University of Tripoli
Faculty of Engineering
Department of Electrical and Electronic Engineering

Fall-2024
EE569 - Deep Learning

Assignment #1

الاسم: محمد محمد محمد النعاس
رقم القيد:2200209040

Part 1 Introducing the Linear Node:

A new “Linear” Node was introduced to the EDF file, this new Node combined the functionality of the Multiply Node and the Addition Node. allowing to reduce the number of nodes present in the 03_logistic_Regression_with_linear.py file. Adding the Linear Node made no effects on the accuracy and loss of the Model.

Part 2 Introducing Batches:

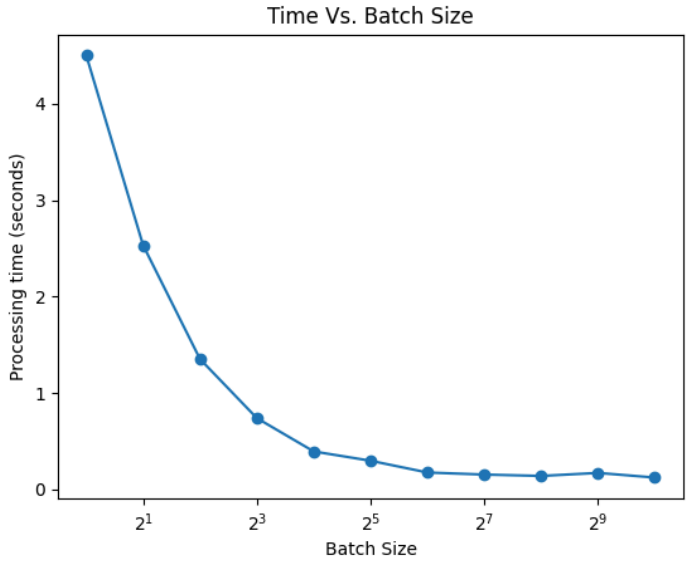
The EDF.py file was copied into EDF_batch.py to modify the function to accept matrices as input data. The Nodes that were changed to facilitate the use of Batching were the Linear Node and the BCE Node in both their forward and backward passes, note that the Multiply Node and Addition Node were not modified as they were made obsolete by the Linear Node. The modified 03_logistic_regression_with_linear_plus_batch.py file contains additional code used to modify the inputs in a way they can be run through the nodes as patches. As well as functionality to run multiple model trainings on the same dataset using a multitude of batch sizes (1,2,4,8,16… and so on). Finally it also records model properties such as final loss, accuracy (which uses modified code to properly function) and training time for each batch size. And uses the data recorded to make comparative graphs that show the effects of batch size on each property.

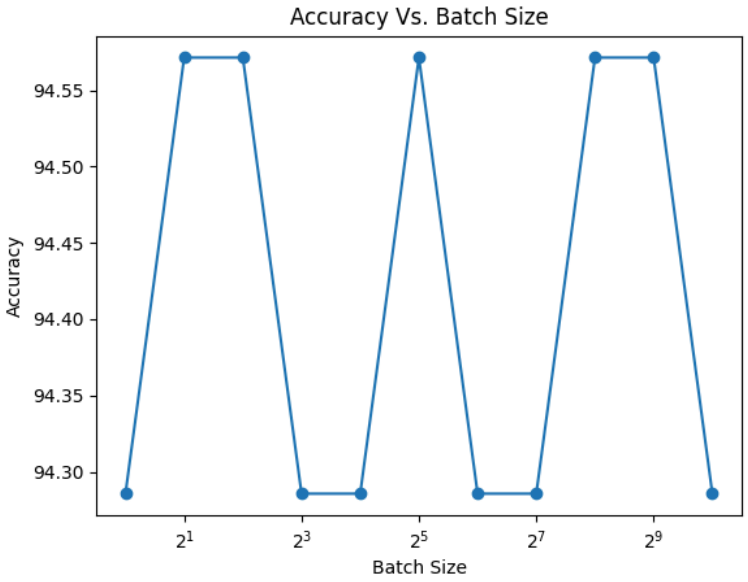
Part 3 The Effects of Batch Size:

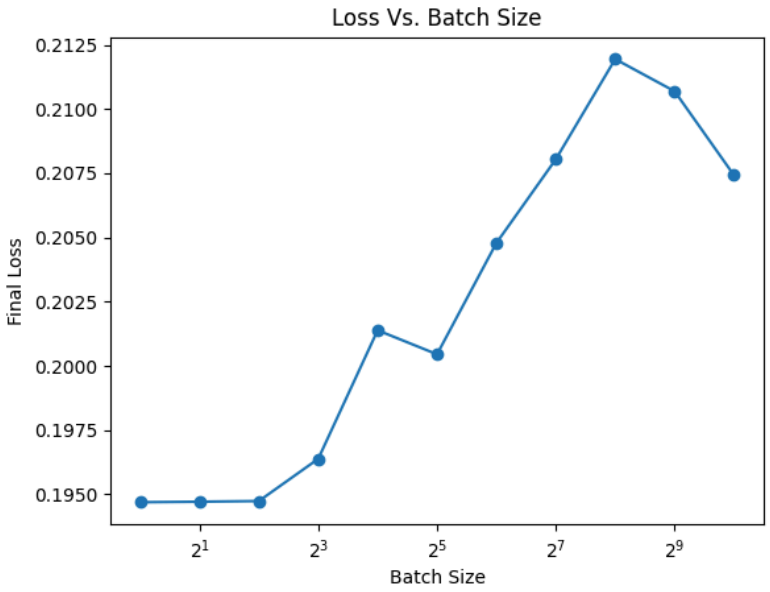
Different Batch sizes have varying effects on the different properties of the model. And certain properties depend both on the properties of the dataset and batch size.
Some of the most prominent and consistent effects of increasing the batch size, is the inversely-proportional relationship between it and model training time. As increasing the batch size reduces the training time by up to half for double the batch size.
Additionally, increasing the batch size increases the memory used. As the bigger batch size means a bigger matrix that needs to be loaded into memory.
Training Loss as well as model accuracy seem to be negatively affected by Batch size, with test accuracy decreasing as batch size increases.

The following figures show the model properties vs. batch for a dataset with the following properties:

```
CLASS1_SIZE = 700
CLASS2_SIZE = 700
MEAN1 = np.array([3, -2])
COV1 = np.array([[6, 0], [0, 3]])
MEAN2 = np.array([-1, 3])
COV2 = np.array([[2, 0], [0, 8]])
```






Loss Vs. Batch Size
Final Loss
0.2125
0.2100
0.2075
0.2050
0.2025
0.2000
0.1975
0.1950
2^1
2^3
2^5
2^7
2^9
Batch Size