**About data**

* Imbalance data: 1703:262
* No of missing values : 134(removed as its less than 7%)
* All variable distribution is somewhat normal, expect for ‘**weekly\_mins\_watched’** as it contain outlier.

Columns taken for modelling: ['gender', 'age', 'no\_of\_days\_subscribed', 'multi\_screen', 'mail\_subscribed', 'weekly\_mins\_watched', 'minimum\_daily\_mins', 'maximum\_daily\_mins',

       'weekly\_max\_night\_mins', 'videos\_watched', 'maximum\_days\_inactive', 'customer\_support\_calls', 'churn']

AIM:

To build churn prediction model

Approach:

* missing value Treatment
* data variable processing and grouping
* splitting data into features and label
* splitting data in test and train(1:4)
* model evaluation hyperparameter tuning
* Explainable AI dashboard

Results:

* Xgboost shows best prediction
* Accuracy on Train Data : 1.0
* Accuracy on Test Data: 0.9088471849865952
* Overfit