Applied AI
Homework-III

Football Player Price Estimator

Introduction:

This project aims at creating an engine to estimate the transfer value of a player in the current transfer market across Europe. Twice a year in Football (soccer) leagues across Europe FIFA lifts the transfer window open. During this period in Summer lasting from end of a league season's to beginning of next one and a month's duration in January, football clubs are allowed to sign the players from other teams participating in the same league or different league across Europe.

In recent times, we have observed how the football market has gone crazy with transfer fees for some players going as high as $150 Million Euros. We see that in every season the world record for the transfer fee is shattered for each of the player positions, which is unusual as defenders and goalkeepers were valued less as compared to forward more attacking players. However, that has changed in recent times. The major concern that these high transfer fee raises is whether it is a good business to buy players as such inflated costs because of inflation in the market.

In the current market of global recession where the football clubs are being run more like a business and less like a sports team, it becomes of prime importance for teams to scout the right player and estimate its accurate market value. This not only helps the club to carry out effective negotiations but also make sure that they can get the money invested in the right players by making a high probable bet to gain profits either from player's selling price some years later or through market value of a player.

* This document is built on the framework from the work done in previous assignments and is updated to accommodate the changes made for implementation using Netica Bayes Net with decision node and its effect on the result.

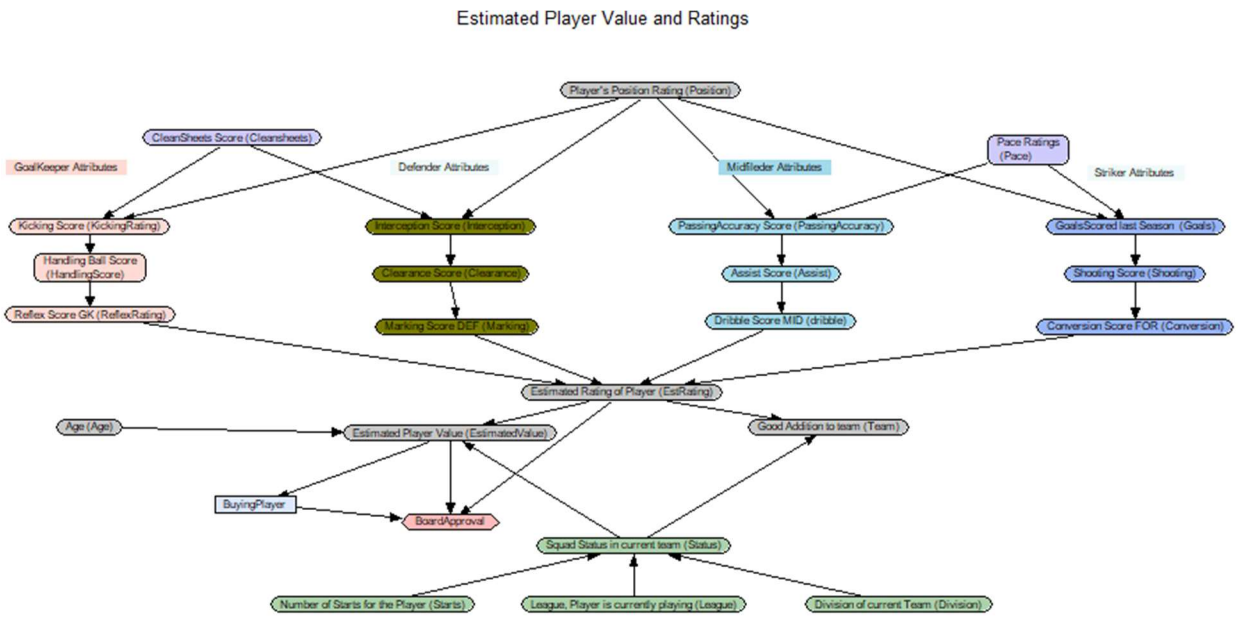
Implementation:

There are many player attributes that could be considered while estimating price however sometimes only attributes are not enough to gauge a player's value. There are many organizations that provide statistics on players and ratings based on their on-pitch performances. These ratings are calculated using very robust and complicated methods and we

shall be using a few among these stats for the sake of simplicity and scope of the project. Following are the factors that we are using to gauge players estimate fees:

