Project Title: Predicting Customer Churn for a Telecom Company

Project Overview: In this project, we aim to build a machine learning model that predicts which customers are likely to churn (i.e., cancel their subscription) for a telecom company. By identifying customers who are at risk of leaving, the company can take proactive measures to retain them, such as offering incentives or improving their customer service.

Project Goals:

1. Collect and preprocess a dataset of customer demographics, usage patterns, and churn status.
2. Train and evaluate different machine learning models, such as logistic regression, random forest, and support vector machines, to predict customer churn.
3. Select the best performing model and fine-tune its hyperparameters to optimize its performance.
4. Deliver a report summarizing the project findings and recommendations for the telecom company.

Data Collection and Preparation:

Collect data from the company's customer database, including demographics (age, gender, income), usage patterns (call duration, data usage), and churn status.
Clean and preprocess the data by handling missing values, scaling numerical features, and encoding categorical features.

Model Selection and Training:
Split the dataset into training and testing sets to evaluate the performance of the models.
Train and evaluate different machine learning models, such as logistic regression, random forest, and support vector machines, using cross-validation and performance metrics such as accuracy, precision, and recall.
Select the best performing model based on the evaluation results and fine-tune its hyperparameters using grid search or randomized search.

Results and Evaluation:

Evaluate the performance of the selected model on the test set and compare it to other models.
Analyze the model's feature importance to identify the most important factors that contribute to customer churn.
Summarize the project findings and provide recommendations for the telecom company to reduce customer churn.

Conclusion and Future Work:

Summarize the project and its outcomes, including the limitations and potential future work.
Discuss potential extensions of the project, such as predicting customer lifetime value or recommending personalized retention strategies.

Project Timeline:

Week 1: Data collection and preparation
Week 2: Model selection and training
Week 3: Model evaluation and fine-tuning
Week 4: Report writing and presentation

Project Deliverables:

Final report summarizing the project findings and recommendations
Code repository containing the machine learning models and preprocessing scripts
Project Resources:

Python programming language
Scikit-learn machine learning library
Telecom customer churn dataset from Kaggle
Project Team:

Project Budget:

No budget required as all resources are open source or provided by the company.