SUPER PREDICTOR OF INDIAN PREMIER LEAGUE (IPL)

1. INTRODUCTION:

The idea revolves around a sports website with a forecaster and a dashboard. We've developed dashboards with Cognos Analytics, and these dashboards are similar and effective for displaying key IPL findings. The Predictor was built with Cognos analysis and focuses on forecasting the first innings score of future IPL matches using data from previous matches.

2. OVERVIEW:

In India, the IPL has a large following. Fans have been encouraged to watch the matches due to the high degree of unpredictability and close confrontations. Fans are not only interested in watching matches ball by ball, but also in predicting future outcomes, and they frequently use online fantasy sports platforms to do so. Such folks may benefit from this prediction. Using this prediction, IPL teams may create or alter their strategy.

2.1. Existing methods or techniques for resolving this issue:

We frequently see previous data and metrics in IPL matches that reflect the anticipated score or, in certain cases, the chance of a side winning the match depending on the present match scenario. We've built a dashboard and a predictor to help with this. The Cognos Dashboard was built utilising historical data from 2008 to 2019 and is useful for showing key insights as specified in the project. The dashboard visualizes following capabilities:

1. To find the team that won the greatest number of matches in the entire IPL.

2. To find the team that lost the greatest number of matches in the entire IPL.

3. Does winning a toss increase the chances of victory.

4. To find the player with the most player of the match awards in the entire IPL.

5. To find the city which that hosted the maximum number of IPL matches.

6. To find the most winning team for each season.

7. To find the on-field umpire with the maximum number of IPL matches.

8. To find the biggest victories in IPL while defending a total and while chasing a total.

9. Which team won the most matches while batting first.

10. Which team won the most matches while batting second.

11. List of teams which have won matches by most runs cumulatively.

2.2 Proposed solution:

To create the predictor, we first looked for the appropriate dataset that met the requirements. After that, we used a gradient boost regressor to analyse the data and apply the correct steps. Based on the features provided by the user, this predictor forecasts the first innings score. After creating the predictor in Cognos Analysis, we created a website with an HTML and CSS interface. HTML and CSS are used to create a form that collects information from the user.

3. Hardware/Software:

Software Designing of the Project:

❖Frontend

• HTML, CSS

❖ Backend

• Cognos Analytics

Hardware Designing of the Project:

• intel CORE i7 10th Gen (Cognos Analytics)

• intel CORE i5 8th Gen (Cognos Analysis)

4. Experimental Investigations:

<u>While working on the dashboard, we conducted an experiment:</u>

We searched on the internet for different sports dashboards to obtain some ideas for the Cognos dashboard. Then we looked for a match-by-match dataset, which was necessary for creating the dashboard. After that, we filled in some blanks. The data was then cleansed (certain information regarding clubs, umpires, and localities were corrected). We also introduced a mandatory column called "Looser." To get the needed results, we also used Cognos Analytics to create the required queries (such as max, count, and so on). Then we made graphs with line graphs, pie charts, column charts, and so on.

## While working on the predictor, we conducted an experiment:

We looked for data from the first innings ball-by-ball that included all of the needed fields for the prediction. The data was then decoded by determining its form. After that, we cleansed the data by following a few procedures. Some team and venue values were rectified, and some rows that were no longer usable were eliminated.

We next created several graphs to see whether there was any link between the variables, and we removed the columns that were less correlated. Then we looked at which columns should be labelled as 'features' and which should be labelled as 'label.' The data was then divided into training and testing groups. After that, we experimented with several models on our training data and evaluated them on our testing data.

