**HR Analytics Problem Statement**

* **HR Analytics**

**Objective**: The HR Department woke up to the utility of machine learning to try predictive analytics in identifying the employees most likely to get promoted

**Achievements**: Highly imbalanced data, treated with SMOTE method. A lot of insights were gathered by in depth Univariate and Multivariate analysis. The data was standardized and parameters were Hyper tuned by using Grid Search and Random Search. Evaluation done on the basis of F1 Score. Achieved: 92%

**Key Skills**: Scikit Learn, Recommender, Collaborative Filtering Algorithm

**Dataset Description**

|  |  |
| --- | --- |
| Variable | Definition |
| employee\_id | Unique ID for employee |
| department | Department of employee |
| region | Region of employment (unordered) |
| education | Education Level |
| gender | Gender of Employee |
| recruitment\_channel | Channel of recruitment for employee |
| no\_of\_trainings | no of other trainings completed in previous year on soft skills, technical skills etc. |
| age | Age of Employee |
| previous\_year\_rating | Employee Rating for the previous year |
| length\_of\_service | Length of service in years |
| KPIs\_met >80% | if Percent of KPIs(Key performance Indicators) >80% then 1 else 0 |
| awards\_won? | if awards won during previous year then 1 else 0 |
| avg\_training\_score | Average score in current training evaluations |
| is\_promoted | (Target) Recommended for promotion |

**Evaluation Metric**

The evaluation metric for this competition is F- Score.

**Predicting Upvotes - Problem Statement**

An online question and answer platform has hired you as a data scientist to identify the best question authors on the platform. This identification will bring more insight into increasing the user engagement. Given the tag of the question, number of views received, number of answers, username and reputation of the question author, the problem requires you to predict the upvote count that the question will receive.

**Dataset Description**

|  |  |
| --- | --- |
| **Variable** | **Definition** |
| ID | Question ID |
| Tag | Anonymised tags representing question category |
| Reputation | Reputation score of question author |
| Answers | Number of times question has been answered |
| Username | Anonymised user id of question author |
| Views | Number of times question has been viewed |
| Upvotes | (Target) Number of upvotes for the question |

**Evaluation Metric**

The evaluation metric for this competition is RMSE (root mean squared error)