Integer Optimisation for Dream 11 Cricket Team Selection

Article *in* International Journal of Computer Sciences and Engineering · November 2020

DOI: 10.26438/ijcse/v8i11.16

Saurav Singla
Mindrops

Swapna Samir Shukla
National University of Singapore

IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol. 8, Issue.11, November 2020 E-ISSN: 2347-2693

Integer Optimisation for Dream 11 Cricket Team Selection

Saurav Singla[1*], Swapna Samir Shukla[2]

[1]Independent Researcher, Gurgaon-122011, India
[2]Independent Researcher, Bhubaneswar-751030, India

**Corresponding Author: sauravsingla08@gmail.com, Tel.: +91-9958793952*

DOI: https://doi.org/10.26438/ijcse/v8i11.16 | Available online at: www.ijcseonline.org

Received: 07/Nov/2020, Accepted: 19/Nov/2020, Published: 30/Nov/2020

Abstract – In the recent years, Dream 11, a fantasy sports platform has taken the Indian gaming landscape into storm by raking in a valuation of 1 million USD. One of the important aspects of participating in a Dream 11 contest is team selection. Though Dream 11 hosts fantasy Cricket, Kabbadi, Football and Basketball in its platform, Fantasy Cricket has gained more users due to its popularity in India. Moreover, Cricket is one such sport that generates large volumes of data, and therefore provides many opportunities for data analysis. A Dream 11 user needs to select the right mix of players to maximize his/her points, and thereby win some cash rewards. The paper describes a retrospective approach to team selection using the real-world data collected from Player performances in the last 10 matches, to propose a Dream 11 Fantasy team for the upcoming match. The technique used is Integer Programming, implemented using the Gurobi library in Python. The team selection problem has also been analyzed through the lens of Markowitz Optimization, which is mostly used to select stocks in a financial portfolio. The concept of risk aversion has been applied to penalize inconsistent performances, as risk taking and risk averse users might want to bet on different odds for the same match.

Keywords— Integer Programming, Binary Optimisation, Team Selection, Cricket

I. INTRODUCTION

Dream11 is a fantasy sports platform based in India that allows users to play fantasy cricket, hockey, football, kabaddi and basketball. It started in the year 2012 and has become the first Indian gaming company to enter the ‘Unicorn Club’ with a valuation of more than USD 1 billion. It has acquired over 70 million users and has achieved the remarkable feat of over 1 million concurrent users when IPL cricket matches are live. To summarize its core business model, the company charges a small amount to participate in fantasy games. It uses the proceeds from user subscription fees to build a corpus which is used to reward the users who build the most successful team. A user has to build his own team for an upcoming match subject to a budget constraint. He/she is required to select players from a pool drawn from both the teams participating in the match with each player bearing a certain cost. A minimum number of players must be selected from each team. The constraints in terms of the number of batsmen, bowlers, all-rounders and wicketkeepers must also be met. Based on the performance of the players thus selected in the match, each player is given a score at the end of the match. The user whose fantasy team gets the highest scores wins the cash reward.
Some people think of fantasy sports as a game which essentially entails a sophisticated form of betting which is unlawful on Indian soil. Petitions have been filed in courts to challenge the legality of the fantasy sports business but the Supreme Court of India dismissed the petitions adjudging that fantasy games and betting can’t be treated at par as they require a lot of domain knowledge and analytical skills to participate.

This study aims to provide a retrospective approach to Dream 11 team selection using Integer Programming. The problem statement is formulated as a classic Knapsack problem, where the decision variable is binary. The optimization has been achieved through the Gurobi library in Python.

The paper has been organized as follows: - Section I gives a brief introduction about Dream 11, followed by introduction of the problem statement and the tools used. Section II gives an overview of some of the work done by researchers with respect to player selection in Cricket. Section III defines the problem statement in a greater detail, and introduces us to decision variables and the constraints. A novel approach of using the Markowitz approach (mostly applied to optimize the selection of a financial portfolio) to recalibrate the optimization model, by penalizing the more inconsistent players is introduced. Section IV analyses the proposed solution with respect to India-England Series.
Section V discusses some of the limitations of this study and further course of work.

II. RELATED WORK

The usage of Analytics and other mathematically intensive approaches have been quite limited in Cricket. While the usage of technology has vastly increased over the last decade, usage of Analytics and optimisation techniques have picked up pace only in the last few years.