1. **Age**: rules are based on the age groups as age has a great influence on the price of the player. Young players with less experience are values less, similar to players who are more the 30 to 35 years old. Players who are young from 20 to 30 years of age are most valued in the transfer market.
2. **Player Position**: The forwards more attacking players have been valued a slightly higher than the players playing in more defensive roles. However, recently there have been some anomalies in this case as fees as high as $70 million were paid for defenders.
3. **Appearance**: More the appearance of a player will corelate to a higher value.
4. **Status in Squad**: If the player whose value is being estimated is an indispensable player for the current team he is in, he will come with a price tag slightly higher as compared a player with similar stats and attributes but with less important role in the team. This is because teams don't like their first team members leave easily, hence they have loftier transfer fees
5. Other attributes such as **pace, reflex, number of goals, interceptions made, assists score, dribbling score etc**., are considered based on the position of the player to calculate his rating and then factors such as if the player is in the **same league, same division etc.,** are also factored in to estimate the current transfer fees of the player.

Network



Decision Node

The decision node here gives us an insight on the whether the club should purchase the player or not. It considers how happy would the board be when the respective club move forward with a player or not, based on the attributes of estimated player rating and his estimated value (nature nodes). The decision node gives us an insight on how the buying process goes on. As we can see the case with the second test case where a good player is not bought because of being expensive. Not all good and excellent player trades go down in the transfer market as it depends a lot on the board's decision to invest in a particular transfer market.

Utility Node

The Utility node here establishes the relationship between Estimated Player Rating, Estimated Player Value (nature node) and Buying Player decision node. Set up such that when the club present the case of purchasing a player which is highly skilled (i.e. values of good and exceptional) in the Rating node and is priced reasonably (i.e. values of Medium and Low) in the Estimated value, the board shows higher affinity towards a decision like that than when club presents a low skilled and higher priced player.

The following network is giving two outputs 1st **Estimated Player Value** and 2nd **Good Addition to the team.** We shall discuss the architecture by dividing the nodes further based on the player's position.

1. **Common Attributes:**
   Attributes that are common to all the players irrespective of positions are grouped together, they include number **of appearances, Division, League** of player's Team. They are summed into property which tells how important is player to the squad **(Status)** according to these variables. These variables are then considered while calculating both outputs, estimated value group and addition to the team.

2. **Player's Position**
   The most important attribute for a player to estimate his/her price is the position at which he/she plays. All other position related attributes are calculated based on this, i.e. the probability is a cross product between their position and one of the Parent Player Attributes.

3. **Parent Position Attributes**
   These are set of two attributes, **Pace and Clean sheets.** Pace is a key attribute for **Midfielder** and **Forward** Player while Clean Sheet is a key attribute for **Goal Keepers** and **Defenders.** This along with the player position forms the base for calculating probability of how good a player is at his/her own position when his attributes are rated.

4. **Position wise Attributes**
   Each Position comes with its own set of attributes, these are:
   GoalKeeper**: Kicking, Handling and Reflex Scores**
   Defender**: Intercept, Clearance and Marking Score**
   Midfielder: **Passing, Assist and dribble Score**
   Forward: **Goals, Shooting and Conversion Score**

   All these attributes are added cumulatively with respect to a position while values of other positions are kept at Below Average level. This aids in calculating the **Player Rating** which is a good metric to gauge the player.

5. **Age**
   As mentioned above age is directly related to the price of the player.

