Data Mining System for Selecting a Winning Cricket Team

GROUP 2

E/15/119
E/15/202
E/15/208

Motivation

- Sabermetrics in baseball (moneyball)
- Oakland Athletics
  - 20 consecutive wins

Why Cricket ??

- Popularity of cricket
- Bets on cricket

ST. MORITZ
ICE CRICKET

Cricket Sabermetrics?

Will it be successful?

Our Focus

- Predict ODI
- What are ODI matches
- How winners are selected

Problem Definition

- Large percentage of bets
- Unsuccessful of previous methods
- Continuous losses of teams
- Time period of match details
- Player performance
    - Batting
    - Bowling

Solution

- Different sets of performance attributes
  - Player wise
    - Batting
    - Bowling
    - Fielding
    - Allrounder
  - Team wise
    - Composition based
- Recent match details

Related Work

Includes..

- Research papers - ODI(One Day International),Twenty20, IPL
  - Player performance
  - Team performance
  - Outcome prediction
- Different technologies

Title/Authors	Match type and Time periods	Player performance Attributes(A)	Team performance Attributes(B)	ML techniques and Accuracy	Technologies used
CricAI: A classification based tool to predict the outcome in ODI cricket (by Kaluarachchi, Amal Chaminda,Varde, Aparna S)(2010)	ODI From 1971		Team, Opponent team, Home/Away, Day/Night, Toss, Bat 1st	Naïve Bayes(best model), Decision Tree Classifiers using C4.5, Bagging and Boosting	JAVA, MYSQL
Applications of modern classification techniques to predict the outcome of ODI Cricket (by Neeraj Pathak)(2016)	ODI 2001-2015		Toss outcome, Day/Night Effect, Home game advantage, Bat First	Naïve Bayes(best model), SVM, Random Forest	R
Predicting the Outcome of ODI Cricket Matches: A Team Composition Base Approach (by Madan Gopal Jhanwar) (2016)	ODI 2010-2014	No. of Matches played, Batting innings, Batting average, No. of centuries, No. of fifties, Bowling innings, wickets taken, >= 5 wickets, bowling average, bowling economy	Toss, venue, strength of team A against team B	SVM, Random Forests, Logistic Regression, Decision Trees, kNN (best model - 0.71 accurate)	

Title/Authors	Match type and Time periods	Player performance Attributes(A)	Team performance Attributes(B)	ML techniques and Accuracy	Technologies used
Identifying the optimal set of attributes that impose high impact on the end results of a cricket match using machine learning (by Somaskandhan Pranavan) (2018)	IPL 2008-2016		High Individual Wickets, Number Of Bowled-Deliveries, No of thirties, Total wickets, Wickets in powerplay, runs in death overs, dots in middle overs, No of 4s, Singles in middle overs	Extra tree, Naive Bayes, SVM SVM(Best) -81%	Java Python DBMS
Performance Analysis of a Selected Cricket Team Using Data Mining and Machine Learning(by Nilucshan Siva)	IPL, T20 2006-2017	Average, Strike rate, Evaluation score	-City,Venue, Match type, outcomes, No of overs, Player of the match, Opposition, Toss winner, Toss decision, Winner	B - Logistic Regression, Decision Tree, SVM, Naive Bayes,KNN, Random Forest SVM(Best) -70.21%	Java R
Novel Method for Cricket Match Outcome Prediction using Data Mining Techniques (by S.A.D.P Subasingha)(2019)	ODI	wicket no, partnership runs	toss effect, ground condition, day-night effect, opponent, total runs	A&B - Naïve Bayes,	HTML Agility library, WEKA tool

