indium
Make Technology Work

Predictive Team Selection using Player Performance Analytics

Digital Services

Success Story

Client

The client is a technology-centric Sports Consultant which advises professional teams across different sports on strategies that lead to performance enhancement.

Overview

Client's team competes in the world's 3rd most expensive sporting league. Most of the team's expense come from paying its superstar players' salaries, without any guarantees on their onfield performance and team chemistry. The allure of a player's star power resulted in teams overlooking hundreds of other promising, but unknown, players who have a far higher ROI. Indium Software performed Advanced Predictive Analytics and Data Mining on the hundreds of available players participating in the auction, to identify and rank the players who would be the best investments for the team.

Status Quo

The client is a technology-centric Sports Consultant which advises professional teams across different sports on strategies that lead to performance enhancement. It is an expert in collecting large troves of first-hand data from each player/game and transforming them into strategic insights for team owners or coaches to act upon.

The specific business need in question was regarding the **world's 3rd most expensive sporting league,** for which each city's team would need to bid for players in the upcoming player auction. Traditionally, team owners would bid for players based on a combination of the player's reputation and the coach's personal opinions. **This led to all teams bidding exorbitant sums** for a small group of famous players who were in many cases not ideally suited for the teams bidding for them. Additionally, there was no bidding consultant capable of advising on the performance or playing style of each of the hundreds of relatively unknown and overlooked but potentially talented players.

Business Requirements

The client partnered with Indium Software for the following:

- **Recommendations** on which players to bid for and the analytical reasoning via statistical evidence.
- **A ranking list** of the most promising players by their playing position using CPIs (Composite Performance Indicators) which are to be developed in conjunction with domain knowledge.
- The rankings should leverage years of highly specific player & game statistics and be **objective, comprehensive** (50+ criteria) and account for players' 'form'.
- Coaches should be able to scan the rankings and infer which of the players **best fit their teams'** needs by digging deep into the accompanying analytical metrics.

Business
Predictive Analytics, Data Mining

Domain
Professional Sports

Tools
Python, SQL Server, XGBoost

Key Highlights

- The team was able to plan its bidding strategy which led to it utilizing only 70% of its bidding budget

- The recommendations narrowed the pool of players from 350 to 20 permitting the coach to target his focus

Solution

Indium Software analyzed the data collected by the client and approached the solution in the following methodology:

- The solution pertained to two cases - Ranking bowlers and batsmen separately using different criteria for each. For both cases, the preliminary steps of data cleaning and data aggregation were performed.

- Data Cleansing – The data was cleansed and formatted by combining unrelated data sets across games, tournaments and country leagues to form a unified, structured database.

- Data Aggregation – In a sport like cricket where multiple data points for a playing medium like batting can be collected, the aggregate statistics for each player can be highly complex. The preliminary set of relevant aggregates were chosen after brainstorming with the client.

- Index creation – To rank the list of players, the team created formulae and algorithms to evaluate player performance using analytics.
  1. Compiled broad aggregate statistics for each individual player.
  2. Ascertained the relevant metrics which drove good player performance for each department role (bowling/ batting) using statistics and domain research.
  3. Statistics, statistics and more statistics. Advanced analytics techniques were leveraged to generate relevant, dependable and detailed statistics which exposed the players' strengths and weaknesses.

- Two methods were used for calculating a Composite Performance Index.
  A) A Descriptive method - using formulae to derive the bowling and batsmen strength.
  B) A Predictive method - using ML methods on historical data to determine the index.

A) Descriptive method
1. What are the concepts behind the batting index?

The Composite Performance Index for batsmen compares each batsmen's performance based on:
- Strike rate in different phases of T20 matches.
- Strike rate compared to other players.
- The recent tournaments' performance.

2. What are the concepts behind the bowling index?
The Composite Performance Index for bowlers compares each bowler's performance based on:
- Strike Rate index, the number of balls taken to get a wicket, and Economic Rate Index which gives the number of runs per over; both in different phases of T20 matches.
- SR and ER compared to other players.
- The recent tournaments' performance.

B) Predictive method
1. The team first created a descriptive index which ranked the list of players and then redefined the calculations for the descriptive index until the indexing rank matched the "original rank". The descriptive index was the dependent variable.

2. Trained a model to ascertain the player characteristics (independent variables) and their weights to determine their ranks.
- Variable's selection was performed by running multiple models such as Linear Regression, Random Forest and Recursive Feature Elimination technique.

3. Iterated the model performing transformations to the variables to test the model since absolute strike rates, economic rates, averages of batting and bowling may have skewed distributions.

Significant features of the Predictive model included its ability to incorporate:

- Comparative analysis between players to adjust statistics so they accounted for variations and discrepancies in the games that different players played.

- Differential weightages of tournaments to account for varying conditions, players' improvements and optimize for the upcoming league.

- Custom methods to average the accuracy of the CPI prediction.

Business Impact

- Most of the top 10 most bid bowlers and batsmen figured in our recommendations.
- The recommendations narrowed the pool of players from 350 to 20 permitting the coach to target his focus.
- An objective and comprehensive ranking of each available player (indicating performance) was presented alongside revealing statistics (indicating team fit).
- The team was able to plan its bidding strategy which led to it utilizing only 70% of its bidding budget.
- Discovered high performing and good-fit players who were not on the team captain, coach or team owner's radar.
- Provided precise statistics of the selected players' strengths and weaknesses to leverage during team training.

📞

INDIA

Chennai | Bengaluru | Mumbai
Toll-free: 1800-123-1191

USA

Cupertino | Princeton
Toll-free: 1 888 207 5969

UK

London

SINGAPORE

+65 9630 7959

🌐

General Inquiries
info@indiumsoftware.com

Sales Inquiries
sales@indiumsoftware.com