Model Metrics File

Model Overview

- Models Evaluated:
    1. Logistic Regression
    2. Random Forest
    3. Gradient Boosting Machines

Logistic Regression Model

- Model Name: Telco Customer Churn Logistic Regression
- Version: 1.0
- Training Date: 4/29/24
- Validation Method: 70-30 train-test split
- Hyperparameters:
    - Solver: liblinear
    - Regularization: L2
- Performance Metrics:
    - Accuracy: 82%
    - AUC-ROC: 0.862

Random Forest Model

- Model Name: Telco Customer Churn Random Forest
- Version: 1.0
- Training Date: 4/29/24
- Validation Method: 70-30 train-test split
- Hyperparameters:
    - Number of Trees: 400
    - Max Depth: 10
    - Min Samples Split: 5
- Performance Metrics:
    - Accuracy: 0.81
    - AUC-ROC: 0.86

Gradient Boosting Machines Model

- Model Name: Telco Customer Churn GBM
- Version: 1.0

- Training Date: 4/29/24
- Validation Method: 70-30 train-test split
- Hyperparameters:
    - Learning Rate: 0.096
    - Number of Trees: 154
    - Max Depth: 3
- Performance Metrics:
    - Accuracy: 0.81
    - AUC-ROC: 0.86

Comparison and Final Selection

- Discussion: The Logistic Regression model provided a solid baseline with high AUC-ROC. Random Forest and GBM were explored to see if improvements could be made, but after Random Search of 10 iterations, they were slightly less accurate than LR mode.
- Selected Model: Logistic Regression is the model of choice, until we can try ML models with higher compute power.