

Approach

1. First I imported the necessary libraries.
2. Read and viewed the data.
3. Checked for missing values.
4. Did variable Typecasting.
5. Generated features like age_range it contain the age group of the users, followers_range, views_range, category_range, video_range these features explain the popularity of the categories and the videos.
6. Encoded the categorical variables.
7. Tried various ML models out of which GBM performed the best.
8. Build a Gradient Boosting(GBM) model.
9. Did predictions and computed the r2 score.
10. Did hyperparameter tuning.
11. Again build a new GBM model with new hyperparameter values.
12. Did predictions and computed the r2 score for the new GBM model.
13. At last, converted the test data frame to a CSV file.