

## Approach

1. First I imported the necessary libraries.
2. Read and viewed the data.
3. Checked for missing values.
4. Did variable Typecasting.
5. Imputed the missing values present in "products\_purchased" Variable.
6. Extracted Features From "signup\_date" and "created\_at" Variables.
7. Dropped "id", "signup\_date" and "created\_at" Variables.
8. Tried various ML models out of which Random forest performed the best.
9. Built a Random forest (RF) model.
10. Did predictions and computed the f1 score.
11. Did hyperparameter tuning.
12. Again built a new RF model with new hyperparameter values.
13. Did predictions and computed the f1 score for the new RF model.
14. At last, converted the test data frame to a CSV file.