**A Player Performance Prediction Application Using Machine Learning to Revive Cricket in Kenya**

Kenya a cricket playing country that had all the eyes of best cricket playing countries who thought Kenya was the best emerging team ever seen by world cricket. In the 1980’s to the early 2000’s Kenya had seen such a big hike of growth in the game, such that the ICC (International Cricket Council) had so much faith in the country that they chose the country to host a World cup trophy in 1996 where Kenya managed to turn out as runners up. Also, a World Cup knock out tournament in 2000 which is a big thing in cricket as the ICC wanted Kenya to gain for attraction and funds to develop the game in the country. Moreover, in the 2003 which is Kenya’s dream year till date where Kenya managed to qualify for the final events which was never done before by any associate nation. In this process Kenya managed to win the strongest teams of that time that is to West Indies(which shock the whole world), India and Sri-Lanka. However, after 2003 is when Kenya started to see the decline that is lost quite a few consecutive games till 2011 where ICC then had to snatch away the ODI status of the country which made Kenya to pass through qualifiers which till date they have not even been able to qualify to get a spot in the world cup.

The major reasons for the decline are, corruption, poor management, and outdated technology. The impact of corruption to the nation have been such disastrous such that it has ruined the game of deserving players and its value to the whole world. Deserving players could not get a chance while the coaches took bribe from players who were not even well performing to get into the team. Moreover, all the funds received from ICC for development in Kenya were misused and the players didn’t even get their salary. This also led to deserving players being forced to work in other places in order to feed themselves and their family, this also led some players to go on strike which also was a harm to the reputation of the game in the country. The misuse of the money (eaten up by the board) also is a case of poor management as this also led to up rest and strikes of players as money was not managed well as they also had poor infrastructure and support to play with.

With outdated technology Kenya is losing out a lot since the main usage of IT and technology in cricket is to analyze players and their performances and to collect and store data. This takes cricket Kenya a step behind than the other countries with advanced technology as they can analyze their players well and can train them on any weaknesses while the Kenyan players are then left with no option but to play with personal instincts which may not be always prudent.

The main objective of the proposed web application is to help revive cricket in Kenya.

The goals of the proposed system are;

* To minimize the effect from corruption.
* To centralize the data collected.
* To enable easy and accurate scoring of matches
* To predict the performance of players.
* To improve the selection of players in the best.

The proposed web application will do everything from scoring to predictions. The scoring done will be ball to ball meaning runs will be tallied on ball-to-ball basis that will be added to total score, batsman individual score and bowlers for the runs given out. And for the batman it scores which side of the ground they are getting out to mostly and for bowlers the side they are hit the most. Then after the season is over batsman scores per match are tallied up and total runs added up and a ranking leaderboard will be generated with the highest run scorer as number 1 and the rest follow. The same for the bowlers the wickets per match will be accumulative and will be added at the end of the season with the highest wicket taker being ranked as first and the order follows. The other good thing with this system will be that Cricket Kenya will have direct access to this data and they can store it wherever they want and thus this makes it a centralized storage platform for them that they can use to avail the data whenever they want.

Once the leaderboards are generated for both the batters and the bowlers the top 7 batsman and the top 7 bowlers and top 3 wicket keepers in the batsman leaderboards to make as 17-man squad. From the 17 players an individual performance prediction for upcoming matches will be done. The prediction will take 3 values that is high performance, average performance and poor performance. Those players that have predicted to have high performance will be selected for the first eleven and if these players with high performance result are exhausted and there are not 11 players selected then the remaining players will be those with a result of average performance. For the batsman the specific metrics used for prediction will be previous score, average, venue, and opponent team for the bowler the metric will be average, strike rate, number of wickets in previous match and the venue. The playing eleven will then be selected that will get the best output for cricket Kenya. This will also motivate the players since they will also just like the system think that they will perform well and then go to the ground with a positive intent which is beneficial for both the team and the players.

The machine learning algorithm that will be used will be random forest that is a classification task algorithm. Random forest is an algorithm that combines the prediction form multiple decision trees by averaging the predictions. The reason for choosing random forest is because first of all it is accurate and reliable since the predictions it makes are based on an average function that deals with all the outliers making the outcome more predictable. Moreover, Random forests have a lot of built in libraries and frameworks that makes the algorithm easy to use. In order to train data using Random Forest first the data needs to be collected and consistent to make its outcome the most appealing. The data set then needs to be split into a training set and a testing set. Then feeding and evaluating the data basted on the testing set if the prediction is satisfying it can then be used to predict the performances and if the outcome us not satisfying then the parameters can be changed and tested again until the predictions are reliable.

With the model working as anticipated the intended problem of corruption will be resolved such that players not appearing on the leaderboards will not be allowed to play for the national team unlike before when players used to pay coaching staff or any member of the board to get into the team. So, then performance becomes key and performance will define recent form which will benefit Kenyan cricket.

Moreover, management of players will also improve as the performing players will get a chance to play and not any player who is not performing or never performing and subbing them out with those that are performing. This is managing the performing players well.

Also, the best 11 will be chosen for the upcoming games with those that are predicted to perform and hence their performance will take Kenya to a winning side hence the selection tool comes in handy if the team wants to win then they will have to go with their best players who will perform in those games.

And finally, as earlier mentioned since the players knowing that they are selected in the first 11 then they will be highly motivated as they know that they have been predicted to perform well and reduce or remove the nervousness of these players.

In conclusion, in order to revive cricket in Kenya such measures and system are required to put in place. If the system acts as planned then the corruption can be done with and allow the players to openly play as long as they are performing well. Performance is the key if there is no performance then it is obvious that, that particular player is taking up space for other potential players and thus the issue of management is again touched. Moreover, the system will add value to the technical and analysis department of the team by providing them with advanced technologies that can also predict performance of plyers and help in selection of the best playing eleven giving cricket Kenya a competitive advantage over other countries. In the long run the system will be benefitting the game of cricket in Kenya it will definitely lead to recognition and interest of citizens leading to development of the game and attraction of more players to base their careers on the game and get famous by playing for the national team.