Perceptron and Logistic Regression

Thomas Duong	**Maggie Ha**	**Dr. Katarina Jegdic**	**Dr. Parker**
Undergraduate, Jr.	Undergraduate, Sr.	Professor	**PI Project Director, MSEIP**
Statistics Major	Mathematics Major	Department of Mathematics and Statistics	Exec. Director
Department of Mathematics and Statistics	Department of Mathematics and Statistics	College of Sciences and Technology	Scholars Academy
College of Sciences and Technology	College of Sciences and Technology	University of Houston-Downtown	College of Sciences and Technology
University of Houston-Downtown	University of Houston-Downtown	Houston, TX	University of Houston-Downtown
Houston, TX	Houston, TX	jegdick@uhd.edu	Houston, TX
duongt13@gator.uhd.edu	hat10@gator.uhd.edu		parkerm@uhd.edu

Abstract

The main goal of this research is to analyze and compare several machine learning algorithms. These algorithms have important applications in our daily life including banking and financial systems, businesses and economics, cybersecurity, medicine and healthcare, and many other fields.

We considered Perceptron, Logistic Regression, and Linear Regression. They are supervised machine learning algorithms with Perceptron and Logistic Regression being used for classification problems and Linear Regression being used for regression problems. By understanding these algorithms and how they are used on big data sets, we can develop more practical and efficient approaches for data analysis and predictions.

The Perceptron algorithm is a single neuron model with sign function as the activation function to make prediction for a binary outcome. The predicted output in perceptron is a hard label with discrete values -1 and +1.

Logistic Regression is a single neuron model with sigmoid function as the activation function to make predictions in the range from 0 to 1. The predicted value is considered the probability of a data instance belonging to the positive class. Compared to Perceptron, Logistic Regression uses a continuous function as an activation function and it gives a probability value as the predicted output.

For Linear Regression, we use gradient descent method to find the minimum value of the loss function which is the Root Mean Square Error. We studied both batch gradient descent method and stochastic gradient descent with scaling. The batch gradient descent gives a better prediction; however, it takes long time for machine to run, and it also depends on the machine capacity in term of memory. Although stochastic gradient descent with scaling is not as accurate as batch gradient descent, it is faster and more practical.

STEM category: Mathematics and Statistics

Keywords: machine learning algorithms, data science

Acknowledgements: Minority Science and Engineering Improvement Program (MSEIP)
Grant number: **P120A220015**, "Boosting STEM Student Success", Dr. Bernadette Hence at U.S. Department of Education, and University of Houston-Downtown.