My approach to predict the engagement score of the video on the user level.

1. I explored the Train Data and Test Data ,and found that there were **6 numerical columns** and **2 – categorical columns** in given data.

2. I have stored the engagement score columns in the target variable.

3. Since I am going to use all **tree based model** therefore I **didn't scaled the numerical columns.**

4. For **categorical columns** , I **encoded** them , both training and testing data, using **label encoder**.

5. Then I created 5-folds of training data for **cross – validation and hyperparameter tuning of model**.

6. Firstly I used **RandomForestRegressor()** model with 700 estimators which has given 0.2945 score on public leader board.

7. Then , I have trained **XGBoost , LightGBM , CatBoost** regressor model.
8. I have tuned hyperparameter of all the models by using **RandomizedSearchCv.**
9. Then, I have **ensembled all the model's prediction** using **average voting ensemble** and I have also used stacking ensemble technique but it's score was less than average voting ensemble , therefore I have kept the prediction of average voting ensemble.

Fig -: Average voting ensemble model