DC’S AUCTION STRATEGY FOR IPL 2025

Team Name - Tailenders
Team Members :
Anubhav Rathi | anubhavrathi.et21@rvce.edu.in
Vasanth Subramanian | vasanthsubramanian0694@gmail.com
Debraj Bose | bosedebraj0609@gmail.com

Problem Statement:

How can the Delhi Capitals (DC) formulate a winning squad for the upcoming IPL season with the help of mega auction 2025?

Rationale For The Problem Statement:

Given the Delhi Capitals' (DC) history of underperformance and the upcoming 2025 IPL mega auction, we are conducting an analysis of data from the past three years to identify a pool of 70-80 players from which DC can select 20-25 players to build a squad capable of winning their first IPL trophy.

Objective:

The objective is to build a dream squad that will lead the Delhi Capitals to victory in the IPL 2025. The team should be composed of skilled players who can quickly adjust and adapt to different conditions and phases of the game. The focus is on creating a consistent and aggressive team, ensuring that each department and phase of the game is covered by highly capable players.

DC’s Expected Retained Players:

- Rishabh Pant
- Axar Patel
- Kuldeep Yadav
- Jake Fraser-McGurk

DC’s Expected RTM player1:

- Tristan Stubbs

Our Approach:

We analyzed IPL data from the past three years to identify a pool of batters and bowlers using the Satty Scale, which incorporates factors such as dot percentage, runs per innings, and strike rate for batters, and economy rate, wickets, etc., for bowlers.

Initially, we compiled lists of batters and bowlers and identified those with significant overall performance intersecting with the performance at Kotla/Arun Jaitley Stadium of the last 3 years (i.e., from 2022). From this list, we selected a few players without applying a strict Satty Scale score threshold.

Next, we refined our selection by examining the overlap with performers outside Delhi, applying a Satty Scale score threshold of 0.3 and 0.5 for batters and bowlers.

Road Map:

1. We got the Dataset from Kaggle and cricsheet.
2. Data preprocessing and cleaning.
3. Data Transformation and Dimensionality Reduction.
4. Use of Scatty scale for Normalization of Data.
5. Data Splitting depends upon different phases of the Game.
6. Intersection of different phases based on players performance in Kotla and overall.
7. Filtering out the players with certain criteria.
8. After all the analysis,we got a pool of 70-80 Players.

OVERALL TOP BATTERS FROM IPL 2022:

The filtering is done based on the batters who played in IPL from 2022 and their performance is analyzed on the basis of phases i.e Powerplay,Middle Overs and Death Over.

After preprocessing and normalization, we got a set of batters. Below you can find the Colab link for the reference.

Top_Batters.ipynb

OVERALL TOP BOWLERS FROM IPL 2022:

The filtering is done based on the bowlers who played in IPL from 2022 and their performance is analyzed on the basis of phases i.e Powerplay,Middle Overs and Death Over.

After preprocessing and normalization, we got a set of bowlers. Below you can find the Colab link for the reference.

Top_bowlers.ipynb

TOP BATTERS IN KOTLA FROM 2022:

The filtering is done based on the Batters who played in Kotla and their performance is analyzed on the basis of phases i.e Powerplay,Middle Overs and Death Over.

After preprocessing and normalization, we got a set of batters. Below you can find the Colab link for the reference.

Top batters in "Kotla" - for and against DC in each phase >=2022

TOP BOWLERS IN KOTLA FROM 2022:

The filtering is done based on the Bowlers who played in Kotla and their performance is analyzed on the basis of phases i.e Powerplay,Middle Overs and Death Over.

After preprocessing and normalization, we got a set of bowlers. Below you can find the Colab link for the reference.

Top bowlers in "Kotla" - for and against DC in each phase >=2022

TOP BATTERS OUTSIDE KOTLA FROM 2022:

The filtering is done based on the Batters who played outside Kotla and their performance is analyzed on the basis of phases i.e Powerplay,Middle Overs and Death Over.

After preprocessing and normalization, we got a set of batters. Below you can find the Colab link for the reference.

Top batters in each phase >=2022 except Kotla

TOP BOWLERS OUTSIDE KOTLA FROM 2022:

The filtering is done based on the Bowlers who played outside Kotla and their performance is analyzed on the basis of phases i.e Powerplay,Middle Overs and Death Over.

After preprocessing and normalization, we got a set of bowlers. Below you can find the Colab link for the reference.

