PREDICTING THE CHANCE OF GRADUATE ADMISSION USING MULTIPLE LINEAR REGRESSION

Mahidi Hasan Mithun(1930432)

September 2023

Abstract

We wanted to find out if we could predict the chance of getting into graduate school using factors like GRE scores, undergraduate GPA, research experience, and other features. Using Multiple Linear Regression to do this. Given the data set, five hundred people applied for graduate programs. After running our analysis, we found that our model could somewhat accurately predict if someone would get admitted or not. While the model can give a good idea about the chances of getting in.

1 Introduction

This project's goal is to build a predictive model using multiple linear regression. Probability of getting a chance to be admitted based on the given feature. The data set contains 500 data points, the last column representing the chance of admission.

2 Methodology

The first data load in Google Colaboratory and remove the first column serial number of the data points. The given dataset was shuffled and split into 80% for training and 20% for testing datasets. The feature was normalized using MinMaxScalar range between 0 and 1. Three variant gradient descent algorithms were used. Batch gradient descent, mini-batch gradient descent batch size 10, and stochastic gradient descent batch size 1. Hyperparameter setting the next. The learning rate or alpha is set at 0.01 and the number of epochs 100,500,1000 and 10000. The iterative process was run for different numbers of epochs and observed which gradient gives a good minimum for MSE. Build the loss-vs-iterations graph for each of the models.

3 Performance of the regression model

The effectiveness of the regression model is quantified by using the Mean Squared Error (MSE) metric.

Epochs	Batch GD MSE	Mini-batch GD MSE	Stochastic GD MSE
100	0.005666838605815978	0.002157955252968609	0.0019161636906942308
500	0.0031766557208861503	0.0019196930941349903	0.0018989194073230718
1000	0.002528748135479099	0.0018788487606554367	0.0018989188181898166

Table 1: MSE for Different Gradient Descent Algorithms and Epochs

For 100000 epochs it takes too much time to run. Among 100,500 and 1000 epochs, Mini-batch Gradient Descent performed the best with the lowest MSE of 0.0018989194073230718.

4 Plot and Result

Generally, stochastic gradient descent makes updates more frequent, and it can converge faster. Its updates are noisy, and it may not settle to the best minimum but sawing around it. For 100 and 500

it is to find the good minimum because it updates frequently. It is beneficial in scenarios where you have fewer epochs. We see in the plot for 100 epochs MSE 0.0057 for batch gradient descent, 0.0021 for mini-batch gradient descent, and 0.0019 for stochastic gradient descent. We clearly see the best model with MSE is stochastic gradient descent. The plot for 500 epochs is the best model with MSE stochastic gradient descent analysis from Graph.

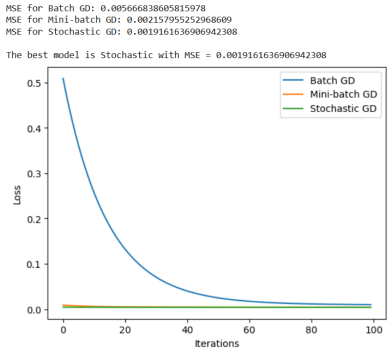

Figure 1: 100 epochs

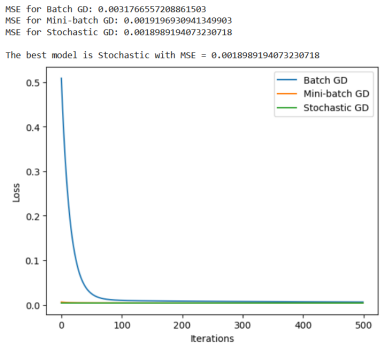

Figure 2: 500 epchos

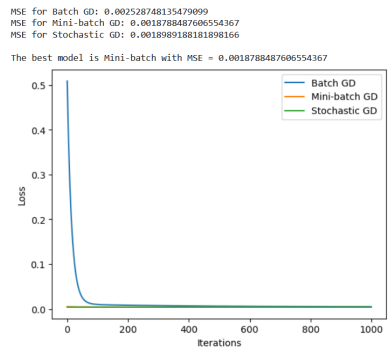

Figure 3: 1000 epchos

We see from graph, that increasing iterations loss will decrease. Some change observation here, for 1000 epochs we see the best model mini-batch gradient descent. Mini-batch gradient descent ends up being faster to converge. Mini-batch gradient descent is less noisy than stochastic gradient descent. 1000 epochs it had enough updates to find the good minimum. Batch gradient descent is the most stable but slowest among the three gradient descents. For 10000 epochs I observed some change here. I see here the good minimum is batch gradient descent. A large number of epochs and, large number of data sets batch gradient can require significant computational power.

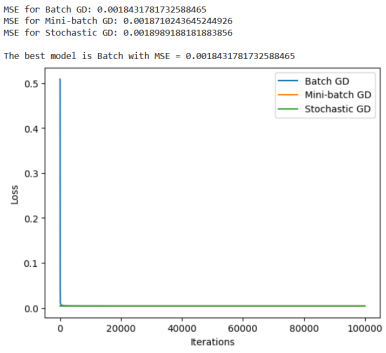

Figure 4: 100000 epchos

5 Conclusion

The objective to predict the "Chance of Admit" for graduate admissions was successfully achieved using regression models. Among three gradient descents, the mini-batch gradient provides the best result with the lowest MSE for 1000 epochs. Batch gradient descent provides the best result with the lowest MSE for 100,000 epochs among these three gradient descents. For 100,000 epochs, the MSE difference between mini-batch gradient descent and batch gradient descent is very small.

References

https://www.investopedia.com/terms/m/mlr.asp
http://www.stat.yale.edu/Courses/1997-98/101/linmult.htm
https://ieeexplore.ieee.org/document/9249747