In the past, many ranking methods have been used for optimal team selection. The ranking methods used specific performance indices such as batting averages, strike rates, number of fours and sixes for evaluating a batsman. Some of these works are of Lemmer [1], Boorah and Mangan [2]. Kimber and Hansford [3], Damodaran [4] proposed a method to calculate batting averages when a player remains not out. Ovens and Bukiet [5], Swartz et.al.[6] have explored mathematical models to optimize the batting order.

When it comes to quantifying bowling performances, Lemmer [7] has proposed a method to calculate bowling averages and bowling strike rates. Lewis [8] have proposed metrics to evaluate performances of players using Duckworth/Lewis percentage values. In the past, many Fantasy league players would simply choose a team, by using these ranking metrics. However, with the introduction of budget and other constraints, a simple ranking method would not have sufficed. Over the years, many researches have used different Machine Learning and optimisation techniques to help fantasy league users in selecting an appropriate team.

In this section, we explore some of the works done by researchers using Machine Learning approaches. Farhana et.al. [9] have proposed a Support Vector Machine (SVM) based solution that generates a ranked list of players based on their abilities, from which a user can select a playing 11. C Deep Prakash [10] has proposed a Random Forests based Recursive Feature estimation algorithm to rank the players. The features that were passed into the Random Forest algorithm were a combination of customised metrices that accounted for both IPL and T20 International matches. These metrices so constructed, removed bias against the players who had been playing for a longer time. Once the performance evaluation metrices were selected on the basis of feature importance, an MGA (Memetic Genetic Algorithm) was used to optimise the team selection for IPL Teams.

Sharp et.al. [11] have formulated an Integer programming method for T20 team selection as a maximisation problem, wherein the decision variable indicates whether a certain player having a specific skill is selected or not. The coefficients for different skills such as batting, bowling and fielding were normalised using a method that utilized certain indices established indices such as averages and strike rates for batsmen and bowlers using a combination of different performance statistics.

Ahmed et.al. [12] have represented player selection as an evolutionary multi-objective optimisation algorithm (NSGA-II [13]). In the later parts of their paper, after taking into account the trade-offs between batsmen and bowlers using Dynamic optimisation, they have proposed a set of high-performance teams that can provide a basis for team selection during IPL Player Auctions.

III. METHODOLOGY

As discussed briefly in Section I, a user gets to build his own team for an upcoming match. From a pool of 30 players (inclusive of both teams), a team of 11 players need to be selected subject to budget and player constraints. Every player has an associated cost. A good player has a high cost and vice versa. Selecting all the good players will exhaust the budget and the team may not be complete. In order to have a balanced team, the Dream11 also imposes restrictions on number of batsmen, bowlers, all-rounders and wicket-keepers.

The objective of any user is to pick a team that maximizes his expected team score. Since the objective of this project is to formulate an optimization problem and solve it, we remove the element of prediction from our model and base our player selection on the mean score and standard deviation of his previous matches. The problem resembles a classic Knapsack problem where the decision variable is binary with $X_{jk} \in \{0,1\}$.

The decision variable not just indicates the selection of a particular player but also indicates the role in which he is selected.

III.I Decision Variables

Any player j can only be selected as a normal player, a vice-captain, a captain or be not-selected at all. We denote the role of the player by adding another subscript k.

$X_{j1} = 1$ if player j is selected as a normal player
$X_{j2} = 1$ if player j is selected as vice-captain
$X_{j3} = 1$ if player j is selected as captain
$X_{jk} = 0$ for all k∈{1,2,3} if player j is not selected

III.II Objective Function

We take the scores of each player Sj as information given to us. The optimization problem at hand involves finding a solution set that maximizes the total team score subject to team, role and player type constraints.

The objective function can be written as:

$$Max \sum_{j=1}^{n} (S_j * X_{1j} + 1.5 * S_j * X_{2j} + 2 * S_j * X_{3j})$$

III.III Constraints

III.III.I Budget Constraint

Each user is given an initial budget to pick players each of whom has a price tag. The players that perform well have a higher price tag. This budget constraint ensures that users cannot select all the best players as their budget is limited.

