Dataset Proposal & Datasheet

Dataset: *Telco Customer Churn Prediction*

Dataset File: WA_Fn-UseC_-Telco-Customer-Churn.csv

Data Source: The dataset is publicly available and can be accessed on Kaggle --

https://www.kaggle.com/datasets/blastchar/telco-customer-churn

Dataset Abstract:

This dataset contains information on a telecommunication company's customers, including their demographic details, account information, services availed, and billing details. It is primarily used to predict customer churn, which indicates whether a customer has discontinued their subscription to the company's services. The dataset includes 7,043 customer records with 21 variables.

The objective of using this dataset is to analyze and predict customer churn based on factors like contract type, monthly charges, tenure, and the services customers subscribe to. Understanding churn patterns can help develop strategies for improving customer retention.

Justification of selecting this Dataset:

The "Telco Customer Churn Dataset" is well-suited for this project because it meets the size and variable requirements. With 7,043 rows and 21 variables, the dataset is compact yet comprehensive enough to support robust modeling. The dataset includes both character variables (e.g., 'gender', 'contract') and numerical variables (e.g., 'tenure', 'monthly charges'), making it suitable for exploratory analysis and machine learning algorithms.

Also, this dataset aligns well with the project's goals by providing the necessary features to predict customer churn and identify the key drivers behind it. By answering questions such as --

1. What factors most influence customer churn?
2. Which customer segments are most likely to churn?

the project will generate actionable insights for reducing churn rates and enhancing customer retention strategies.

Datasheet:

Name	Description
customerID	Unique identifier for each customer
gender	Gender of the customer (Male/Female)
SeniorCitizen	Whether the customer is a senior citizen (1 = Yes, 0 = No)
Partner	Whether the customer has a partner (Yes/No)
Dependents	Whether the customer has dependents (Yes/No)
tenure	Number of months the customer has stayed with the company

Name	Description
PhoneService	Whether the customer has a phone service (Yes/No)
MultipleLines	Whether the customer has multiple lines (Yes/No/No phone service)
InternetService	Type of internet service availed by the customer (DSL/Fiber optic/No)
OnlineSecurity	Whether the customer has online security add-on (Yes/No/No internet service)
OnlineBackup	Whether the customer has online backup add-on (Yes/No/ No internet service)
DeviceProtection	Whether the customer has device protection add-on (Yes/No/ No internet service)
TechSupport	Whether the customer has tech support add-on (Yes/No/ No internet service)
StreamingTV	Whether the customer has streaming TV service (Yes/No/ No internet service)
StreamingMovies	Whether the customer has streaming movie service (Yes/No/ No internet service)
Contract	Customer's contract type (Month-to-month, One year, Two year)
PaperlessBilling	Whether the customer has paperless billing (Yes/No)
PaymentMethod	Customer's payment method (Electronic check, Mailed check, Bank transfer, Credit card)
MonthlyCharges	The amount charged to the customer monthly
TotalCharges	The total amount charged to the customer
Churn	Whether the customer churned (Yes/No)