BITS F464 : Machine Learning
Assignment - 1
Linear Perceptron Algorithm

Omkar Pitale
2019A7PS0083H

Aditya Chopra
2019A7PS0178H

Anuradha Pandey
2019A7PS0265H

March, 2021

1 Introduction

Logistic Regression is a method when the variable to be determined(the dependent variable) is categorical. It is used to assign observations to a discrete set of classes. Logistic regression transforms its output using the logistic sigmoid function to return a probability value which can then be mapped to two or more discrete classes.

2 Model Description

2.1 Model Design

The Model is defined as a class, with two instance variables, w, the weights and b, the bias of the model. The parameter w can be initialized with a Gaussian distribution, zeros or ones. b is initialized to 0. The model has two Gradient Descent Algorithms: Stochastic Gradient Descent (SGD) and Batch Gradient Descent (BGD). The optimizers, takes in a feature set and optimizes the model over it. Loops have been avoided as much as possible in favour of vectorization of mathematical operations. The only loop is over the number of epochs that the training runs over.

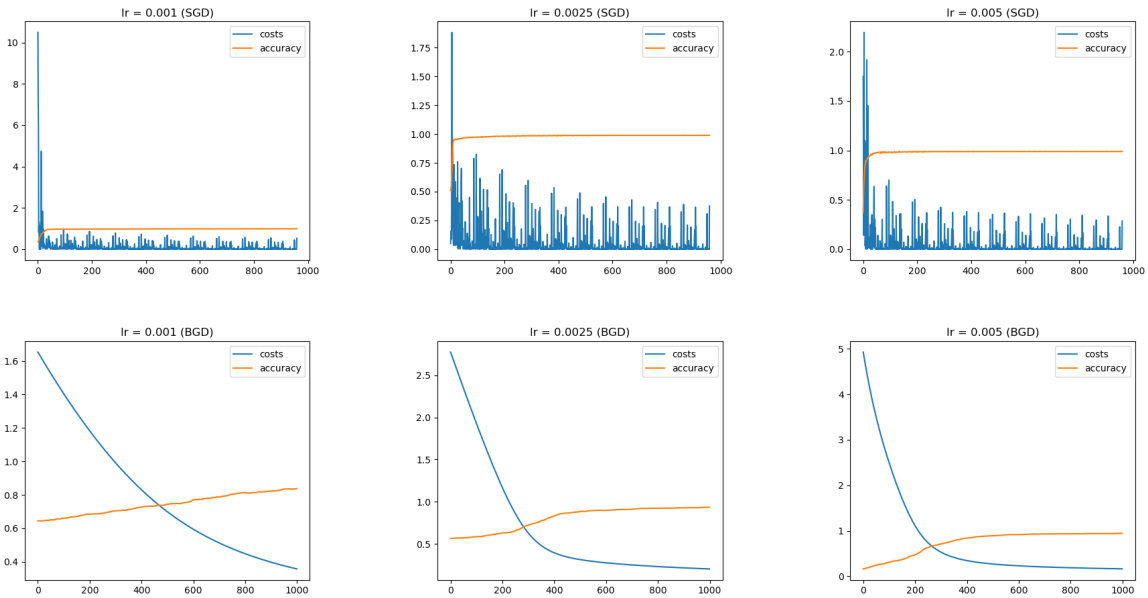

Figure 1: Accuracy vs Iterations Plot for Stochastic and Batch Gradient Descents for various Learning Rates, $\eta \in \{0.001, 0.0025, 0.005\}$

3 Results

1. Four Measures of Performance: Accuracy, Precision, Recall and F1-Score were considered
2. The goal of logistic regression is to find a D dimensional hyper-plane that separates the two classes. We use the D dimensional weight vector to obtain the feature importance.
3. Feature 1, i.e. attr1 has the highest negative importance.