If Pj denotes the price of the jth player, the budget constraint can be written as:

$$\sum_{i=1}^{m}\sum_{j=1}^{n} P_j * X_{ij} \leq 100$$

III.III.II Role Constraint (Captain, Vice-Captain, Normal, or not selected)

Player can be only chosen as captain or vice-captain or normal or not selected at all.

$$\sum_{i=1}^{m} X_{ij} \leq 1 \qquad \forall j \in \{1,2,3 \dots, n\}$$

Total normal players have to be 9 in number, 1 captain has to be selected and 1 vice-captain has to be selected i.e. $b = [9,1,1]$

$$\sum_{j=1}^{n} X_{ij} = b_i \qquad \forall i \in \{1,2,3\}$$

III.III.III Team Constraint

A user is not allowed to build a fantasy team using players from only one team. A maximum of seven players can be selected from a team.

$$X_1 team_p{}' + X_2 team_p{}' + X_3 team_p{}' <= 7$$

p=1,2

III.III.IV Player Type Constraint

Dream11 platform specifies different ranges for the number of batsmen, bowler, all-rounder and wicket-keepers that a fantasy team must have. The required ranges are as follows:

Table 1. Player types and their selection range

	Constraint Type			
	Batsman	**Bowler**	**All-rounder**	**Wicket-Keeper**
Selection Range	3 – 6	3 – 6	1 – 4	1 – 4

The same can be expressed mathematically as follows:

$$lb_t <= X_1 type_t{}' + X_2 type_t{}' + X_3 type_t{}' <= ub_t$$

$$t = 1,2,3,4 \quad lb = \begin{bmatrix} 3 \\ 3 \\ 1 \\ 1 \end{bmatrix} ub = \begin{bmatrix} 6 \\ 6 \\ 4 \\ 4 \end{bmatrix}$$

Here **type** and **team** are the matrices which contain which player belongs to which type and nation. If a player is of type ‘batsmen’ then that particular column will take value of 1 and all others take a value of 0, a sample can be seen from the table below. The same logic follows for the other type of players as well.

Table 2 illustrates the player type matrix for nine players with the following type information.

P1->Bat, P2->WK, P3->Bowl, P4->Bowl, P5->Bat, P6->Bat, P7->Bat, P8->Bowl, P9->AR

Table 2. Player type matrix illustration

Player Type	Player Selection								
	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**
Bat	1	0	0	0	1	1	1	0	0
Bowl	0	0	1	1	0	0	0	1	0
All-rounder	0	0	0	0	0	0	0	0	1
Wicket-keeper	0	1	0	0	0	0	0	0	0

Likewise, the team information for the following sequence is depicted in Table 3.

P1->Team1, P2->Team1, P3->Team2, P4->Team1, P5->Team2, P6->Team2, P7->Team2, P8->Team2, P9->Team1

Table 3. Player team matrix illustration

Player Type	Players Selection								
	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**
Team1	1	1	0	1	0	0	0	0	1
Team2	0	0	1	0	1	1	1	1	0

III.IV Mean Variance Improvement

For the above Integer Programming formulation, the optimization approach is deterministic, meaning we assume the next time the score of each player will be equal to the mean performances of the previous matches.

However, that is seldom the case. The variability of the score also needs to be considered. We need to punish players with very large variance as large variance shows that the player is very inconsistent with the scores.

For this particular optimization, we use a Mean-Variance Optimization. Mean-variance analysis is the process of weighing risk, expressed as variance, against expected return.

This is a risk hedging technique commonly used in investment strategies to come up with portfolios with equities which have less variance and are also independent of each other.

If we want to incorporate mean-variance analysis, the objective function needs to be added with a term that punishes variance.

New Objective Function

Maximize

$$\sum_{j=1}^{n}\left(S_j * X_{1j} + 1.5 * S_j * X_{2j} + 2 * S_j * X_{3j}\right) - \lambda\left(\sum_{j=1}^{n}\left(X_{1j} + X_{2j} + X_{3j}\right)\sigma_j\right)$$

The σ_j term is the standard deviation of a player j and $\sum_{j=1}^{n}(X_{1j}+X_{2j}+X_{3j})$ makes sure that we consider all the types of choices of the player (whether normally selected or selected as a vice captain or selected as a captain). Here the value of λ depends on how risk averse or risk taking the user of the fantasy league is. The higher the value of λ, the more risk averse the user is and vice versa.

IV. RESULTS AND DISCUSSION

We begin with a set of 30 players for which we calculate the mean and variance of scores using the historical scores for last 10 matches. We are also given the information on their roles as batsman, bowler, wicket-keeper and all-rounder and their prices.