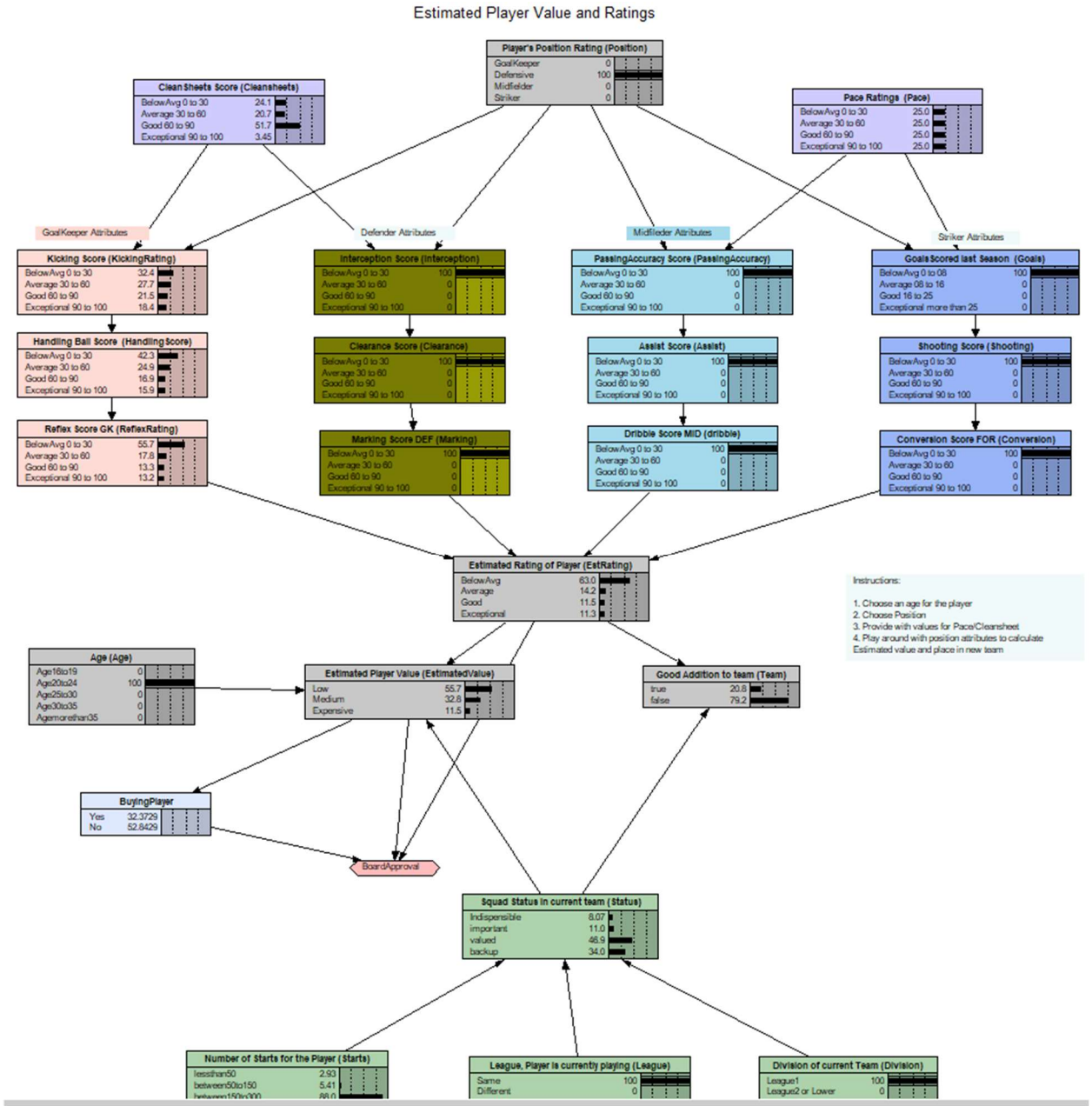
Output

1. Estimated Value of the Player

   Categorized as Low, Medium and Expensive, this nodes takes Age, Rating and Squad Status into consideration to estimate the value of the player.
2. Good Addition to Team
   A matric to determine if the player will be good addition to the team, based on the probabilities around estimated rating and Squad Status.

Test Cases:

1. **Player Position: GoalKeeper** Stats: **Below Average** **Age: 16to19**
   Output: Estimated Rating: **Below Average(54.5)**
   Estimated Value: **Low(93.2**)
   Good Addition to team: **False(77.4)**
   Decision Node: Buying Player: No



