

Sport performance analysis

Project Report

Name: Branko Chomic

Date: 14/04/2016

Table of Contents

Introduction

GUI

Problem encountered

Project features

What have I learned?

What was not achieved?

Recommendations

What would I done differently?

How would I improve project?

Libraries used

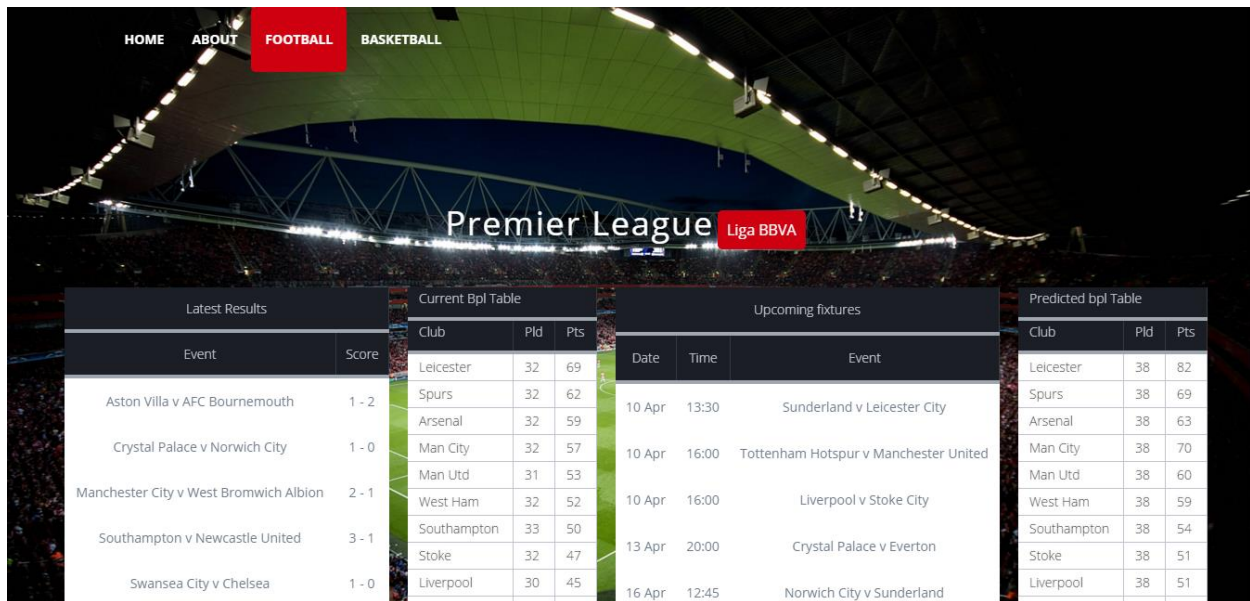
Conclusion

Acknowledgments

Introduction

I have developed a sport prediction web application which helps people in decision making when placing bets. Users that will visit my web app will be able to view predictions for the upcoming fixtures predictions such as correct score, goals scored in a match and recommended team choose to win. It will also provide information like latest predictions, latest results, upcoming fixtures date and time when they will be played, current standings of the teams, predicted winner of competitions such as (Premier league, La Liga, NBA) users can also check the stats based on which algorithm predicts the outcome and also users will be able to view live score from the matches which are updated every minute. The web app includes 2 sports basketball and football and mobile and computer versions are available.

