

CS 4361/5361 Machine Learning

Fall 2020

Practice Exam 1

1. Write the function *remove_constant_attributes(X)* that receives a 2D array *X* representing a dataset and returns the 2D array resulting from removing from *X* all attributes that have the same value for all examples (for example, the pixels in MNIST for which all images have a value of 0).
2. Write the function *most_common_error(y,p)* that receives 1D arrays *y* and *p* containing the actual and predicted target function values and returns two integers *actual* and *predicted* describing the most common error made by the classifier. For example, in MNIST, the most common error is to misclassify 4 as 9, so your function should return 4,9 in that case.
3. Write the function *most_important_feature(X,y)* that receives a 2D array *X* and a 1D array *y* and uses the *linear_regression* program to determine which is the most important feature in *X* in order to predict *y*. For example, with the particles data set, your function should return 4.
4. Write code to evaluate the performance of the Naive Bayes classifier on the MNIST data set using square and product features, as we did with linear regression. Does performance improve?