Project Proposal

Title: Data Mining system for Selecting a winning Cricket Team

Aim of the proposed project:

The aim of this project is to apply Data Mining techniques to analyse the available player data & the performance of their respective teams in ODI (One Day International - 50 overs) cricket matches to come up with a method to assemble a winning team. And to predict the winnability of all the assembled teams.

Background/Short Description of the project (Be concise and comprehensive):

Sabermetrics is the study of the game baseball, through statistical analysis. This technique has led the team Oakland Athletics (American professional baseball team) to 20 consecutive wins. Which was a world record. After witnessing this great victory, cricket as well as many other sports tried to adopt this approach.

Several analyses have been done on using sabermetrics approaches to cricket. They have been successful mostly in IPL (Indian Premier League - 20 overs) and twenty20 matches. Most of these efforts have been focused on identifying different performance metrics based on the domain knowledge of cricket. But unfortunately, this has not succeeded in some cases either due to the lack of impact from the performance metrics. Therefore under this project our main focus is on ODI matches. We aim to analyse the player data and their team wise performance to assemble a winning team for a ODI match while minimizing the use of domain knowledge of cricket to decide the winning team.

Expected Outcomes (In point form state what is planned to achieve):

This semester:

- Identify what the past researches have done and not done under this topic
- Finding datasets with a satisfactory amount of data (ODI match data)
- Cleaning the dataset and converting them to a format that can be used in a machine learning analysis

Next semester:

- Identifying best attributes that can describe a player who can led the ODI match to win
- Identifying best metrics that describes overall team performance.
- According to the best attributes obtained from players, analysing their team performance.

References

1. P. Somaskandhan, G. Wijesinghe, L. B. Wijegunawardana, A. Bandaranayake, and S. Deegalla, “Identifying the optimal set of attributes that impose high impact on the end results of a cricket match using machine learning,” 2017 IEEE Int. Conf. Ind. Inf. Syst., pp. 1–6, 2017.
2. N. Siva, A. Perera, G. Thejawansha, A. Bandaranayake, and S. Deegalla, “Performance Analysis of a Selected Cricket Team Using Data Mining and Machine Learning” Final year project research paper by E13 batch
3. https://en.wikipedia.org/wiki/Sabermetrics

Peer Review

(Must evaluate the proposal providing suggestions for improvement)