SN	Loss	Accuracy	Precision	Recall	F1-Score
1	0.419	0.9879	0.9829	0.9885	0.9857
2	0.503	0.9854	0.984	0.984	0.984
3	0.335	0.9903	0.989	0.989	0.989
4	0.419	0.9879	0.9767	0.9941	0.9853
5	0.251	0.9927	0.9944	0.989	0.9917
6	0.419	0.9879	0.9891	0.9838	0.9864
7	0.252	0.9927	0.9826	1	0.9912
8	0.335	0.9903	0.989	0.989	0.989
9	0.587	0.983	0.97	0.9949	0.9823
10	0.251	0.9927	0.9943	0.9886	0.9914
Mean	0.377	0.9891	0.9852	0.9901	0.9876

Table 1: Performance with Batch Gradient Descent ($\eta = 0.1$, epochs=1000) over 10 random samples

SN	Loss	Accuracy	Precision	Recall	F1-Score
1	0.671	0.9806	0.9771	0.9771	0.9771
2	0.503	0.9854	0.9802	0.99	0.9851
3	0.335	0.9903	0.9832	0.9944	0.9888
4	0.335	0.9903	0.9777	1	0.9887
5	0.335	0.9903	0.9941	0.9826	0.9883
6	0.503	0.9854	0.9716	0.9942	0.9828
7	0.419	0.9879	0.9835	0.989	0.9862
8	0.419	0.9879	0.9818	0.9878	0.9848
9	0.084	0.9976	0.9943	1	0.9972
10	0.168	0.9951	0.9897	1	0.9948
Mean	0.377	0.9891	0.9833	0.9915	0.9874

Table 2: Performance with Stochastic Gradient Descent ($\eta = 0.01$, epochs=50) over 10 random samples

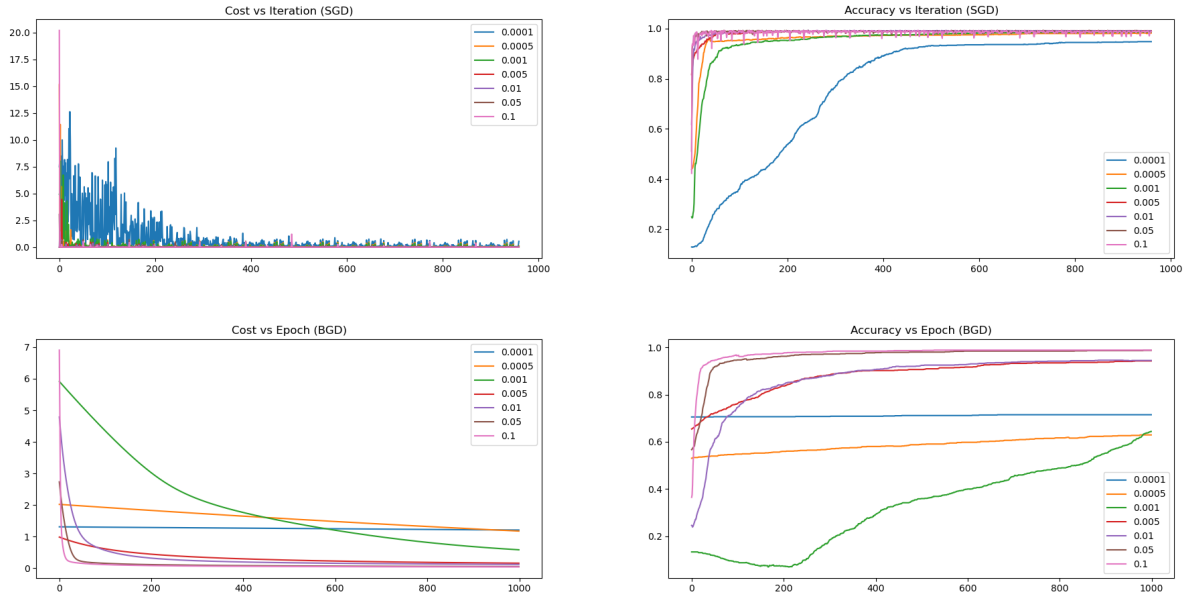

Figure 2: Performance Comparison of SGD and BGD