Predicting Customer Churn in a Telecom Company

This project utilised machine learning to predict customer churn in a telecom company, providing valuable insights for improving customer retention strategies. Future work includes further model optimisation and deployment for real-time predictions.

Library used:

* **pandas: data manipulation and analysis**
* **sklearn: machine learning model and evaluation metrics**

Steps:-

* Data collection and exploration

Imported customer data from csv file.

Explored dataset by displaying the first few rows, basic information, and summary statistics.

Checked for missing values.

* Data preprocessing

Filled missing values in the “TotalCharges” column and converted it to numeric format.

Applied one-hot encoding to transform categorical variables into numerical format.

* Feature engineering

Defined features and target variables.

Selected relevant columns for features and identified the churn column as the target.

* Model building

Split the data into training and testing sets.

Initialised and trained a Random Forest Classifier on the training set.

Made predictions.

* Model evaluation

Evaluated model performance using accuracy score, confusion matrix, and classification report.

Achieved an accuracy of approximately 80%, indicating the model’s effectiveness in predicting customer churn.