University of Tripoli Faculty of Engineering

EEE 569: Introduction to deep learning

Assignment 1 Part (A)

Reem Ben Guma

Introduction:

This part of the assignment focuses on enhancing the functionality and performance of a logistic regression model by implementing foundational computation nodes, introducing batching, and analyzing batch size effects.

Data:

Data set contains two linearly separable classes.

- Class 1: $\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mu = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$
- Class 2: $\sum = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\mu = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$

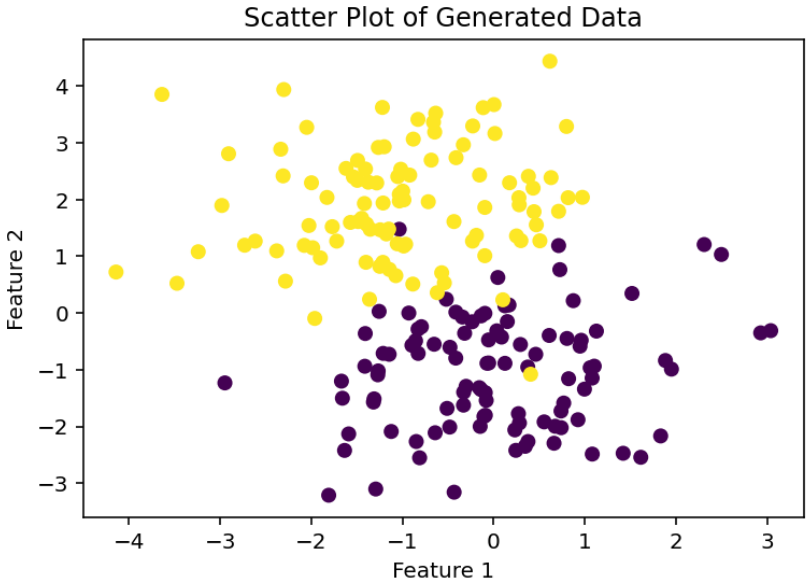

Figure 1: scatter plot of data set.

Logistic Regression with Single Sample Training:

Using single sample training (Stochastic Gradient Descent), one-hundred epoch, and 0.01 learning rate, testing the model on 0.25% of data set unseen samples gives accuracy of 96.0%, where the last epoch has loss of 0.0215.

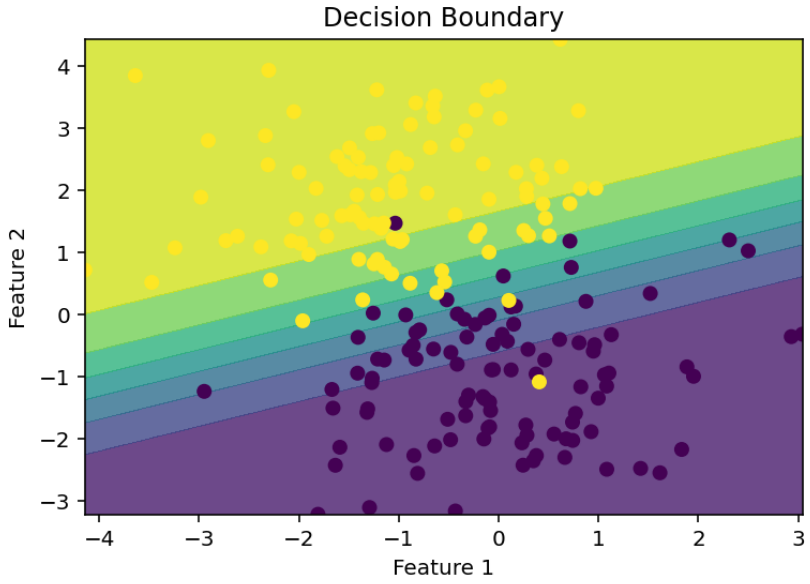

Figure 2: Decision boundary.

investigating batch size impacts:

Figures below show the impact of increasing batch size on models' performance:

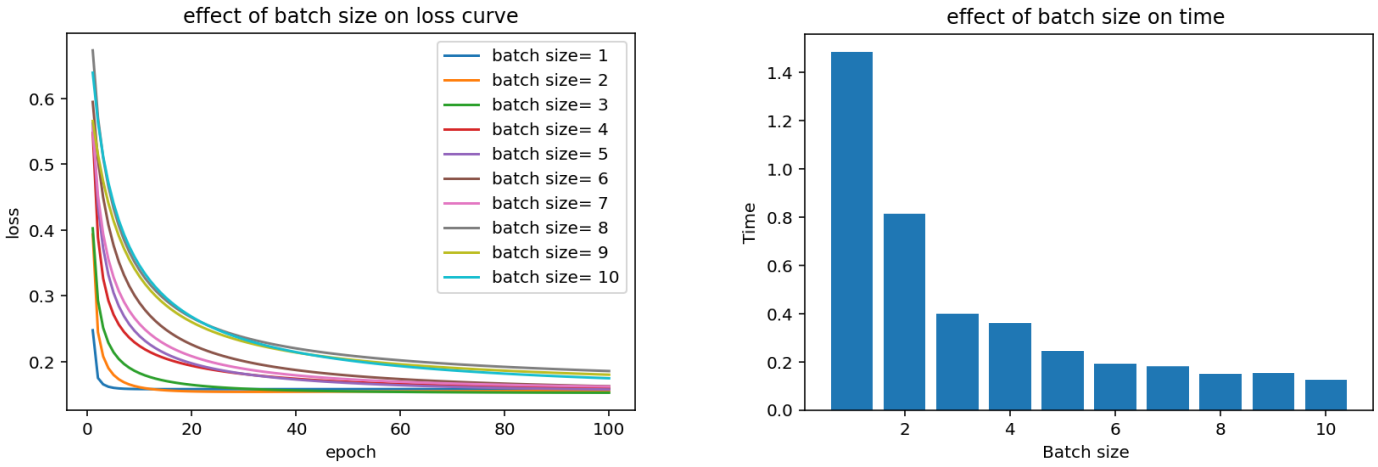

Figure 3: Batch Size vs. Training Loss for Batch Sizes 1, 2, ..., 10.

