



Assignment#2

11.03.2020

Ahmed Nasser abdelkareem (9)

Problem Statement

In this assignment we will be addressing the task of multi nominal classification of handwritten digits from the famous MNIST dataset.

The MNIST dataset consists of 60,000 training images and 10,000 test images. Our classes are the digits 0-9.

You are required to build 2 models to solve this problem:

- A simple using fully connected layers.
- A model using Convolutional Neural Network (CNN) before applying the simple model.

you should observe the difference in accuracy between the 2 models.

Fully connected NN

- 1) Compile your model, you'll want to experiment with both the choice of optimizer and learning rate and evaluate how these affect the accuracy of the trained model?

- I tune three parameters (optimizer, loss, activation function)

```
optimizers = ['SGD', 'RMSprop', 'Adagrad', 'Adam', 'Adamax']
```

```
activation = ['sigmoid', 'tanh', 'relu']
```

```
losses =
```

```
['categorical_crossentropy', 'mean_squared_error', 'mean_squared_logarithmic_error', 'binary_crossentropy', 'sparse_categorical_crossentropy']
```

- 2) Evaluate accuracy on the test dataset, what is your observation on the accuracy on the test dataset and the accuracy on the training dataset?

- I found the most appropriate parameters:

```
- Model ( Adam relu sparse_categorical_crossentropy ) with  
Test accuracy: 0.9756
```

```
- Model ( Adam tanh sparse_categorical_crossentropy ) with  
Test accuracy: 0.9739
```

```
- Model ( RMSprop relu sparse_categorical_crossentropy )  
with Test accuracy: 0.9739
```

```
- Model ( RMSprop tanh sparse_categorical_crossentropy )  
with Test accuracy: 0.9727
```

CNN

- 1) Compile the CNN model with an optimizer, learning rate and regularizer of choice.

- I tune three parameters (optimizer, loss, activation function)

```
optimizers = ['RMSprop', 'Adam']
```

```
activation = ['tanh', 'relu']
```

```
losses = ['sparse_categorical_crossentropy']
```

- 2) You are required to show the running time, number of parameters and number of multiplication of your model in estimation and in training over the same number of epochs.

Number of multiplication in CNN:

- + (26*26*3*3*24)
- + (24*3*3*36*11*11)
- + (748*128)
- + (128*10)

- 3) What is the highest accuracy you're able to achieve using the CNN model, and how does the accuracy of the CNN model compare to the accuracy of the simple fully connected network?

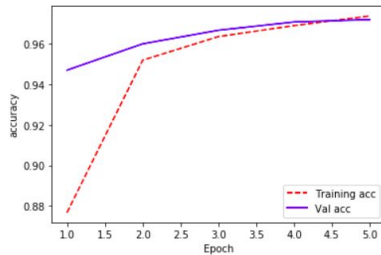
- Model (Adam relu sparse_categorical_crossentropy) with Test accuracy: 0.9904 and Test loss: 0.030511482527873886 In time: 214.91840028762817 the highest accuracy using CNN.
- Model (Adam relu sparse_categorical_crossentropy) with Test accuracy: 0.9756 the highest accuracy using FC.

- 4) What is the result of using Relu and tanh activations, which of them produced a better result?

- Relu has better results than tanh

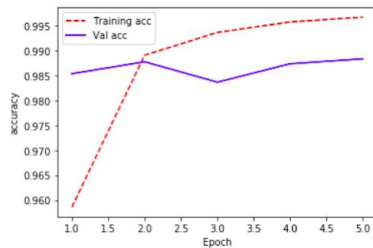
5) Try different convolution filter sizes, stride lengths, and pool layers. Give plots for the effect of those changes.

- Try to change stride length



Model (filter_size:24,36 ,strides: 2 ,pooling_size: 2) with Test accuracy: 0.9721 and Test loss: 0.091050905758515 In time: 48.679410457611084

- Try to change pooling size



Model (filter_size:24,36 ,strides: 1 ,pooling_size: 1) with Test accuracy: 0.9884 and Test loss: 0.04034636159711227 In time: 1432.1209542751312

[[stride length compared to convolution filter size](#)]

Note

- We should tune learning rate after using Adam and RMSprop optimizer