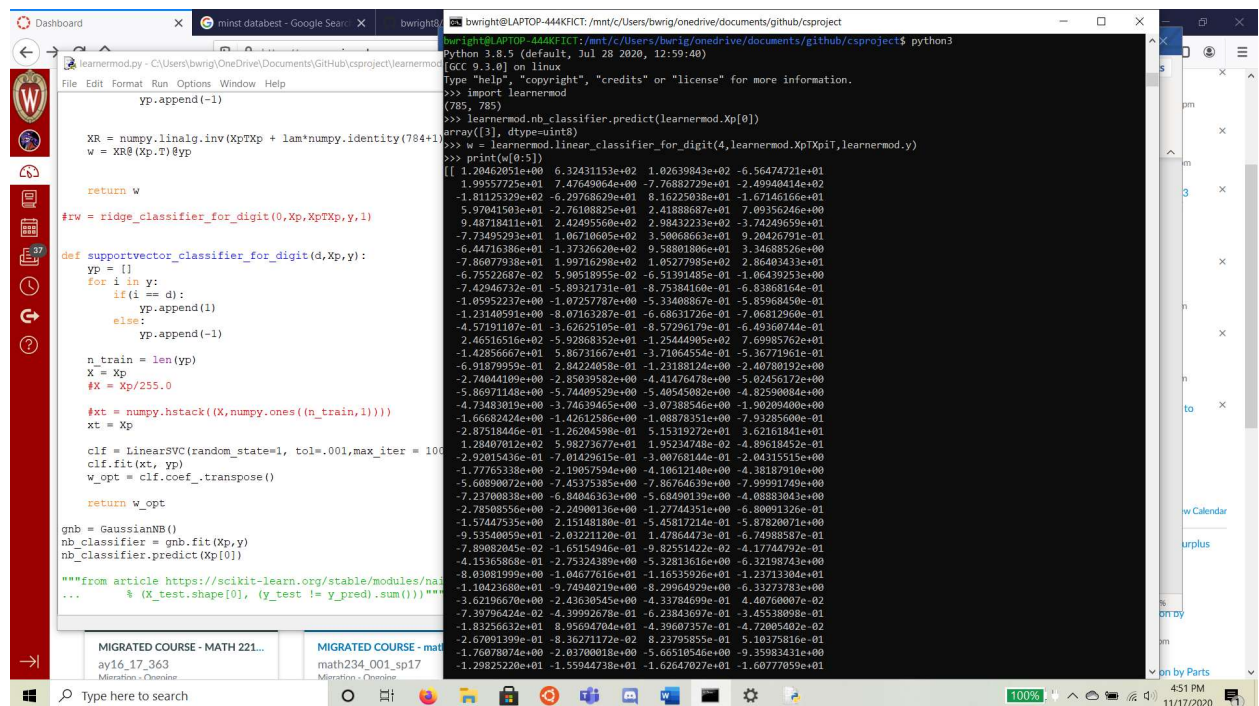


Benjamin Wright
November 17, 2020
Matrix Methods in Machine Learning

Since writing the project proposal, I am par for the course on updates to the project. I have loaded and processed the data from the MNIST dataset using a special loading program. I have implemented classifiers for linear and ridge regression, as well as implemented classifiers based on support vectors and naïve Bayes' (this puts me ahead of my goal in terms of implementing classifiers). A python module (currently called "learnermod" in the GitHub project) builds a classifier based on a given method and digit. Below is a screenshot of using the predictor built by the naïve Bayes' method and determining the linear classifier corresponding to the digit "4" with the module.

Yet to do is to begin analyzing the data and writing the final report and selecting a parameter for the ridge regression (probably by segmenting training data into cross-validation sets). That is, by December 1st, I should have implemented all classifier methods (only neural network left to do) and compiled all the data for analysis in the final report – specifically testing the classifiers on the testing set. I will have begun writing the final report by then.

The project is maintained at <https://github.com/bwright8/cspjct/tree/master>.



The screenshot shows a Windows desktop environment. On the left, a code editor window displays Python code for a classifier. The code includes imports for numpy and sklearn, and defines functions for linear and ridge regression classifiers. The main part of the code is a function `supportvector_classifier_for_digit` that takes a digit `d` and a set of points `yp` as input. It uses `LinearSVC` from sklearn to train a classifier on the data. The code also includes a comment about the source of the data: `from article https://scikit-learn.org/stable/modules/naive_bayes.html`. On the right, a terminal window shows the output of the code. The output includes the version of Python (3.8.5) and the version of the `learnermod` module (0.3.0). It also shows the output of the `supportvector_classifier_for_digit` function, which is a list of 1000 values representing the predicted class for each point in the dataset. The values are mostly 0, 1, 2, 3, 4, 5, 6, 7, 8, and 9, indicating that the classifier is predicting the digit for each point.