Can you recgonize who is the real player in basketball court who take the key impression for the win when they do not get a score, rebound and steal?

AIMS

Now we are at a big data era and we can use large amount volume of data with the developing of technologies such as Data warehouse, OLAP and Cloud Computing to help us make more values. In many areas it has been implementing for a long time and has been proved that these kinds of practice can help people make more values. For example, the recommedation system improve the sales of the product in e-business area. So we want to learn some sucessful experience from them and implement it in our project.

EPV is a kind of data that more detailed and vivid describe how possible in a specific basketball situation can a player earn how much point which is influence by a basketball event. For instance, the best 3-point shooter in NBA take the ball in his hot area and no defender in front of him, so in this siuation*(hot area, no defender are all the basketball events)* we can see that he has a great possibility to get the three point which means in this basketball situation*(all the basketball event make up a specific basketball situation)* the EPV of this player is infinite close to 3.

The aim of this project is to develop a model that can help a NBA basketball team to evaluate who is the best player in what time period of game to execute which tactic by calculating the EPV of every possesion of the ball.

The overall aim can be divided into four objectives:

(1) Collect the player' data in different kinds of basketball situations.

(2) Evaluate several kinds of possession of the ball model to calculate the probability distribution of the next trend of the ball holder

(3) Evaluate several kinds of potential model which can best caculate the EPV*(EPV is a advanced data which can be calculated by base data)*.

(4) Use and deploy the selected model.

BACKGROUND

The existing methods to solve this problem is the competitive risk model. The model mainly uses survival analysis to identify multiple risks of death and changes in time, from which i found opportunities applicable to basketball.

In our project we build EPV model which based on the competitive risk model. In the EPV model we turned "the duration of human life" into "NBA possession of the ball" , "multiple causes of death" into "events on the basketball court." and 'risk' into the probability of various incidents at different times on the court." The EPV at the current moment of each possession of the ball is the weighted average of the possible outcomes of all future offensive options in this situation. To calculate this value we need a probability distribution model that can calculate the next trend of the ball holder for a given player's spatial position on the court, so that we can understand the trend of the ball and the development probability determined by their current status. We call this model the “the possession of the ball” model.

After clarify the definition and function of the possession of the ball model we can now explain that how the EPV model works. First to get the probability distribution of the decision to be made by a particular player in a particular situation. Second to calculate the value of the EPV after the player has made that decision. Taken together, we can know the EPV of a possession the ball at any time and the attack configuration characteristics that generate it.

RESEARCH PROJECT

In this section address the following issues:

• Describe the significance of your project. Why is this an important problem that needs to be solved?

Some players are totally undervalued because they do not get too much award, high statistic records but actually their impression to the win of the game may only be recognized by their temamates on the court, the audience will always undervalue these kinds of player, sometimes, maybe even their coach. So we want to bring a new model and new data format to reevaluate their impression to the team because the undervalue of other due to the low simple statistic record is too unfair for them.

• Describe the innovation behind your proposed project. What is new and/or interesting about what you want to do?

If we compare traditional basketball (statistical) game to chess, you will find what interesting is that we focus too much on the movement of each step and lose sight of the overall strategic arrangement of those movements. Just as winning or losing in chess is often not the last step, so is every possession in basketball. The last shot doesn't mean everything. So what is new in our project is that we can focus somethings more detailed and deep in the basketball court to bring us a new sight of analyzing.

• Briefly outline how you are going to solve the problem. For example, you can provide a reasonable breakdown of the tasks, such as: (i) Construct Training Set, (ii) Build Classifiers using Python+some libraries, and (iii) Evaluation.

(1) Collect and Construct the data from NBA official website.

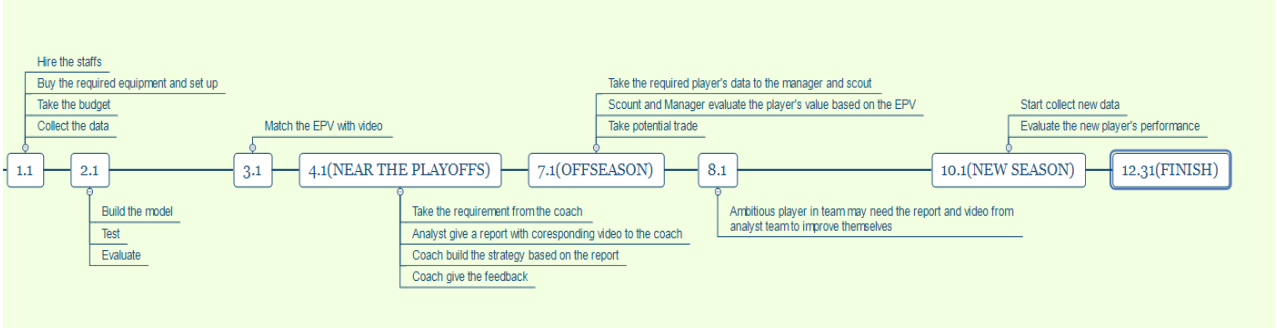
(2) Build the possession of the ball model using the training data in Python and get the propability for the use in the EPV model.

(3) Reference the competitive risk model.

(4) Modify parameters in competitive risk model in order to build the EPV model.

(5)Evaluate the EPV model.

• Give timelines or a Gantt chart of how the subprojects would fit together. Remember you should aim for a project that would take a year to complete.



• Describe what you expect the outcome of the proposed project would be and ideally, how the investor would benefit.

The outcome of our project would be a analyzing report that can show the EPV of a player in every time in every game in all kinds of situations.

Ideally the basketball team would get a new understanding of the what player's impression to their team and the role player play in the game. They can also more easily make a choice to chooes one of them when new plan to sign a group of new player. Finally the team make more wins in a new season.

BUDGET

Personnel (For a NBA basketball team):

-A group of Data Analysts(They must all know competitive risk model)

200k dollors * 4

-A people who know how to operate super computer and also a Data Analyst

250k dollors

-Video recording engineer

150k dollors

Equipment:

-Camera to record the game details

10k dollors * 16

-Super computer(Not the most advanced but need it have good performance on calculating)

50 – 100 k dollors

PERSONNEL

Data Analysts: They will do the most basic and also important work which is to build, test and evaluate the model.

People who know how to operate super computer: He will help the Data analyst to run the data base on the model they build by operating super computer.

Video recording engineer: He will provide the video and image to the Data analyst when they want to match the EPV they work out with the coresponding time point showing on the video.

REFERENCES

Ackoff, R. L. Ackoff's Best: His Classic Writings on Management. New York: Wiley, 1998.

Ackoff, R. L. The Art of Problem Solving: Accompanied by Ackoff's Fables. New York: Wiley, 1978.

Argyris, C. "Resistance to Rational Management Systems" in Decision Making in a Changing World. NewYork: Auerbach, 1971, pp. 13−26.

www.NBA.com

VIDEO PITCH

https://www.youtube.com/watch?v=WGrF9nNmlN8&feature=youtu.be