The effect of batch size on the training process varies depending on the problem. As shown in Figure 3, increasing the batch size causes the gradients to converge towards a local minimum. Additionally, as observed in the time bars, the significant improvement in training time compared with the minor difference in loss values, making larger batch sizes preferable.

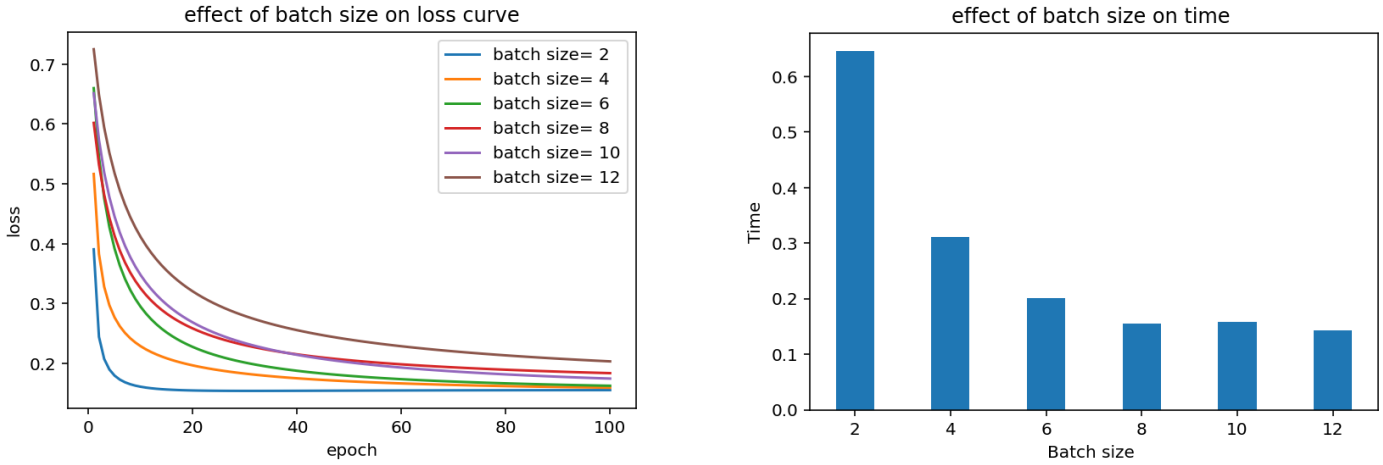

Figure 4: Batch Size vs. Training Loss for Batch Sizes 2, 4, ..., 12.

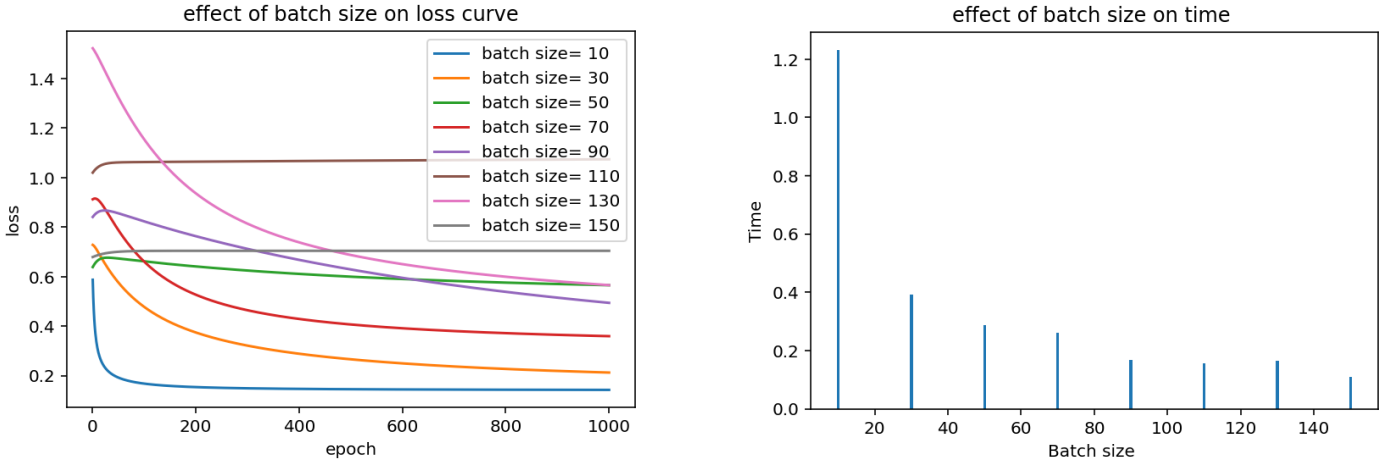

Figure 5: Batch Size vs. Training Loss for Batch Sizes 10, 30, …, 150

Even though the training process involves 1000 epochs, it doesn't take much time due to the use of larger batch sizes. As noted above, the improvement in running time between batch sizes 10 and 30, with only a small difference in loss, makes a batch size of 30 more preferable.