ML Intern Final Assignment

This assignment is designed to evaluate your ability to analyze data, build machine learning models, and provide business insights. Follow the instructions carefully and submit all deliverables through the provided Google Form link.

Assignment Title:

Predict Customer Churn and Build Insights for Retention

Scenario:
You are tasked with analyzing customer data for a subscription-based service provider to identify patterns, predict churn, and propose retention strategies. Additionally, build a recommendation engine to enhance customer engagement.

Tasks

1. Exploratory Data Analysis (EDA)

- Load the provided dataset and conduct a detailed EDA.
- Identify and handle missing values, if any.
- Perform univariate and bivariate analysis to understand patterns (e.g., churn distribution, relationship between tenure and monthly charges).
- Visualize important trends using graphs and charts.

2. Feature Engineering

- Create new features (e.g., average spend per month, usage patterns).
- Encode categorical variables and scale numerical ones.

3. Predictive Modeling

- Build and compare the performance of the following models to predict churn:
    - Logistic Regression
    - Decision Trees
    - Random Forest
- Evaluate models using metrics such as accuracy, precision, recall, and F1-score.
- Include a confusion matrix for the best-performing model.

4. Model Interpretation

- Use feature importance techniques (e.g., SHAP or feature importance plots) to interpret the results.
- Identify which factors contribute most to customer churn.

5. Recommendation Engine

- Build a collaborative filtering-based recommendation engine to suggest additional services/products to customers based on their usage patterns.
- Provide insights into how the recommendation system could improve customer retention.

6. Deployment Simulation

- Create a Flask API to expose your churn prediction model. The API should:

    - Accept customer details as input.
    - Return the churn probability.
- Include instructions to test the API.

7. Submission Deliverables

1. **Code**: Submit all scripts for data preprocessing, model training, evaluation, and deployment.
2. **Jupyter Notebook**: Include explanations and visualizations in a well-documented notebook.
3. **Report (PDF)**: Summarize:
    - EDA findings.
    - Model comparison and performance metrics.
    - Business insights and recommendations.
4. **API Instructions**: Provide a README with steps to test the API locally.
5. **Google Drive Link**: Upload your entire project folder (code, notebook, and report) to Google Drive and include the link in the Google Form.

Dataset

Download the dataset for this assignment from the following link:

ML - Assignment Dataset

Submission Instructions

- Complete the assignment and upload all deliverables to Google Drive.
- Submit your project folder link in the Google Form below:

  https://forms.gle/AbNhcP1aoGYDMovv7

Note:

Leverage AI tools to enhance your work, but do not rely on instant solutions or generic outputs. Showcase your unique effort and problem-solving skills by providing well-thought-out and original answers. Your approach and execution matter the most in this assignment!