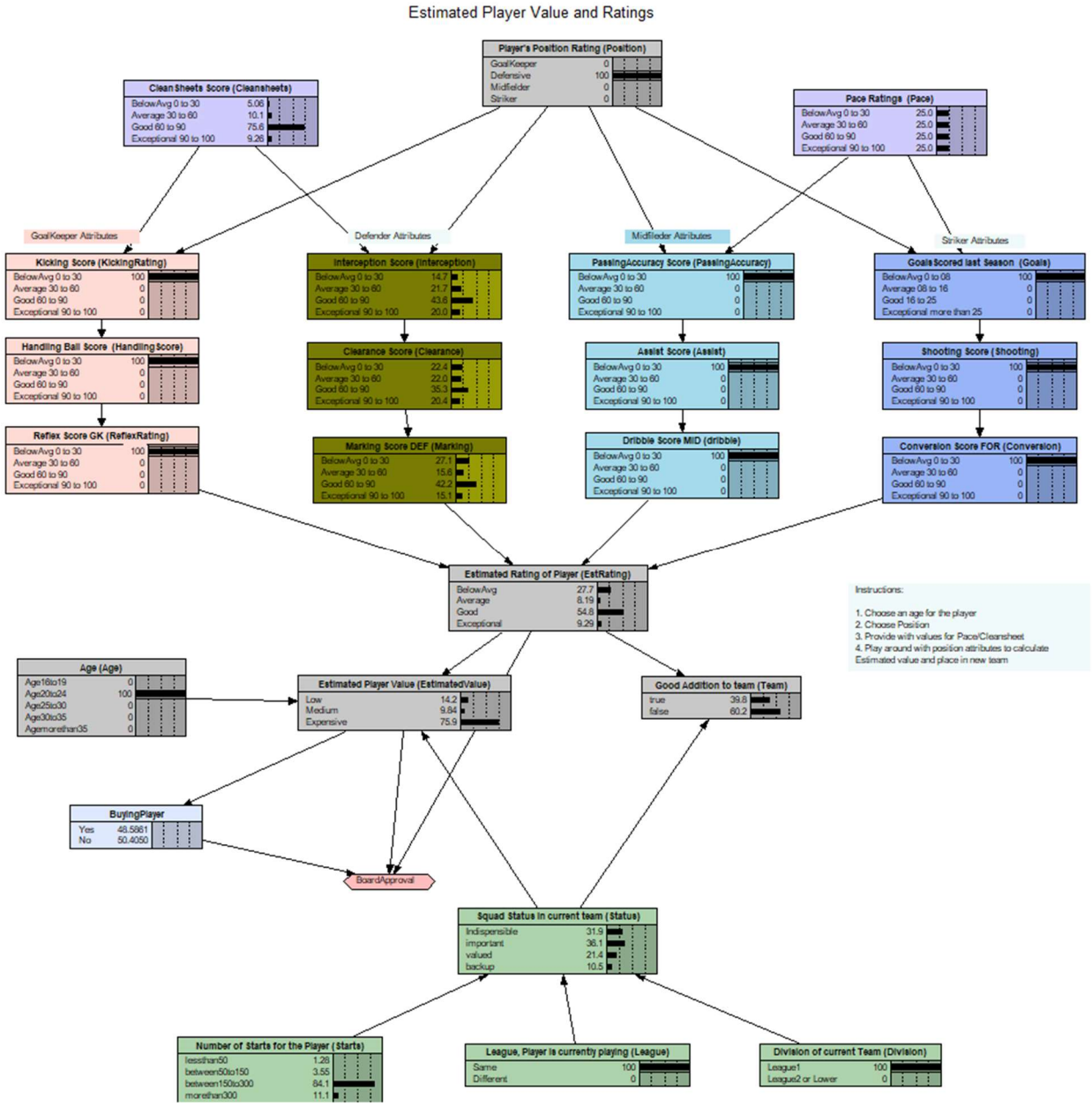
2. **Player Position: Defender** Stats: **Good/Avg** **Age:** : **20to24**

Output: Estimated Rating: **Good(48.6)**

Estimated Value: **Expensive(65.8**)

Good Addition to team: **True(54.8)**

Decision Node: Buying Player: No (50.40 Borderline No, as the player is good but expensive)



