

**A PROJECT REPORT
ON
“ ANALYSIS OF AUCTION IN IPL”**

**Submitted to
KIIT Deemed to be University**

In Partial Fulfillment of the Requirement for the Award of

**BACHELOR’S DEGREE IN
COMPUTER SCIENCE AND ENGINEERING**

By

2105188

BASTAB BARIK

2105842

SUVRA MUKHERJEE

2105853

DIPANJAN KAMILLA

21051596

SHRUTI DUTTA

21052528

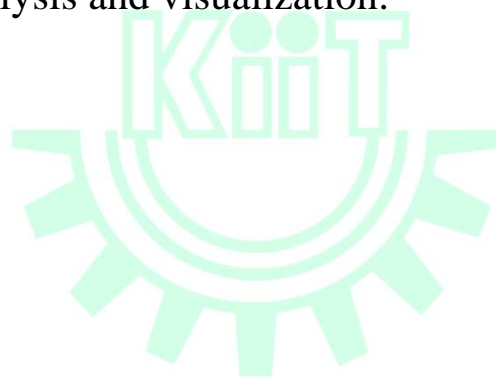
SARTHAK GHOSH

**UNDER THE GUIDANCE OF
Dr Soumya Ranjan Mishra**



Abstract

A single player's massive data output across an entire line means that data science and machine learning are becoming more and more important in cricket every day. In order to assess and forecast results from the Indian Premier League (IPL) cricket competition, this project makes use of Python and machine learning techniques. To obtain insights, preprocessing is done on historical IPL data and exploratory analysis is carried out. Feature engineering extracts important features impacting match outcomes. Using the Python programming language and a number of tools, including Pandas, NumPy, and Matplotlib, this project explores the analysis of IPL 2022 auction data. Insights into player values, team tactics, and market patterns that surfaced during the auction are to be gained through exploratory data analysis and visualization.



Introduction

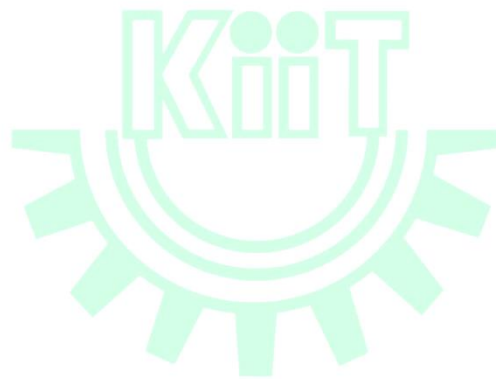
Bidding wars erupted during the IPL 2022 auction as teams fought to acquire the most promising players for their teams. The goal of this project is to use Python and related libraries to analyze the IPL 2022 auction data in order to find patterns and trends in player acquisitions, club expenditure, and other pertinent variables. Through the explanation of its decisions in a format that is both human-readable and interpretable, XML makes it possible to comprehend how a machine learning model generates predictions. The preparation and division of the data into training and testing sets comes first. After that, we train a machine learning model with the h2o package and use the lime package to characterize the predictions it made.

Last but not least, we include the explanations to help you comprehend how the model produced its forecasts. This might prove beneficial for machine learning and data science professionals who aim to develop more trustworthy and transparent artificial intelligence systems.



Objectives

- To evaluate how player prices are distributed.
- To determine which teams paid the most amount for each player.
- To investigate the connection between player auction prices and performance metrics.
- To evaluate the tactics and expenditure habits of the team.
- In order to spot any patterns or trends in the player acquisitions.



Methodology

1. Data Collection:

Consulting with reputable sources or the official IPL websites to acquire information about the 2022 auction.

2. Data Processing:

Preparing the data for analysis requires cleaning, handling missing values, and proper formatting.

3. Exploratory Data Analysis:

For descriptive statistics, visualizations, and the creation of insights, use Pandas in exploratory data analysis (EDA).

4. Statistical Analysis:

Utilize NumPy for statistical analysis and numerical computations to investigate the correlations between variables.

