**Telecom Churn - ML Group Case Study**

**Chiranshu:**

Data analysis, cleaning and featuring generation

1. Filtered high value customers
2. Created churn flag
3. Removed month 9 attributes
4. Analysis of feature columns
5. Derived new columns

Data Preparation

1. Test train split
2. Scaling of variables StandarScaler()
3. Applied Imbalanced data technique (SMOTE)

Total features till now 107

PCA

1. Applied PCA for dimensionality reduction
2. 55 components explain 96% variance

Tried following algorithms

1. Logistic regression
2. Logistic regression on after oversampling data
3. RandonForest with default hyperparameters
4. XGboost( with default hyperparameters)

Yet To do:

1. SVM with default parameters
2. SVM with hyper parameters tuning
3. Decision tree classifier
4. Check for other methods