Abstract

Customer churning has various problems in a number of different industries. Due to the loss of customers the companies not only have a lower revenue but there is a vast effect on the reputation and client number. Therefore, it has become crucial for businesses to anticipate customer churn. In this project, we want to create a model for predicting customer churn using artificially generated data.

The project will involve creating artificial data that closely resembles the traits of actual customer data. To create a dataset that accurately reflects the customer data, we will develop a framework for the generation of synthetic data. To make sure the synthetic data is appropriate for machine learning algorithms, it will be pre-processed and given engineered features. Using the simulated data, we will assess how well different machine learning algorithms predict customer churn. The model's performance will be contrasted with that of conventional models developed using real-world data. Later we will deploy the maximum accuracy machine learning model into a web application in a flask for companies to input data and predict customer churn prediction.

The project's findings will shed light on how well artificial data can be used to predict customer churn. Businesses will benefit from better customer retention strategies and more informed choices. Additionally, the framework developed for the project can be applied to numerous other applications where a lack of data is a bottleneck. The project's contribution goes beyond just predicting customer churn; it also offers a general method for creating synthetic data that can be used in other situations.