5. Visualization:

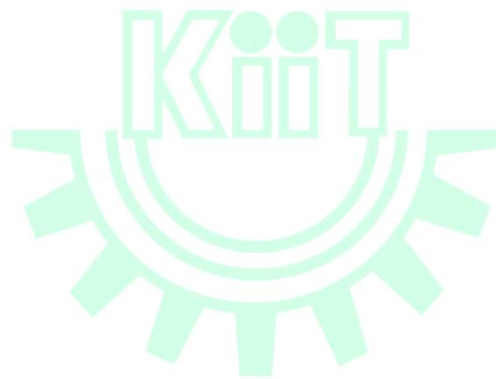
Usage of Matplotlib to generate histograms, box plots, scatter plots, and heatmaps, among other visual representations, to effectively communicate findings.

6. Interpretation:

Make inferences from the data and pinpoint important revelations about player values, team tactics, and market dynamics.

Tools and Technologies Used

- The Python programming language
- Pandas for analysis and data manipulation
- NumPy for statistical analysis and numerical calculations
- Matplotlib for displaying data
- Jupyter Notebook for interactive documentation and analysis



Result and Findings

- **Distribution of Player Prices**

Analysis shows that the distribution of player prices is skewed, with a few high-value outliers and the rest of values falling inside a specific price range.

- **Most Pricey Participants**

Determine which players are the most expensive and examine the teams that have acquired them to provide insight into high-end talent acquisition tactics.

- **Relationship between Performance and Price**

To comprehend the factors that influence valuation, examine the relationships between player performance indicators (such as batting average and bowling economy) and their auction prices.

- **Team Expense Trends**

Examine how other teams distributed their funds, looking for trends in expenditure and tactical methods used throughout the auction.

- **Patterns and Trends**

Emerging trends or patterns in player acquisitions, such as a penchant for particular player types, nationalities, or playing styles, should be noted.

The following graphs are obtained from execution of the code.

```
In [167]: plt.pie(type.values, labels=type.index, labeldistance=1.2, autopct='%1.2f%%',
               shadow=False, startangle=60)
plt.title("Role of player Participated", fontsize=17, loc='right')
# plt.plot()
plt.show()
```

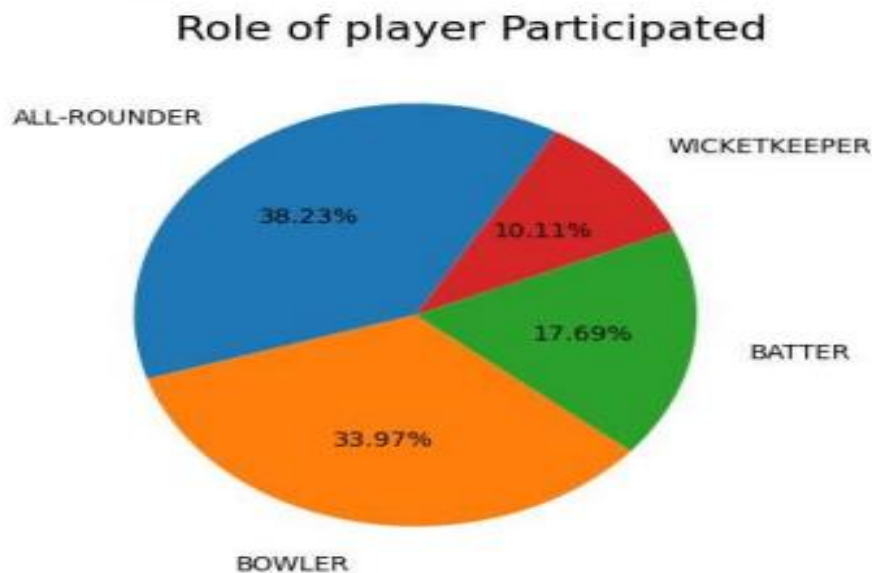


Fig. 1. Role of Player Participated

```
In [168]: plt.figure(figsize=(10,5))
fig=sns.countplot(ipl,x='status',palette=['orange','pink'])
plt.xlabel('Sold or Unsold')
plt.ylabel("Number of players")
plt.title("sold vs Unsold", fontsize=16)
fig.bar_label(fig.containers[0])
plt.show()
```



