

# Assignment Report

Syed Alle Mustafa

October 5, 2022

## 1. TASK 1

### (a) Problem Statement:

- i. Implement three layer neural network
- ii. Use categorical cross entropy for loss
- iii. Iterate over batches
- iv. Perform forward pass to compute unit outputs.
- v. Compute needed gradients
- vi. Apply Computed gradients to update network weights.
- vii. Compute training and validation loss and accuracy of each cycle.

### (b) Proposed Solution:

Balance weights dataset is loaded from UCI datasets since it is a multiclass dataset containing classes Balance, Left and Right. So After cleansing, normalizing and loading and partitioning dataset, a sequential model was used with two sigmoid layers and one softmax layer, were used to achieve the best performance form the given data set. More details are discussed in the implementation part.

### (c) Implementation Details:

- i. Loaded Balance Scale Dataset.
- ii. Normalizing train and test datasets.
- iii. Splitting train dataset into 2 parts, one for training and one for validation.
- iv. initializing a sequential model for training.
- v. initializing training and validation accuracy matrices.
- vi. Initializing epochs to 80;
- vii. Enumerating datasets batches.
- viii. training a batch and finding loss value and saving them on gradient tape.
- ix. Finding out gradients using loss functions.
- x. optimizing the weights using gradients.
- xi. Checking accuracy and validation and repeating this steps till the epoch times.

### (d) Results: The result of training and validation is as follows in the graph, the result shows that the accuracy and loss of both training and validation set is good.

## 2. TASK 2

(a) Problem Statement:

- i. Perform the solution of above problem with using computational graphs.

(b) Proposed Solution:

A class for computation graph is defined with functions Sigmoid, Softmax, Sigmoid Derivative, Softmax Derivative, Categorical cross entropy and Derivative of CCE. All these functions perform the Activations, loss calculation and derivative of the operation they perform. Then there are 2 main functions, forward pass and backward pass. forward pass takes the data  $x$  and find its dot product with weight1 and then perform bias addition and then find the sigmoid of the result, this step is repeated but with out of sigmoid1 and weight2. After this a softmax function is perform and then the loss values are calculated. In the backward pass the derivatives are found out in reverse order of forward pass and in the end we subtract the derivatives of loss with respect to weight times the learning rate from the actual weight and repeat this process several times.

(c) Result:

result of both the validation set during the training and testing set of the final model are quite impressive.