After that, we experimented with several models on our training data and evaluated them on our testing data. Each model was subjected to cross validation. Finally, the models were fine-tuned for greater accuracy. The gradient boost regressor was chosen as our final model since it provided better accuracy than other models.
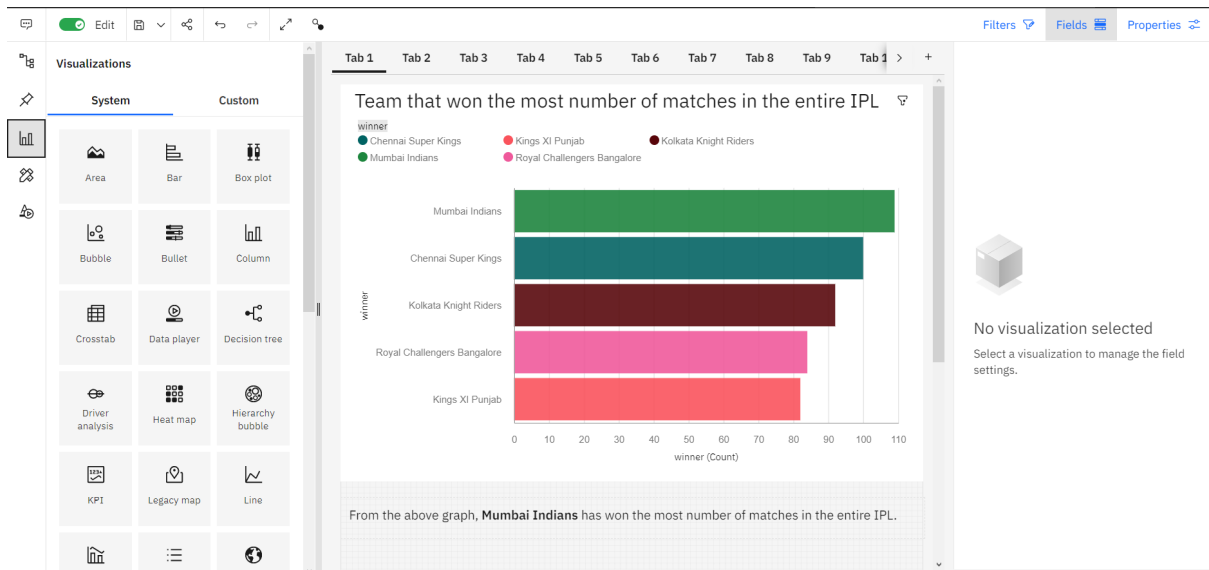
5. Flowchart:

Flowchart of the solution(dashboard):
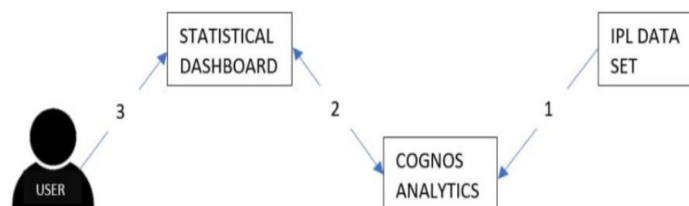
6. Result:

Screenshots of the Dashboard:

From the above graph, **Mumbai Indians** has won the most number of matches in the entire IPL.

Open menu Edit

Filters Fields Properties

Tab 1 | Tab 2 | Tab 3 | Tab 4 | Tab 5 | Tab 6 | Tab 7 | Tab 8 | Tab 9 | Tab 1

Visualizations

System | Custom

Area | Bar | Box plot
Bubble | Bullet | Column
Crosstab | Data player | Decision tree
Driver analysis | Heat map | Hierarchy bubble
KPI | Legacy map | Line
Line and

Team that lost the most number of matches in the entire IPL

winner
● (no value) ● Delhi Capitals ● Kochi Tuskers Kerala ● Pune Warriors
● Rising Pune Supergiant ● Rising Pune Supergiants

winner (Count)

No visualization selected

Select a visualization to manage the field settings.

From the above graph, **Rising Pune Supergiants** has lost most number of matches in the entire IPL.

