

Project Report: Developing and Deploying a Customer Churn Prediction Model

1.introduction:

The goal of the project was to develop a machine learning model to predict customer churn based on historical customer data.

2. Data Preprocessing:

- The dataset, named customer_churn_large_dataset.xlsx, contained information on customer attributes, interactions, and whether they churned or not.
- Initial data exploration revealed the structure of the dataset, column names, and data types.
- Missing data are checked if they are present or not.
- Categorical variables were encoded, and the dataset was split into training and testing sets.

3. Feature Engineering:

- Rescaling the features to have a similar scale, such as having a standard deviation of 1, to make sure the model considers all features equally.
- Transforming categorical features into a numerical representation.

4. Model Building:

- Three different machine learning algorithms were explored: Random Forest, Logistic Regression, and Neural Network.
- The models were trained and validated on the training dataset.
- Model performance was evaluated using various metrics, including accuracy, precision, recall, and F1-score.

5. Model Optimization:

- Hyper parameter tuning was performed to enhance model predictive performance.
- Techniques like randomized search were used to find optimal hyper parameters.

6. Model Deployment:

- The Flask framework was used to create a web application that exposes an API for model predictions.
- The trained model was pickled using joblib and loaded within the Flask app.
- An API endpoint was defined to handle incoming requests, pre-process data, make predictions, and return results.

Overall, the project helped me to understand the end-to-end process of developing and deploying a machine learning model for practical business use, highlighting the importance of data pre-processing, model selection, optimization, and making the model accessible through a user-friendly interface.