

Ryan Torelli
11481335
February 7, 2018
CptS 437: Homework #2

4. Logistic Regression for Handwritten Digit Recognition

(a) The `logistic_regression()` function learns a weight for each feature and returns the weights. The function initializes weights to zero and, for each iteration up to a maximum number of iterations, computes the gradient,

$$-\frac{1}{N} \sum_{n=1}^N \frac{y_n \mathbf{x}_n}{1 + e^{y_n \mathbf{w}^T(t) \mathbf{x}_n}}$$

and updates weights by adding the product of the negative gradient and learning rate,

$$\mathbf{w}(t+1) = \mathbf{w}(t) + \eta \mathbf{v}_t \text{ where } \mathbf{v}_t = -\mathbf{g}_t$$

(b) The `accuracy()` function classifies data according to a logistic regression model and returns the percentage of correctly-classified data. The function classifies data by passing the product of transposed weights and features,

$$\mathbf{w}^T \mathbf{x}_n$$

into the logistic function,

$$\frac{e^s}{1+e^s}$$

in order to compute a probability between zero and 1. If the probability is greater than 0.5, the data classifies as 1; otherwise, the data classifies as -1.

(c) The `thirdorder()` function applies a third-order transform to data and returns the transformed data. For data containing two features and a bias, the function generates ten features,

$$1, x_1, x_2, x_1^2, x_1 x_2, x_2^2, x_1^3, x_1^2 x_2, x_1 x_2^2, x_2^3$$

(d) The results of the accuracy trials show that the third-order polynomial model has equivalent or superior accuracy to the linear model for both training and test data. Therefore, the third-order polynomial model has better fit to the data and is recommended.

Table 1. The accuracy of the linear model for training and test data with varying parameters.

Case	Max_iter	Learning_rate	Train acc	Test acc
Max_iter test 0	100	0.2	0.835	0.828
Max_iter test 1	200	0.2	0.924	0.901
Max_iter test 2	500	0.2	0.966	0.941
Max_iter test 3	1000	0.2	0.974	0.950
Learning_rate test 0	1000	0.1	0.966	0.941
Learning_rate test 1	1000	0.2	0.974	0.950
Learning_rate test 2	1000	0.5	0.979	0.962

Table 2. The accuracy of the third-order model for training and test data with varying parameters.

Case	Max_iter	Learning_rate	Train acc	Test acc
Max_iter test 0	100	0.2	0.924	0.899
Max_iter test 1	200	0.2	0.958	0.941
Max_iter test 2	500	0.2	0.970	0.948
Max_iter test 3	1000	0.2	0.975	0.955
Learning_rate test 0	1000	0.1	0.971	0.948
Learning_rate test 1	1000	0.2	0.975	0.955
Learning_rate test 2	1000	0.5	0.978	0.965