Given this information, we formulate this optimization as an integer programming problem concerning only the mean score of players. We then propose to improve upon this selection by punishing the uncertainty in performance through the inclusion of variance into the objective function along with a risk aversion factor, λ that can be calibrated based on user risk preferences. Our decision variable is a (30x3) matrix which indicates whether a player is selected in a particular role or unselected. We can easily derive the team matrix (30x2) and player type matrix (30x4), and write down the constraints as indicated in the previous section

Case 1. Inconsistent performance not penalized
In the first optimization, we keep the risk aversion parameter at zero to ignore the variance term. This means the selection of players is only based on their average scores and uncertainty of scores i.e. inconsistent performance is not penalized.

Gurobi Output with risk aversion = 0

```
The Captain of the Team is V Kohli(BAT) from INDIA
The Vice Captain of the Team is KL Rahul(BAT) from INDIA

Other Players:-

R Sharma(BAT) from INDIA
J Root(BAT) from ENGLAND
J Roy(BAT) from ENGLAND
J Bumrah(BOWL) from INDIA
L Plunkett(BOWL) from ENGLAND
J Archer(BOWL) from ENGLAND
U Yadav(BOWL) from INDIA
J Buttler(WK) from ENGLAND
K Jadav(ALL) from INDIA
```

Case 2. Inconsistent performance penalized

We improve the optimization model by including a penalty for inconsistent performance. We calculate the standard deviation of scores for previous 10 matches as the measure of inconsistency in performance and include in the objective function by coupling it with a risk aversion factor. Our model is able to reshuffle the players and pick players with less inconsistent performance.

Gurobi output with risk aversion = 1

```
The Captain of the Team is V Kohli(BAT) from INDIA
The Vice Captain of the Team is KL Rahul(BAT) from INDIA

Other Players:-

J Denly(BAT) from ENGLAND
J Roy(BAT) from ENGLAND
M Ali(ALL) from ENGLAND
J Bumrah(BOWL) from INDIA
L Plunkett(BOWL) from ENGLAND
B Kumar(BOWL) from INDIA
J Archer(BOWL) from ENGLAND
U Yadav(BOWL) from INDIA
```

Case 3. Sensitivity Analysis

Often, the team selection depends on the playing conditions. The team management would like to shuffle the team composition depending on the type of pitch, the match is being played on. At times, the team management would like to give rest to some of their high performing players, so that they don't get burnt out by playing too many matches. Doing so ensures that they are physically and mentally prepared for critically intense matches. On the other hand, the Team Management would like to drop under-performing players.
While the scenarios can be many, the scope of this report is limited to 3 scenarios only.

Scenario 1. Sometimes, the pitch on which the match will be played on, can be curated to become more batsman friendly. The team management would like to include more batsman in the team. In such cases, minimum requirement for the number of batsmen, in a team can be increased.

A sample output is shown below:

```
The Captain of the Team is V Kohli(BAT) from INDIA
The Vice Captain of the Team is KL Rahul(BAT) from INDIA

Other Players:-

J Denly(BAT) from ENGLAND
R Sharma(BAT) from INDIA
J Root(BAT) from ENGLAND
J Roy(BAT) from ENGLAND
L Plunkett(BOWL) from ENGLAND
J Archer(BOWL) from ENGLAND
U Yadav(BOWL) from INDIA
J Buttler(WK) from ENGLAND
K Jadav(ALL) from INDIA
```

Scenario 2. At times depending on weather conditions, the team which is batting second, is at a disadvantage. The Dew on the pitch ensures that the ball experiences unpredictable swings. In this case, the bowling team should include more bowlers in the team.


```
The Captain of the Team is V Kohli(BAT) from INDIA
The Vice Captain of the Team is KL Rahul(BAT) from INDIA

Other Players:-

R Sharma(BAT) from INDIA
J Root(BAT) from ENGLAND
J Bumrah(BOWL) from INDIA
L Plunkett(BOWL) from ENGLAND
B Kumar(BOWL) from INDIA
J Archer(BOWL) from ENGLAND
U Yadav(BOWL) from INDIA
J Buttler(WK) from ENGLAND
K Jadav(ALL) from INDIA
```