Fig. 2. Graph to show the contrast between sold vs unsold players


```
In [170]: plt.figure(figsize=(10,7))
fig=sns.countplot(ipl[ipl['status']=='sold'],x='Team')
plt.xlabel("Teams Names")
plt.ylabel("Number of players ")
plt.title('Player Bought by each team ',fontsize=14)
plt.xticks(rotation=50)
fig.bar_label(fig.containers[0])
plt.show()
```

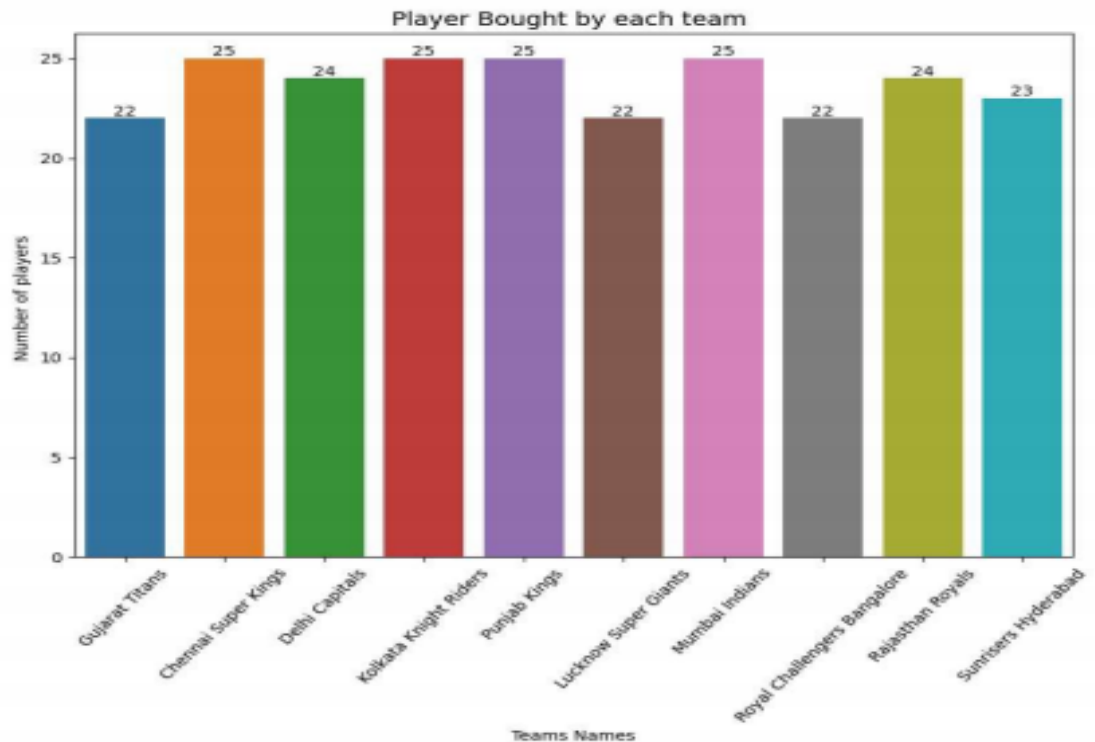


Fig. 3. Graph showing the No.of players bought by each team in IPL

```
In [180]: plt.figure(figsize=(20,10))
fig=sns.countplot(ipl[ipl['Team']!='Unsold'],x = 'Team',hue=ipl['TYPE'])
plt.title("player in each team")
plt.xlabel('Team Names ')
plt.ylabel("Nuber of Player")
plt.xticks(rotation=50)
for p in range(ipl.groupby(['TYPE'])['TYPE'].count().count()):
    fig.bar_label(container=fig.containers[p])
plt.show()
```

