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.

Data Dictionary

|  |  |
| --- | --- |
| **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)**