Model Documentation for Telco Customer Churn Prediction

Project Overview

- Objective: Predict customer churn to enhance retention strategies.
- Stakeholders: Telecommunications company's marketing and customer service departments.

Dataset Overview

- Source: Data from a telecommunications company.
- Entries: 7,043 customers.
- Features: 21 attributes including customer demographics, account information, and services.

Data Preprocessing and Feature Engineering

- Cleaning: Addressed missing values and errors.
- Feature Engineering: Created features aimed at improving model interpretability and performance, such as encoding categorical variables and normalizing numerical features.

Model Development

- Models Tested: Logistic Regression, Random Forest, and Gradient Boosting Machines.
- Model Selection: Logistic Regression was chosen for initial benchmarking due to its simplicity and effectiveness.

Model Evaluation and Interpretation

- Key Features: Senior Citizen, Partner, Dependents, Internet Service were highly influential in predicting churn.
- Challenges: Limited compute power affected extensive hyperparameter tuning.
- Recommendations: Future iterations should explore more extensive grid search techniques with increased computational resources.