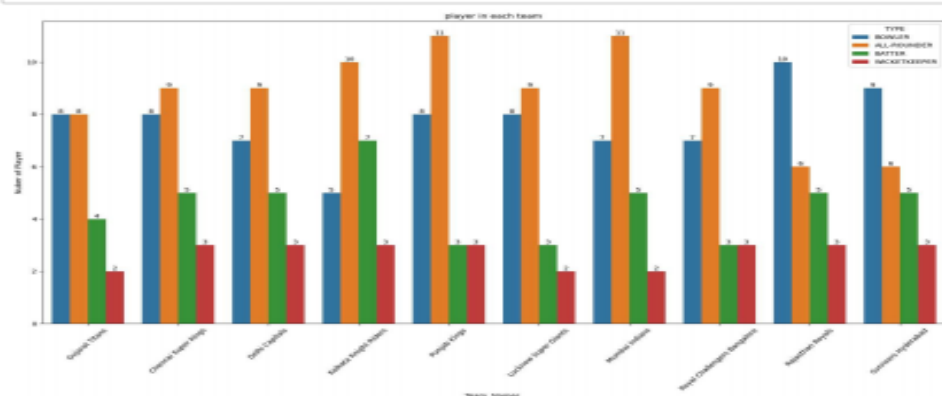


Fig.4. Graph showing players who have been retained by their teams in IPL

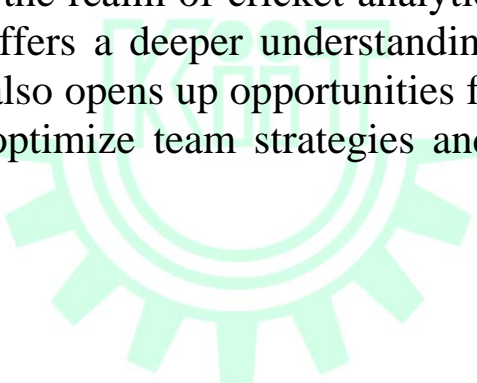
Conclusion

The examination of the IPL 2022 auction data provides valuable insights into various aspects such as market trends, strategic approaches adopted by teams, and the valuation of players. Leveraging Python and its associated libraries, we can delve deep into the intricacies that shaped the composition of teams for the upcoming IPL season. By employing sophisticated analytical tools, we can unravel the dynamics underlying the auction process, shedding light on the factors influencing bidding strategies and player selections.

In recent times, there has been a growing interest in explainable machine learning (XML) techniques, primarily due to the increasing adoption of complex algorithms that impact real-world scenarios. XML focuses on developing models that not only deliver accurate predictions but also provide transparent explanations for the decisions they make. This transparency is crucial, especially when these decisions have significant ramifications for individuals or society as a whole. As machine learning algorithms continue to permeate various domains, the demand for interpretability and accountability in decision-making processes has escalated, making XML an increasingly preferred approach in diverse applications, including sports analytics like in the case of IPL auctions.

Future Works

Predictive modeling in the context of sports, particularly in the Indian Premier League (IPL), involves using statistical techniques to forecast player performance based on auction prices. This approach aims to provide insights into how player acquisitions impact team performance throughout the IPL season. By delving into the relationship between player costs and on-field outcomes, researchers can uncover valuable patterns and trends that influence team success. Furthermore, the integration of advanced machine learning algorithms can enhance the depth and accuracy of such analyses, enabling more nuanced predictions and strategic decision-making in the realm of cricket analytics. Exploring these avenues not only offers a deeper understanding of the dynamics within the IPL but also opens up opportunities for leveraging data-driven insights to optimize team strategies and player selections for future seasons.



References

- Official IPL website
- Cricket news websites
- Documentation and tutorials for Python, Pandas, NumPy, and Matplotlib.