Scenario 3. The team management has decided to rest V Kohli, who has been the choice for 'Captain' for both risk-taking and risk-averse Dream 11 participants. Who shall be the new Captain?

```
The Captain of the Team is KL Rahul(BAT) from INDIA
The Vice Captain of the Team is R Sharma(BAT) from INDIA

Other Players:-

J Root(BAT) from ENGLAND
J Roy(BAT) from ENGLAND
J Bumrah(BOWL) from INDIA
L Plunkett(BOWL) from ENGLAND
J Archer(BOWL) from ENGLAND
U Yadav(BOWL) from INDIA
D Karthik(WK) from INDIA
J Buttler(WK) from ENGLAND
K Jadav(ALL) from INDIA
```

It's important to note that, the team selection depends on the risk-taking appetite of a Dream 11 User. A user, who is willing to take less risk may try to go for highly consistent players, resulting in a low overall score. It's important the test the robustness of our proposed method in terms of risk-reward trade off.

To test the robustness of our model, we create a random set of scores to see how well our model performs on a set of randomized data. We check the players selected, expected score, observed scores and the percentage difference between the expected and the observed for different values of **λ**.

Table 4. Team Composition per Risk Aversion Index of a Dream 11 User

λ	**Players Selection**									
	VK	SD	JD	RS	JR	KR	JY	JV	AH	E
0										
1										
2										
3										
4										
5										
6										
7										
λ	**Players Selection**									
	BS	HP	TC	M	JB	LP	BK	KY	MS	Y
0										
1										
2										
3										
4										
5										
6										
7										
λ	**Players Selection**									
	AR	C	JA	UY	DK	MD	JB	JU	KJ	R
0										
1										
2										
3										
4										
5										
6										
7										

Legend	Not Selected	Selected	Captain	Vice-Captain

Table 5. Expected and Actual Total Scores per Risk Aversion Index

λ	**Scores**		
	Expected Scores	**Actual Scores**	**%Difference**
0	4127	4408	6.8%
1	3909	4095	4.8%
2	3809	3909	2.6%
3	3809	3909	2.6%
4	3701	3896	5.3%
5	3600	3732	3.7%
6	3600	3732	3.7%
7	3388	3525	4.0%

From a pool of 30 players from India and England, a team of 11 players is chosen based on the risk-aversion index (0 being least risk averse to 7 being highly risk averse). The difference between the expected score, as predicted by the model and the observed score is then noted.

We observed that highest possible value for our case actually comes with **λ=0**. The difference decreases first with **λ** =3 and then starts to slowly increase again and the expected score starts to drop faster. This is because after a certain point of time, the punishment term for variance starts to dominate. Thus, for our case a nice spread comes around **λ** of 2 or 3 with a high enough score and a low percentage difference.

Even though in our case the best possible outcome comes when **λ** =0 for later matches we may not turn out so lucky. Here the observed value is greater than the expected, which is why the higher variation is in our favour here. However, in later matches if the expected is larger than the observed, the high variation will give us larger loss. Thus, the risk averse way is to minimize the difference between the expected and the observed and thus making the expected scoring a bit more accurate which gives us a clear idea of whether with the expected score, we should invest in the particular match or not.

V. CONCLUSION AND FUTURE SCOPE

The paper presents a novel method of using Integer Optimisation using Gurobi to predict an optimal team for a

Dream 11 user. The model is recalibrated to take into account the risk-aversion of a Dream 11 User and is seem to outperform Machine Learning and Integer programming models, that focus only on ranking metrics and/or overall scores.

However, the present method focusses only on using the points system outlined by Dream 11 to calculate scores, and is therefore biased against new players. A further improvement could involve using exponential weighted averages for scores, with the last match being given the highest weight and the weights decrease as the matches get older. This would remove the bias against newer players. It would also ensure that the in-form players are selected, instead of out-of-form players with an illustrious record.

Further improvements, could also involve introducing additional variables to this model such as the pitch reports, weather and dew data, which play a major role in deciding the composition of the team.

