content based on the fan's behavior, and if the way the advertisement of certain content with certain types of fans could have impact on the engagement

between the fans and rising player) and for each combination(fan vs Players, Fan vs rising players), would be tested. For instance, type 1 of clusters of fan and player number 1, type 2 of clusters of fan and player number 1, type 3 of clusters of fan and player number 1 and so on. Then we would know what kind of advertisement the company should target for the certain player.

7. Model 7: Bayesian modeling to testing the hypothesis

   Given{Result from model 5 and model 6}

   Use{Bayesian}

   To{Test if the hypothesis of model 5 and model 6 are valid}

   Model 7 is an optional model that can be used to test if the two questions, engagement based on the player's performance could be encourage from advertising certain content based on the fan's behavior, and if the way the advertisement of certain content with certain types of fans could have impact on the engagement between the fans and rising players, are legitimate. Here, different combinations are set up and compared to see if the way the advertisement an impact on how much the increased level of fans engagement could really be.

Summary

The seven models above answer all the question that the Wimbledon would need. These includes, using the past data to create profile of players' performance, the ranking, and identifying the players that are not famous yet but has potential to be rising stars. Second, the model categorizes fans into groups, based on the player's performance, the potential matches combination, their interests in particular plays, and their purchasing behaviors. Then each cluster would be analyzed to see what they are interested the most. Thus, the advertising team of Wimbledon would know what the fan likes and is able to show the correct advertisement to each group. Therefore, it could maximize the engagement of the fan.

Using analytics approach in marketing and adverting is very common nowadays. Any big enterprise would collaborate with the analytics teams to understand consumers behavior, which involves the qualitative and quantitative analysis, in order to maximize their profit. However, in

the real world, creating a solution to the goals that were set is a real challenge. There will be more factors to considerate. Since the IBM does not provide their solutions in this case study article, my proposal of using analytics models learned in class is a good example of solving these problems.

Reference

https://www.ibm.com/case-studies/all-england-lawn-tennis-club-ibm-ix/