Multi-class Classification and Neural Networks

LBYCP29 – Laboratory 7

Gervin Guevarra, Allen Koizumi, Nicholas Moreno, Charleston Uy

Department of Electronics and Communications Engineering

Gokongwei College of Engineering, De La Salle University

Manila, Philippines

*Abstract*—This laboratory report presents the implementation of one-vs-all logistic regression and neural networks to discriminate the hand-written digits. Similar concepts from the previous experiment were used, but as opposed to the previous laboratory experiment, which only concerns itself with 1s and 0s, this experiment deals with digits from 0 to 9.

Keywords—Logistic regression, one-vs-all, multi-class classification

# Introduction

The previous laboratory exercise used a huge data consisting of handwritten 1s and 0s. The task was to teach the machine to differentiate the 1s from the 0s, based on the training data. This was implemented as a binary classification as we are only concerned with two types or categories of data.

However, not all applications involve only two types of data. In this experiment, another huge data set containing 5000 samples of handwritten digits (which is a part of the larger MNIST dataset) is used. This experiment attempts to teach the computer to recognize the digits from one another by using a one-vs-all classification method, which takes one type of data and compares it against the rest and is repeated for other types. Moreover, the program must exhibit high prediction rate after it being introduced to the said sample data set.

# Objectives

The experiment aims to achieve the following objectives

* To obtain a high prediction accuracy, which is about 94.9%
* To implement a one-vs-all classifier and observe its behavior.

# Data and Results

After completing the required assignments in the code, we were able to achieve a 95.1% accuracy which is greater than our primary objective which is 94.9%. Additionally, we were able to obtain <> to <> that shows the cost function curve for each digit, and we can infer that they all share similar characteristics like how they decrease over time, but we can also notice some slight differences that can tell that <> is the hardest to learn while <> is the easiest, due to the nature of their curves.

# Analysis and Conclusion

After performing the experiment, the group concluded that multi-class classification is simply performing logarithmic regression on a one is to all basis meaning we will first classify a single class of data compared to the rest of the data, and this is done to each element of the data set until there is a classifier for every element of the data set. After which, our hypothesis has been properly trained to handle this data set, but this hypothesis cannot stand on its own and must use a neural network to be able to properly classify the data as a 0, 1, 2, etc. Now, there several ways to train a neural network depending on the application, and we decided that using the max function is ideal for this application since we want to use the label that the hypotheses deems to be of best fit for the given image. After training the neural network, we are then finished with devising a classifier for multiple digits.

# Bibliography

|  |  |
| --- | --- |
| [1] | A. Ng, "Spervised Learning," n.d. [Online].  Available: http://cs229.stanford.edu/notes/cs229-notes1.pdf [Accessed 22 September 2015]. |
| [2] | A. Ng, "Machine Learning,"  n.d. [Online]. Available: https://class.coursera.org/ml-003/lecture.  [Accessed 22 September 2015]. |
| [3] | Institute for Digital Research and Edcation, "Stata Data Analysis and  Exampes", 2015 [Online]. Available: http://www.ats.ucla.edu/stat/stata/dae/logit.htm.  [Accessed 22 September 2015]. |

# Appendix