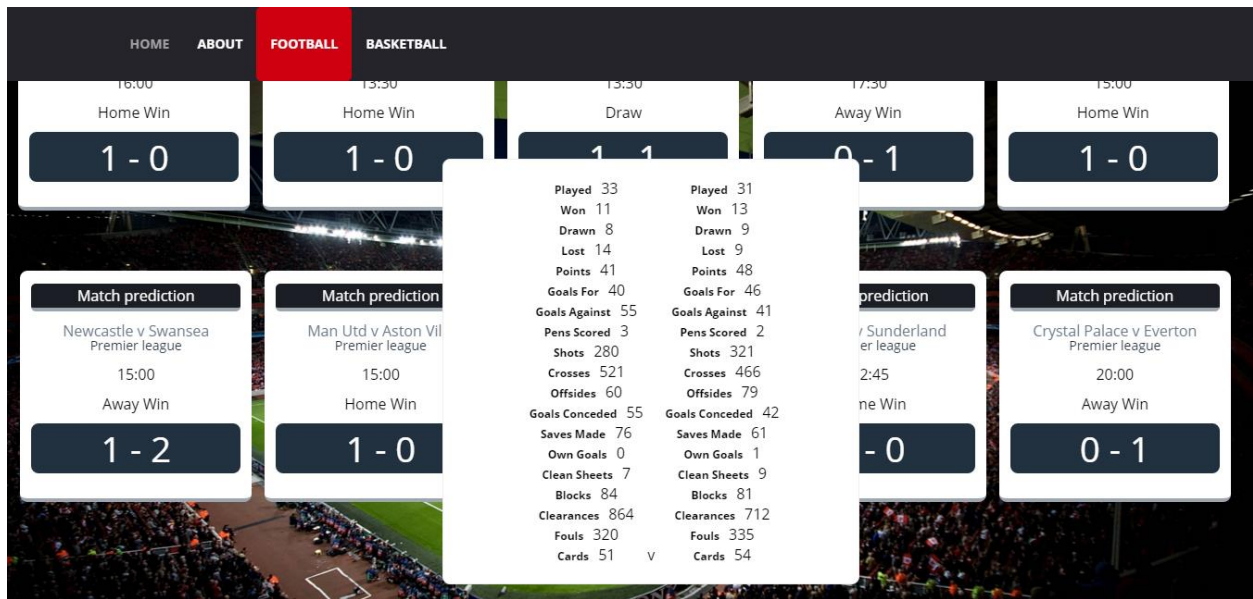
GUI



Users will be able to view latest results, upcoming fixtures, current tables and predicted league table by clicking on Liga BBVA users will switch to Spanish league and all the content change. Users can also choose Basketball as the sport they want to view.



This is a GUI for prediction I wanted something different than just a simple table with predictions in it so I decided to redesign the table that I created for predictions at first and create these boxes which show different information of the game like time, 1x2 prediction, correct score line prediction, name of the game, name of competition.



When clicked on upcoming prediction box with statistics for each team involved will be displayed with home team stats on the left and away team stats on the right. This is feature is useful for users so they are able to see based on what statistics are these prediction predicted.

Problems encounter and how were they solved?

Choosing the right data for predictions

Choosing right data to create testing data sets to give me most accurate results was one of the problems as the training set is the most important when it comes to prediction as the algorithm learns on the training set so the stats had to be chosen carefully to be able to achieve accurate results of at least 55%.

How I solved this problem?

This problem was solved by doing a lot of research on the internet and also use my own experience where I sat down and on piece of paper I wrote down the statistics that i would consider to be the most important when placing bets. For example in football the most important statistics to determine who has better

chance of winning are stats such as(position in the table, attack how many goals scored, goals conceded(defense), saves made , shots on target and many more).

Choosing the right algorithm

I had problem choosing the right algorithm to use for predictions of matches as there are lot of algorithms which could be used .

How I solved this problem?

By doing a lot of research on the internet looking up different machine learning algorithms and trying to figure out which one would be the most suitable for my project. After doing a lot of research I decided to implement 3 algorithms in my project and do a lot of testing trying to figure out which one can give me the most accurate results. After testing each algorithm with 50 matches I decided to go with multilayer perceptron which gave me accuracy percentage of 57 .

Updating content of the website automatically

I had a problem with this as I have never done it and had no idea how it could be achieved. I needed my parser to run every day and run a program that updates live scores to run every minute.

How I solved this problem?