Edit

Filters Fields Properties

Tab 1 | Tab 2 | Tab 3 | Tab 4 | Tab 5 | Tab 6 | Tab 7 | Tab 8 | Tab 9 | Tab

Visualizations

System | Custom

Area | Bar | Box plot
Bubble | Bullet | Column

No.of teams won the match after winning the toss

winner (Count distinct)

No.of teams loss the match after winning the toss

toss_winner

Dashboard properties

∧ Canvas

Layout positioning   Relative ∨

Page size

Preset   Letter ∨

Fit page   ◯

Width   Height
612 px   792 px

Layout positioning   Relative ∨

Page size

Preset   Letter ∨

Fit page   ◯

Width   Height
612 px   792 px

Area | Bar | Box plot
Bubble | Bullet | Column
Crosstab | Data player | Decision tree
Driver analysis | Heat map | Hierarchy bubble
KPI | Legacy map | Line

No.of teams loss the match after winning the toss

winner (Count distinct)

From the above graph;
No. of teams that won the toss as well as the match (a) = 10
No. of teams that won the toss but not the match (b) = 5
=> a > b

Therefore, majority of the teams have won the toss as well as the match. By this we can conclude that **winning the toss increases the chance of winning the match.**

Layout positioning   Relative ∨

Page size

Preset   Letter ∨

Fit page   ◯

Width   Height
612 px   792 px

Grid
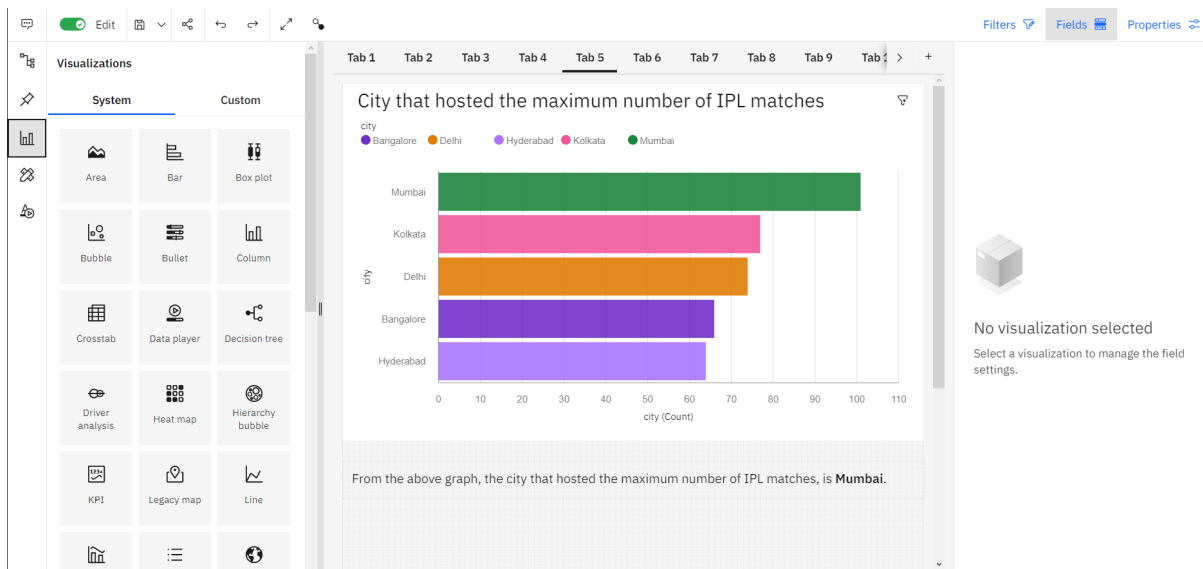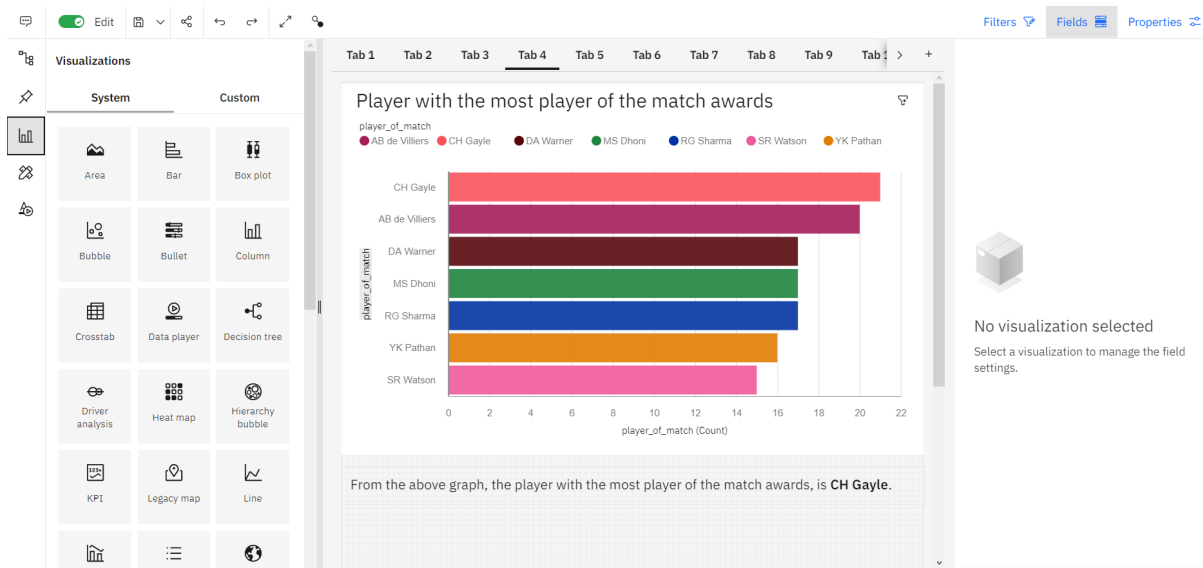
Show grid   ◉
Snap to grid   ◉
Snap to objects   ◉