Top bowlers in each phase >=2022 except Kotla

Scatty Scale For Batters:

Pair-wise Matrix		SR	RPI	BPD	dot_percentage
	SR	1.00	0.33	1.00	0.33
	RPI	3.00	1.00	2.00	0.33
	BPD	1.00	0.50	1.00	0.50
	dot_percentage	3.00	3.00	2.00	1.00
	sum	8.00	4.83	6.00	2.17

Normalised		SR	RPI	BPD	dot_percentage	Criteria_Weights
	SR	0.13	0.07	0.17	0.15	0.13
	RPI	0.38	0.21	0.33	0.15	0.27
	BPD	0.13	0.10	0.17	0.23	0.16
	dot_percentage	0.38	0.62	0.33	0.46	0.45

Scatty Scale for Bowlers:

Pair-wise Matrix		dot_percentage	SR	Eco	Avg
	dot_percentage	1.00	1.00	0.50	2.00
	SR	1.00	1.00	1.00	2.00
	Eco	2.00	1.00	1.00	2.00
	Avg	0.50	0.50	0.50	1.00
	sum	4.50	3.50	3.00	7.00

Normalised		dot_percentage	SR	Eco	Avg	Criteria_Weights
	dot_percentage	0.22	0.29	0.17	0.29	0.24
	SR	0.22	0.29	0.33	0.29	0.28
	Eco	0.44	0.29	0.33	0.29	0.34
	Avg	0.11	0.14	0.17	0.14	0.14

FINAL POOL OF PLAYERS:

After going through each department and each phase,we finalized the pool of players who will be our target in the upcoming auction to find the right balance of our squad.

Below here you can find the Colab links and csv files for final pool of players

Tailenders.ipynb

Batters

Bowlers

ML Prediction:

After getting data of about 74 players(both bowlers and batters),we thought of predicting the best batters and bowlers from this pool of 74 players.

Using **supervised learning** model i.e Random Forest Classifier, we gets 13 Batters and 12 Bowlers.The data set was divided into 70% training dataset and 30% test dataset.

Below here you can find the Colab links for the predicted player using ML algorithm.

Prediction_through_ML.ipynb

Final_Predicted.xlsx

Result & Conclusion:

1. Performance Analysis at Kotla vs. Outside Kotla:

- **Batsmen:** The analysis of player performance at Kotla and outside Kotla provided insights into how environmental and pitch conditions impact individual performances. Some players who excelled in Kotla's specific conditions might not have performed as well outside, and vice versa. This differentiation was crucial in identifying consistent performers across varied conditions.

- **Bowlers:** Similarly, bowlers showed varying degrees of effectiveness depending on whether they were playing at Kotla or other venues. The ability to adapt to different pitch conditions was a significant factor in determining the best bowlers.

2. Selection of Top Performers:

- Using the analysis, we identified the top-performing batsmen and bowlers. The criteria for selection were based on consistent performance metrics such as average scores, runs, strike rates for batsmen, and wickets, economy rates for bowlers.

- The selected pool consisted of 74 players who demonstrated outstanding performances in their respective roles.

3. Machine Learning Model Application:

- **Batsmen:** Machine learning algorithms were applied to the dataset of batsmen to predict and rank players based on an "avg score" average (Satty Score Avg), which was an aggregate metric considering various performance indicators. The model helped in refining the selection by ranking players in descending order, ensuring that only the best-performing batsmen were highlighted.

- **Bowlers:** A similar approach was taken for bowlers. The ML model identified the most effective bowlers, taking into account their performance in different conditions and overall consistency, ranked by the Satty Score Avg.

4. Final Pool of Best Players:

- The application of machine learning models resulted in a final list of the best batsmen and bowlers, sorted by their Satty Score Avg in descending order. This list represents the top players who are most likely to perform exceptionally well, regardless of the conditions they face.

Conclusion:

- The integration of comprehensive EDA with machine learning models helped in comprehensive analysis of the dataset, combined with machine learning models, has identified a **final pool of 74 players**, further refined by **machine learning**, which brings it down to **13 Batters and 12 Bowlers(**including **all rounders)**.
These players, ranked by their Satty Score Avg, represent the best talent in the league. By leveraging this data-driven approach, Delhi Capitals (DC) can strategically select these players to form a formidable squad. This meticulous selection process will significantly enhance their chances of winning the next IPL season, giving them the competitive edge needed to dominate the tournament.