

# 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
<b>Classification vs Regression</b>	
<b>Classification vs Regression</b>	Student is able to correctly identify which type of prediction problem is required and provided reasonable justification.
<b>Exploring the Data</b>	
<b>Data exploration</b>	Student response addresses the most important characteristics of the dataset and uses these characteristics to inform their decision making. Important characteristics must include: <ul style="list-style-type: none"><li>• Number of data points</li><li>• Number of features</li><li>• Number of graduates</li><li>• Number of non-graduates</li><li>• Graduation rate</li></ul>
<b>Preparing the Data</b>	
<b>Identify feature and target columns</b>	Code has been executed in the iPython notebook, with proper output and no errors.
<b>Preprocess feature columns</b>	Code has been executed in the iPython notebook, with proper output and no errors.
<b>Split data into training and test sets</b>	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.
<b>Training and Evaluating Models</b>	
<b>Model Application</b>	The pros and cons of application for each model is provided with reasonable justification why each model was chosen to explore.

<b>Model Performance Metrics</b>	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.
<b>Choosing the Best Model</b>	
<b>Choosing the Optimal Model</b>	Justification is provided for which model seems to be the best by comparing the computational cost and accuracy of each model.
<b>Describing the Model in Layman's Terms</b>	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.
<b>Model Tuning</b>	The final model chosen is correctly tuned using gridsearch 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.
<b>Tuned F1 Score</b>	The F1 score is provided from the tuned model and performs <i>approximately</i> as well or better than the default model chosen.
<b>Quality of Code</b>	
<b>Functionality</b>	Code reflects the description in the documentation.