May 01, 2019

# UTR InternShip Analysis Report

# Introduction

The goal of this analysis is to rank the players from best to worst based upon the dataset provided. The very first thing I did was to go through the understanding of the UTR algorithm on the website and understand every aspect of the Tennis Ranking system of UTR.

# ALGORITHM

After having a brief understanding of the algorithm, I studied the dataset provided. In terms of Ranking, the attribute that must be considered is how experienced is the player with regards to the number of games he/she played. So, in order to analyze the data, I categorized the players into three categories:

* Beginner
* Intermediate
* Advanced

A beginner player is the one who has played a total matched of 15 or less.

An intermediate player is a player with total matches more than the beginner player but less than 30 matches.

A player can be considered as an advanced player if he/she has played at least 30 matches in the entire dataset.

The implementation of this algorithm can be found on my Github as well:

<https://github.com/sahilsood/UTRInternship>

# BASE SCORE

It is important that the ranking of the players take place with the players of the same category i.e. a beginner player should not be ranked higher than the advanced player and his ranking should be considered only with comparison to the performance of other beginner players. Therefore, for this purpose I introduced Base Score.

Base Score is the minimum score a player gets once he reaches a certain level. For instance:

If beginner player with matches less than or equal to 15 will have a Base Score of 15.

An intermediate player will have a Base Score of 50. Whereas, an advanced player will have a Base Score of 100. The aim of introducing the base score in the algorithm is to distinguish the three types of player in terms of ranking. An advanced player will always be ranked better than the intermediate player and a beginner player.

# Player points

Player points will be calculated based upon the number of wins and the number of losses. If a player wins, he/she will get plus points added to his base score. And similarly, if a player loss he/she will have a deduction of points from his/her base score.

**How do we calculate the Points gained or lost?**

The better approach to calculate how we assign the points to players should be based upon their experience. If a beginner player with less experience beats an advanced player, he/she should gain double points. However, if the players with same level of experience play with each other, they will be competing for just a standard point.

# Rankings

Based upon the Player Points, the algorithm calculates the Rankings of the players in the dataset. There are a total of 1052 players ranked from best to worst. The result is stored in a csv file named ‘Rank.csv’.

# tie breaker

There will be a time when certain players will be having same points based upon their matches. To overcome this, we have to take account of the margin of victory in terms of their win sets. The player with higher margin of victory gets a higher rank than a player with less margin.

Note- This approach is not implemented to get the rankings but is definitely a better way to eliminate the tie between two players.

# Data credibility

It is important to have a credible data which can be trusted and verified with the actual dataset and analysis. So, I imported the dataset and Visualized it using Tableau. I matched my results with the Top Players visualized in Tableau graphs.

The Visualization of the internship can be found on my public Tableau Dashboard at <https://public.tableau.com/profile/sahil.sood#!/>

Sahil Sood

571-494-0580

[ssood3@uncc.edu](mailto:ssood3@uncc.edu)

<https://sahilsood.xyz>