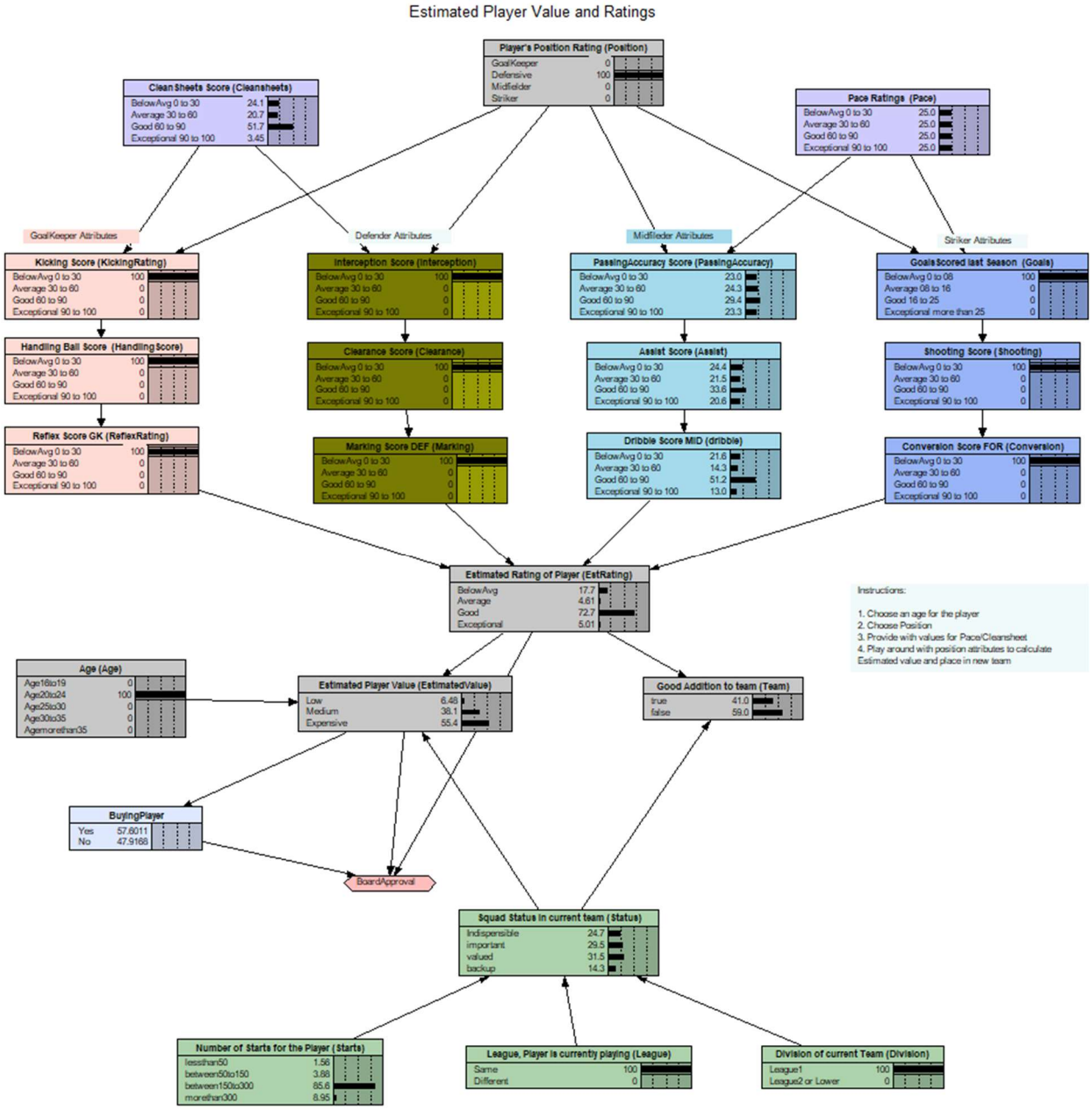
3. **Player Position: Mid** Stats: : **Good/Avg** **Age: 30to35**

Output: Estimated Rating: **Good(48.6)**

Estimated Value: **Expensive(65.8**)

Good Addition to team: **True(54.8)**

Decision Node: Buying Player: Yes (60 good midfielder who's a bit on expensive get approval to but from the board.)



Limitations:

With respect to the scope of the project, not all the contributing attributes of the player was considered, this could influence the transfer fee of the player. Only the topmost important were chosen here. This implementation of Bayer Network does gives a good estimate for calculating the category to which the player would perform and not an approximation for the amount yet.

****Instruction for running the code**:

If you are calculating the estimates for a defender, then please Give the position value of 100. Please try out different values for defender's attribute and stat's such as appearances etc. but keep the stats of all the other positions set to below average for best performance.

Download the file and open it with latest version of Netica on your System
Compile the Program and follow instructions Provided.