First research was done on how it could be achieved while doing research I found out that I have to run my code on cloud platform or a server. I started doing some research on that I found out that I could use python anywhere for that which I had a bit of experience with but it turned out it wont work as python anywhere doesn't support java and I needed something that could run java

program . After while I found amazon services which is cloud platform which is open source and have a lot of useful features. Amazon services is also very easy to use as amazon provides users with a lot of tutorials which can help you in building web applications . Steps that I did in amazon services to deploy website and update content automatically

1. Buy domain – Domain was purchased on Namecheap.com
2. Create a bucket in s3 instance on amazon and give same name as domain name
3. Go through few tutorials of changing and setting up bucket to act as a webapp
4. Upload html, css, javascript files to the bucket
5. Redirect the domain name home page to bucket
6. Create ec2 instance which is Linux virtual machine running on cloud
7. SSH into it using Git bash
8. Export java applications into jar files
9. Upload website and jar files into Linux instance
10. Create crontab to run jar files automatically every day
11. Create python file to execute command that will replace all files updated in the bucket

Submitted project have following features:

- Live scores for football leagues(Premier league, Spanish league)
- 1X2 predictions
- Predicting current score in matches
- Predicting final league table
- Providing statistics for each prediction
- Displaying latest results, upcoming fixtures, current standings of each team in league table
- Content updated automatically
- Checks how many prediction were correctly predicted last week and displaying them on page
- Time and date upcoming fixtures

- Users can choose from 3 different leagues(NBA, Premier League, Liga BBVA) and two different sports(Football, Basketball)

What was not achieved?

- At the beginning of this project my goal was to make predictions for at least 3 different sports which was not achieved as I have completed only 2
- The second thing I didn't achieve in this project was a making predictions at least 65 percent accurate.

What have I learned?

Understanding of how amazon services works

- How to set up s3 instance
- How to set up ec2 instance
- Set up DNS
- How to associate bucket where the web app files were stored with a domain
- How to send files from ec2 instance to a bucket
- How to use crontab
- How to update content of web app automatically

Machine learning

- What it is
- How it works
- Working with different machine learning algorithms such as multilayer perceptron, Bayesian network and linear regression
- How to create training and testing data sets

WEKA

- How to use classes and methods
- How to implement different algorithms
- How to train data
- How to classify instances

Jsoup parser and Java

- How to parse content from websites dynamically
- Selecting tags, deleting tags
- Improved skills in java
- Fixing issues

Gui

- Improved skills in html, css , jquery
- Create responsive design for desktop/mobile
- Fixing GUI issues

What would I do differently if I would start again?

- Instead of starting with creating gui design I would start with parsing data from websites as this is the most important part of the project and takes a lot of time.
- I would add more features into my web application
- Improve predictions try to get accuracy to high 60s
- Do more testing for predictions with different learning rate, hidden layers to try to get most accurate results

How would I improve my web application ?

- Improve predictions accuracy
- Add percentage for each prediction indicating how sure the classifier that the prediction will happen
- Add more leagues and sports
- Add latest news for each sport
- Adding odds for each fixture

What recommendation would I give to people starting similar project?

- Do a lot of research
- Plan everything from the start to finish
- Start with the most important and most time consuming thing which in this project was to parse all the data
- Do research on machine learning algorithms
- Learn how to use library for machine learning
- When designing GUI look at some tutorials first
- Amazon services tutorials are very helpful

Libraries that I have used with this project

Name: WEKA, accessed 17 January 2016

http://filehippo.com/download_weka/

- Is a machine learning library for java

Name: Jsoup, accessed 3 January 2016

<http://jsoup.org/>

- Jsoup is a Java library for working with real-world HTML. It provides an API for extracting and manipulating data using css and jquery like methods

Name: Apache commons, accessed 10 December 2015

<https://commons.apache.org/>

- This library was used for manipulating strings and using classes such as StringUtils

Conclusion

I know I could add more things to it and make it better but I am happy how it is but I am very happy how it turned up to be at the end. It was lot of fun working on this project, over the course of this project I have learned a lot of useful things which can help me in future projects.

Acknowledgments

I would like to thank my supervisor Greg Doyle for guiding me through this project helping me with ideas, suggestions and giving me always more work to do so I can improve my programing skills.