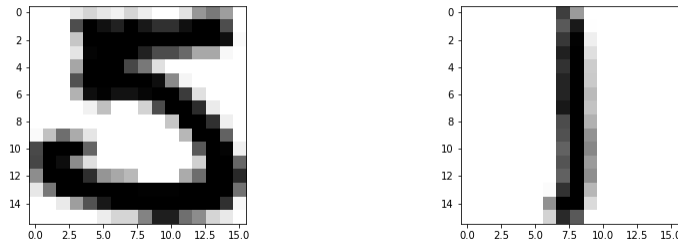


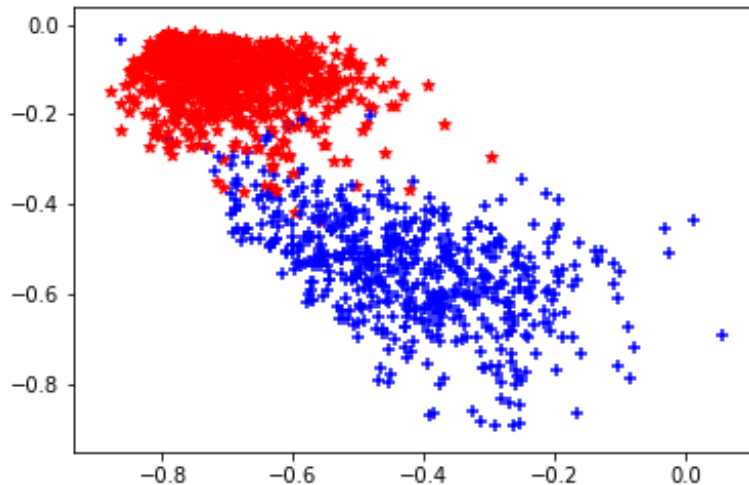
Ryan Torelli
11481335
January 24, 2018
CptS 437: Homework #1

2. Perceptron for Handwritten Digit Recognition

(a) The `show_images()` function shows each one of two 16 x 16 grayscale images. The function displays each image using a reverse grayscale colormap, 'gray_r', to show writing as dark.



(b) The `show_features()` function plots 1561 labeled data points. Each data point contains two features, symmetry and average intensity, such that labeled data is plotted on two axes. Data labeled as digit '1' is plotted in red with marker '*' and data labeled as digit '5' is plotted in blue with marker '+'. In the below scatter plot, data representing the digit '1' clusters in the upper left quadrant, whereas data representing the digit '5' clusters in the lower right quadrant, consistent with high symmetry and low intensity for '1' but low-medium symmetry and medium-high intensity for '5'.



(c) The `perceptron()` function learns weights for each feature and a bias. The function initializes weights to zero and in each iteration up to a maximum number of iterations, updates weights. The update randomly selects a misclassified data point and adds the product of the data point's features, correct label, and learning rate to the weights,

$$w(t+1) = w(t) + x*y*learning_rate$$

For ten pairs of maximum iterations and learning rate, training and testing accuracy are as follows:

Case	Max_iter	Learning_rate	Train acc	Test acc
1	10	0.1	0.936	0.908
2	30	0.1	0.979	0.965
3	50	0.1	0.976	0.958
4	100	0.1	0.867	0.849
5	200	0.1	0.980	0.958
6	200	0.1	0.982	0.962
7	200	0.2	0.980	0.955
8	200	0.3	0.981	0.953
9	200	0.4	0.962	0.936
10	200	0.5	0.924	0.896

The results of the trials show patterns with respect to arguments of maximum iterations and learning rate. For a constant learning rate, the number of iterations reaches a maximum accuracy, followed by perturbations of noisy data. For a constant number of iterations, the learning rate reaches a maximum accuracy and then degrades. In addition, there is a relationship between training and testing accuracy; that is, in each trial, training accuracy is always greater than testing accuracy.

(d) The `show_result()` function plots 424 labeled data points separated by a line. The line describes weights derived from training the `perceptron()` function on the training set of 1561 labeled data points for 10 iterations at a learning rate of 0.5. As before, data labeled as digit '1' is plotted in red with marker '*' and data labeled as digit '5' is plotted in blue with marker '+'. In the below scatter plot, it is evident that the test dataset is not perfectly separable. The weights output by the `perceptron()` function, therefore, cannot divide data with perfect accuracy.

