EXPERIMENT REPORT

Student Name	Anika Chauhan
Project Name	AT1 Week 1
Date	18/08/2013
Deliverables	<chauhan_anika-14177885-week1_xgboost.ipynb> XG Boost https://github.com/anikachauhan30/at1

---

1. EXPERIMENT BACKGROUND	
Provide information about the problem/project such as the scope, the overall objective, expectations. Lay down the goal of this experiment and what are the insights, answers you want to gain or level of performance you are expecting to reach.	
1.a. Business Objective	The goal of this project is to accurately predict the probability of a player being drafted to the NBA based on their performance statistics from the previous matches. These results will help get the right players to the top.
1.b. Hypothesis	This is a simple hypothesis in which all the features are considered while predicting the probability. This was done so as to look at the holistic picture.
1.c. Experiment Objective	We expect a model which predicts accurate probabilities so as to be able to pick the best players.

2. EXPERIMENT DETAILS	
Elaborate on the approach taken for this experiment. List the different steps/techniques used and explain the rationale for choosing them.	
2.a. Data Preparation	Describe the steps taken for preparing the data (if any). Explain the rationale why you had to perform these steps. List also the steps you decided to not execute and the reasoning behind it. Highlight any step that may potentially be important for future experiments The null values were removed and those classification models were used which were robust enough to handle missing data is any.
2.b. Feature Engineering	Describe the steps taken for generating features (if any). Explain the rationale why you had to perform these steps. List also the feature you decided to remove and the reasoning behind it. Highlight any feature that may potentially be important for future experiments
2.c. Modelling	XGBoost model was used for its robustness and high score of the performance metric – AUROC score.

3. EXPERIMENT RESULTS	
Analyse in detail the results achieved from this experiment from a technical and business perspective. Not only report performance metrics results but also any interpretation on model features, incorrect results, risks identified.	
3.a. Technical Performance	AUROC was use in this case as we were interested in the probability of the prediction being correct rather than the prediction itself.
3.b. Business Impact	The model gives a very high AUROC score of 99.87%. This presents a very positive picture but in the worst case scenario it could be a result of overfitting.
3.c. Encountered Issues	

---

4. FUTURE EXPERIMENT	
Reflect on the experiment and highlight the key information/insights you gained from it that are valuable for the overall project objectives from a technical and business perspective.	
4.a. Key Learning	While XGBoost is a good model to use here, we can take it a step further by altering the parameters.

4.b. Suggestions / Recommendations	Given the results achieved and the overall objective of the project, list the potential next steps and experiments. For each of them assess the expected uplift or gains and rank them accordingly. If the experiment achieved the required outcome for the business, recommend the steps to deploy this solution into production.