**Telco Churn Prediction – Capstone Project**

**Problem Statement**

Customers in the telecommunications industry have the freedom to choose their service provider and can easily switch to a different company. As a result of the high level of competition in this market, 15–25% of customers switch providers annually. Since it costs 5–10 times as much to acquire a new customer as it does to keep an existing one, customer retention has become much more vital than customer acquisition.

For many established service providers, the most important thing is to keep their best, most profitable customers. Telecom providers may significantly lower customer churn by identifying and focusing on retaining those most at risk of leaving.

For this project, we will look at customer-level data from a major telecom company, build prediction models to find clients who are at risk and find the most telling signs of customer turnover.

**Your focus in this exercise should be on the following:**

* **Data** - Extract meaningful information from the data and predict whether the firm will go bankrupt.
* **Model** - To get an accuracy of 80% or higher using Logistic Regression, Random Forest, SVM, and Gaussian Naive Bayes.

**Timeline**

We expect you to do your best and submit a solution within a week.

**Deliverables**

Please share the following deliverables in a zip file.

A report (PDF) detailing:

* Description of design options and models.
* Performance evaluation of the model.
* The source code used to create the pipeline.

**Tasks/Activities List**

Your code should contain the following activities and analyses:

* Collect the data from the zip file linked here.
* Split the dataset into dependent and independent variables.
* Perform train, test split.
* Apply StandardScaler() to the train and test dependent variable.
* Fit the best parameters.
* Model Prediction
* Model Validation Statistics
* Create a FastAPI app
* Create routes for getting predictions
* Create a route for retraining the model

**Success Metrics**

Below are the metrics for the successful submission of this case study.

* The model's accuracy on the test data set should be 80% or higher.
* The model should be available through a FastAPI app
* The model should be retrainable using FastAPI

**Bonus Points**

* You can package your solution in a zip file with a README describing how to install and run the end-to-end pipeline.
* You can highlight your documentation skills by explaining how they benefit our organization.