Data preprocessing:

Since we already have the dataset, we can start preprocessing directly. We can start by looking at the data to get a clear idea about rows, columns and data types we are dealing with, and then clean it by handling missing values, check for consistency and if needed we can normalize or standardize it to help ensure the model’s performance.

Data Analysis:

After we have prepared the data, we can start analyzing it. We can do some visualizations that help to identify data trends like scatter plots to observe correlations, derive some insights and we can do regression analysis to see if there is a correlation between the rate of churn and another factor. At the end of this step, we should have a good idea about the reason for churn among our clients so next we should make use of predictive models to help if a suggested solution would give a satisfying result for the next 3 months.

Model Selection and Training:

Choose appropriate machine learning model, such as regression model. Split the dataset into training and test sets, then train the model using the training data. Evaluate model performance like accuracy and precision to determine how well the model predicts customer churn. Finally Interpret outputs to identify factors may cause churn. For example, if the model shows there is a specific service or demographic factor has a high impact on churn, this insight can inform us to take actions.

Deployment:

Finally, we can deploy the model to use for analyzing real-time data. After deployment, it’s essential to keep tracking the model performance to ensure it continues to perform well with real-time data.