GROUP ASSIGNMENT – 3 | DUE: JANUARY 31, 11:59 PM

Language: R
Page Limit: 5 pages
Submission: Report + Code

Football (Some people call it - "Soccer") is a fun game to play and watch. It is the world's most popular sport, played by over 250 million people in more than 200 countries. Because of their vast popularity and fan following, football clubs have very high valuations, and each club pays astronomical transfer fees for top players.

Now after watching Moneyball, the UC Davis soccer team manager has decided to hire a team of MSBA students to predict the future market value of each soccer player to give the team the required commercial and strategic advantage.

Your team's responsibility is to analyze the provided data and characteristics of all the players (Attached with this file) and suggest the best model to predict the future market values of these football players.

Goal:
*Build a model to predict the future market values (**var: Market Value**) of the football players*

Guidelines for the assignment:
It should have the following components so that the report is comprehensive.

1. Data Analysis (20 pts) - Analyze, summarize and visualize the data and the interaction between different variables and features.

2. Feature Selection (20 pts) - After the analysis, work on understanding the correlation between the independent and dependent variables and test with the different model transformations to figure out which features should be dropped. (Clearly explain how you selected the features (variables) you chose).

3. Model Selection (40 pts) - For this assignment, we are focusing on Regularization techniques. Compare the L1 and L2 Norm regularization techniques (Lasso and Ridge, respectively), different penalty terms, and model transformations to build and visualize the optimal model. (*Use Cross Validation to select the optimal tuning parameters and compare it with the AIC model selection*)

4. Model Evaluation (20 pts) - Divide the data into test and train and evaluate the model along with fit, residuals, and standard error plots.

5. Conclusion (Optional, 5 Bonus pts) - Summarize your thought process and discuss future steps and model improvements.

Resources:
Data file: ***FIFA_Player_List.csv****:* Overall Score, Potential Score, Market Value, Weekly Salary, Height, Weight, Age, Preferred Foot, Ball Skills, Defense, Mental, Passing, Shooting, Goalkeeping.