Related work

Similarities

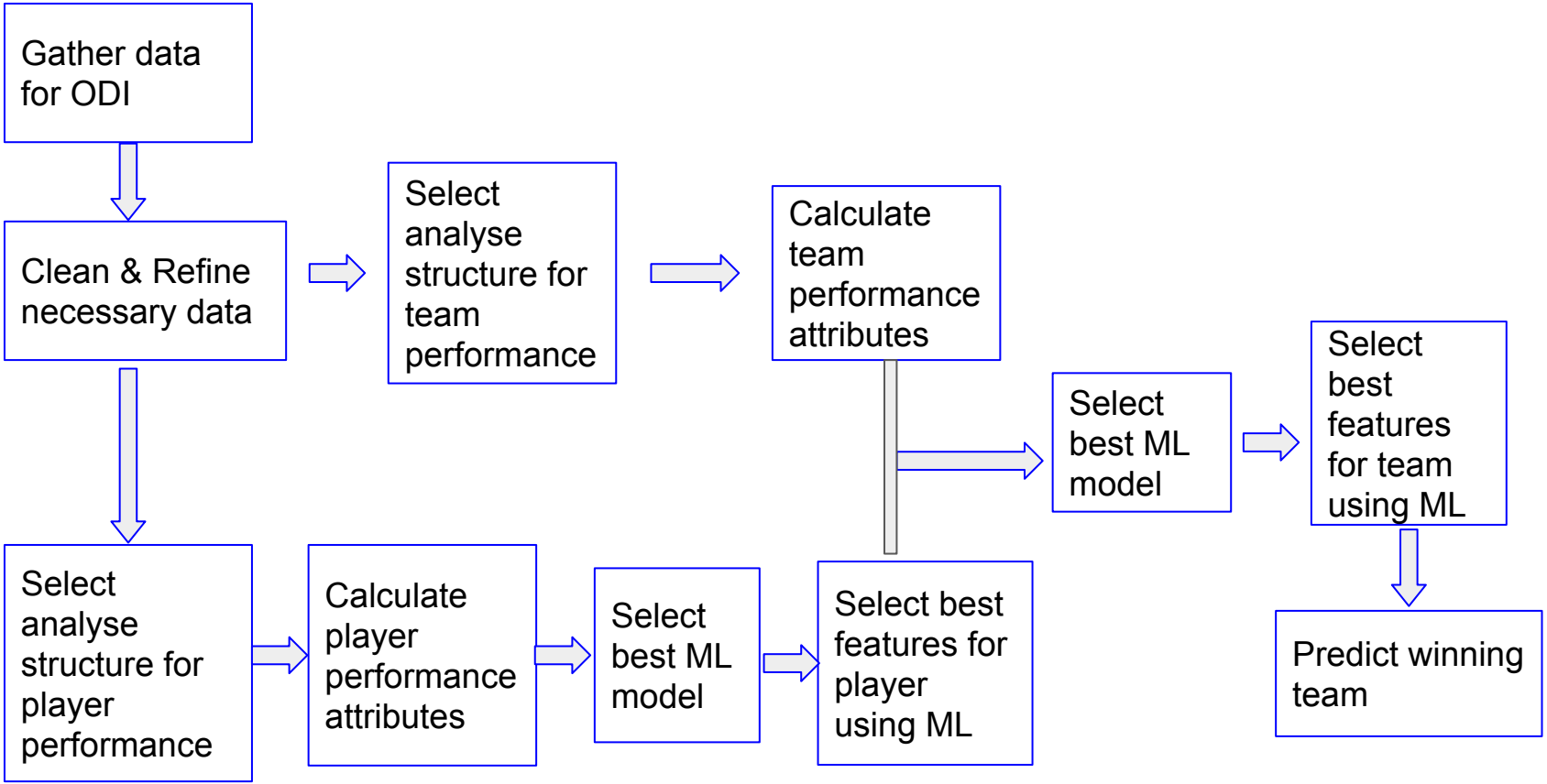
- Some of the team performance attributes
- Some of the player performance attributes
- Similar technologies

Related Work

Dissimilarities

- Player performance attributes
- Team performance attributes
- Time period of matches (from 2020 to past)

Proposed Methodology


Gather data for ODI
Clean & Refine necessary data
Select analyse structure for team performance
Calculate team performance attributes
Select best ML model
Select best features for team using ML
Predict winning team
Select analyse structure for player performance
Calculate player performance attributes
Select best ML model
Select best features for player using ML

Technologies To Be Used

Milestones

- 3 milestones in Predicting the winning team
  - Player wise performance
    - Best attributes to describe each individual player
  - Team wise performance
    - Best attributes to describe the team
  - Winning
    - Predicting the outcome considering team performance and other factors

THANK YOU !!

Q.
A.