

Assignment #3: Supervised Learning

INFO 629 - Drexel University

May 10, 2024

Instructions: This exercise will mimic what we have seen in the examples during our lectures on supervised learning. I have provided example code on working with several classes of these models using the ‘sklearn’ package. In addition to providing the models, the package provides the necessary code used for data-preprocessing (using the pipelines as shown in the examples) as well as the evaluation metrics.

In this exercise, we will work with an expanded dataset for the ‘will you wait’ problem. Your task is to run a variety of models and compare their performance metrics. Specifically, I encourage you to experiment with **alternate data pre-processing** pipelines.

Build your models using the provided training dataset and evaluate using the provided validation set. When you are happy with your results, please provide the following files with your submission:

- A document detailing your development process. What variable transformations and models did you try? What did you find worked well or failed miserably? Also provide your model’s performance metrics on the validation set.
- A saved version of your model pipeline that I can load and run on the (unseen) test set. Please adhere to the sklearn standard models and pipelines to make sure I’m able to run your code.

All individuals will be required to submit a write-up of their approach in a PDF formatted ‘{first_name}_{last_name}_HW3_Solutions.pdf’, along with a zip file of code used in the format ‘{first_name}_{last_name}_HW3_Code.zip’.