Supervised Learning - Building an Intervention System Project Rubric

Overview

This rubric is here to help you understand the expectations for how your project will be evaluated. It is the same rubric that the person evaluating your project will use. You should look at the rubric before you begin working on this project and before you submit it.

Criteria	Meets Specifications
Classification vs Regression	
Classification vs Regression	Student is able to correctly identify which type of prediction problem is required and provided reasonable justification.
Exploring the Data	
Data exploration	Student response addresses the most important characteristics of the dataset and uses these characteristics to inform their decision making. Important characteristics must include: ● Number of data points ● Number of features ● Number of graduates ● Number of non-graduates ● Graduation rate
Preparing the Data	
Identify feature and target columns	Code has been executed in the iPython notebook, with proper output and no errors.
Preprocess feature columns	Code has been executed in the iPython notebook, with proper output and no errors.

Split data into training and test sets	The training and testing sets are randomly generated from entire dataset. Data is split using a 75%/25% training/testing split, or 95 data points are kept for testing.

Training and Evaluating Models

Model Application	The pros and cons of application for each model is provided with reasonable justification why each model was chosen to explore.
Model Performance Metrics	All the required time and F1 scores for each model and training set sizes are provided within the chart given. The performance metrics are reasonable relative to other models measured.

Choosing the Best Model

Choosing the Optimal Model	Justification is provided for which model is the best by directly & explicitly comparing and contrasting the numerical computational cost and accuracy for each model from the recorded data.
Describing the Model in Layman's Terms	Student is able to clearly and concisely describe how the optimal model works in laymen terms to someone what is not familiar with machine learning nor has a technical background.
Model Tuning	The final model chosen is correctly tuned using grid search with at least one parameter using at least three settings. If the model does not need any parameter tuning it is explicitly stated with reasonable justification.
Tuned F1 Score	The F1 score is provided from the tuned model and performs *approximately* as well or better than the default model chosen.

Quality of Code

Functionality	Code reflects the description in the documentation.