Footer

Show footer   ◉

7. Advantages and Disadvantages of proposed solution

<u>Advantages:</u>

➢ Frequently delivers unparalleled forecasting accuracy.

➢ Lots of versatility - can optimise on a variety of loss functions and has numerous hyper parameters tweaking possibilities, making the function extremely adaptable.

➢ There is no need to pre-process data; it frequently works well with category and numerical values as is.

➢ Imputation is not necessary to handle missing data.

Disadvantages:

> This will continue to improve in order to reduce any mistakes. This can lead to overfitting by exaggerating outliers.
> Less interpretive in nature, however this may be easily remedied using a variety of techniques.

8. Applications:

> Fantasy Cricket Platforms- Fans who want to anticipate future outcomes might go to fantasy sports platforms online. Such folks may benefit from this prediction.
> Teams in the Indian Premier League (IPL) can utilise this predictor before a match against a specific opponent. They are able to estimate the opponent's score and adjust their strategy accordingly.
> Sports Analysis- You may use our dashboard to examine and analyse historical IPL numbers and forecast future results.

9. Conclusion:

Finally, we integrated all of the domains (frontend, backend, dashboard, and predictor) to build a website for IPL fans that includes a predictor and dashboard in one location. The Cognos Dashboard is excellent for showing key IPL findings. Predictor calculates a final total based on factors such as batting team, bowling team, venue, overs, runs, wickets, runs last5 and wickets last5.

10. Future Scope:

> We may update the data after each IPL season by adding the data from that season to improve the accuracy of our model.
> We can add a login/sign in option to our website to provide extra functionality.
> We can incorporate a chatbot to allow consumers to simply engage with our website.

11. Bibliography:

<u>References of websites referred:</u>

https://youtu.be/usX1PuOht_w

https://www.ibm.com/docs/en/cognos-analytics/11.1.0?topic=stories-get-started-dashboards

Dataset used in predictor-

https://www.kaggle.com/rushikeshlavate/xgboost-for-predicting-ipl-s-1st-inning-score