REFERENCES

[1] Lemmer HH. A measure for the batting performance incricket players.SAfr J Res Sport Phys Educ Recreation26(1):**55–64, 2002**

[2] Borooah, V. K., and Mangan, J E, "The„Bradman Class‟: An Exploration of Some Issues in the Evaluation of Batsmenfor Test Matches, 1877–2006.", Journal of Quantitative Analysis in Sports, 6 (3), Article **14, 2010.**

[3] Kimber, A C and Hansford, A R, "A Statistical Analysis of Batting in Cricket," Journal of Royal Statistical Society, **156, pp 443 –455, 1993.**

[4] Damodaran, U, "Stochastic Dominance and Analysis of ODI Batting Performance: The Indian Cricket Team, 1989-2005," Journal of Sports Science and Medicine, **5, pp 503 –508, 2006.**

[5] Ovens M and Bukiet B. A mathematical modelling approachto one-day cricket batting orders.J Sports Sci Med5: 495–502.Preston I and Thomas J. Batting strategy in limited overscricket.Statistician49(1): **95–106, 2006**

[6] Swartz TB, Gill PS, Beaudoin D and de Silva BM (2006). Opti-mal batting orders in one-day cricket.Comput Opns Res33:**1939–1950, 2006**

[7] Lemmer HH (2006). A measure of the current bowling performancein cricket.S Afr J Res Sport Phys Educ Recreation28(2): **91–103, 2006**

[8] Lewis, A., "Extending the Range of Player-Performance Measures in One-Day Cricket," Journal of Operational Research Society, **59, pp 729-742, 2008.**

[9] Farhana Siddiqui, Hasan Phudinawala, Chetan Davale, Soham Pawar," Innovative Idea for Playerelection usingSupport Vector Machine(SVM)" International Journal of Computer Sciences and Engineering, Vol. **7**, Issue-**4**, pp.**841-843**, **2019**

[10] C. Deep Prakash "A New Team Selection Methodology using Machine Learning and Memetic Genetic Algorithm for IPL-9", International Journal ofElectronics, Electrical and Computational System **Volume 5 –Issue.4, April 2016**

[11] Sharp, Gary & Brettenny, Warren & Gonsalves, John & Lourens, Michelle & Stretch, R. Integer optimisation for the selection of a Twenty20 cricket team. Journal of the Operational Research Society. 62. 10.1057/jors.2010.122, **2011**

[12] Ahmed, F., Deb, K., & Jindal, A. .Evolutionary multi-objective optimization and decision making approaches to cricket team selection.In Proceedings of the Second International Conference on Swarm, Evolutionary, and Memetic Computing. Berlin, Heidelberg: Springer-Verlaghttp://dx.doi.org/10.1007/978-3-642-27242-4_9, **2011**

[13] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan. A fast and elitist multi-objective geneticalgorithm: NSGA-II.IEEE Transactions on Evolutionary Computation, 6(2):182–197, **2002**.

[14] Nihar Patel, Mrudang Pandya," IPL Player's Performance Prediction" International Journal of Computer Sciences and Engineering, Vol. **7**, Issue-**5**, pp.**478-481**, **2019**

[15] Harsha Perera, Jack Davis & Tim B. Swartz Optimal lineups inTwenty20 cricket, Journal of Statistical Computation and Simulation,86:14,2888-2900,**2016** DOI:10.1080/00949655.2015.1136629

[16] Taha HA (2003).Operations Research, An Introduction. 7th edn,Prentice Hall: Upper Saddle River, New Jersey

[17] Lakkaraju, P., and Sethi, S., "Correlating the Analysis of Opinionated Texts Using SAS®Text Analytics with Application of Sabermetrics to Cricket Statistics," Proceedings of SASGlobal Forum 2012, **136, pp 1-10, 2012.**

AUTHORS PROFILE

Mr. Saurav Singla is a Senior Data Scientist and a Machine Learning Expert. He has fifteen years of comprehensive experience in statistical modeling, machine learning, natural language processing, deep learning, and data analytics. He has a Master of Science from the University of Westminster.
He has been recognized for maximizing performance by implementing appropriate project management tools through analysis of details to ensure quality control and understanding of emerging technology.
Outside work, Saurav volunteers his spare time for helping, coaching, and mentoring young people in taking up careers in the data science domain. He has created two courses on data science, with over twenty thousand students enrolled in it. He regularly authors articles on data science.

Mr Swapna Samir Shukla is an independent researcher in Machine Learning and Data Science. He completed Bachelor of Technology from National Institute of Technology, Rourkela and has a Master of Science from National University of Singapore.
He has around 6 years of industry experience across various verticals such as Financial Services, Healthcare Analytics, working as a Software Engineer and Business Analyst.
Outside work, Samir researches about innovative solutions to tackle some business problems using his domain expertise and data science skills. His areas of interest include Recommender Systems, Financial Risk Analytics, Quantitative Risk Management.

