**Razorpay Case Study: Telco Customer Churn Prediction:**

A) Description of the pipeline and design choices:

The pipeline used for this example consists of 9 steps

1. Problem Definition
2. Data Collection
3. Exploratory Data Analysis (EDA)
4. Feature Engineering
5. Train/Test Split
6. Model Evaluation Metrics Definition
7. Model Selection, Training, Prediction and Assessment
8. Hyperparameter Tuning/Model Improvement
9. Application Deployment Using Flask as API

B) Performance evaluation of the model

For performance assessment of the chosen models, various metrics are used:

1. Confusion matrix- Displays a grid of true and false predictions compared to actual values
2. **Accuracy score-** Shows the overall accuracy of the model for training set and test set
3. **Precision-Recall- Shows** the diagnostic ability by comparing false positive rate (FPR) and false negative rate (FNR) for different thresholds of class predictions. It is suitable for data sets with high class imbalances (negative values overrepresented) as it focuses on precision and recall, which are not dependent on the number of true negatives and thereby excludes the imbalance
4. **F1 Score-**Builds the harmonic mean of precision and recall and thereby measures the compromise between both.

* **C)** Discussion of how to scale up the pipeline to process tens of billions of data points

# Data collection

# Data storage

# Analytics engine

# Monitoring and Quality

* **D)** Discussion of future work

Data is the currency of the future, and it is growing exponentially, the world continues to become more connected every second and in 2020, every human on average created at least 1.7 MB of data per second. AI and automation can help manage this mega data. Through AI, machine learning, and automation, telecom companies can reduce the burden on manual intervention, manage volumes of data efficiently, and increase response speed, allowing them to put the large amount of data to good use.

Predictive analytics is an important edge for telecom companies. Some of the major use-cases of predictive analytics can be seen in answering questions like- how long before a consumer leaves, how to retain them, profiling consumers, developing products for financial inclusion and fraud prevention, based on what the market needs and what’s profitable, and set prices that offer a competitive edge. All of this helps telcos be on top of their game- instead of reacting to changes, it helps them predict those changes in advance and get the advantage of speed and agility. By forecasting demand, anticipating network load, and tailoring capacity, telcos would be able to offer more personalized, efficient bundles to consumers. The uncertainties of the past year are a testament to the need for such personalization - the requirements of the industry with respect to telecom services keep changing, a significant example being the shift from enterprise networks to residential networks in the hybrid working scenario.