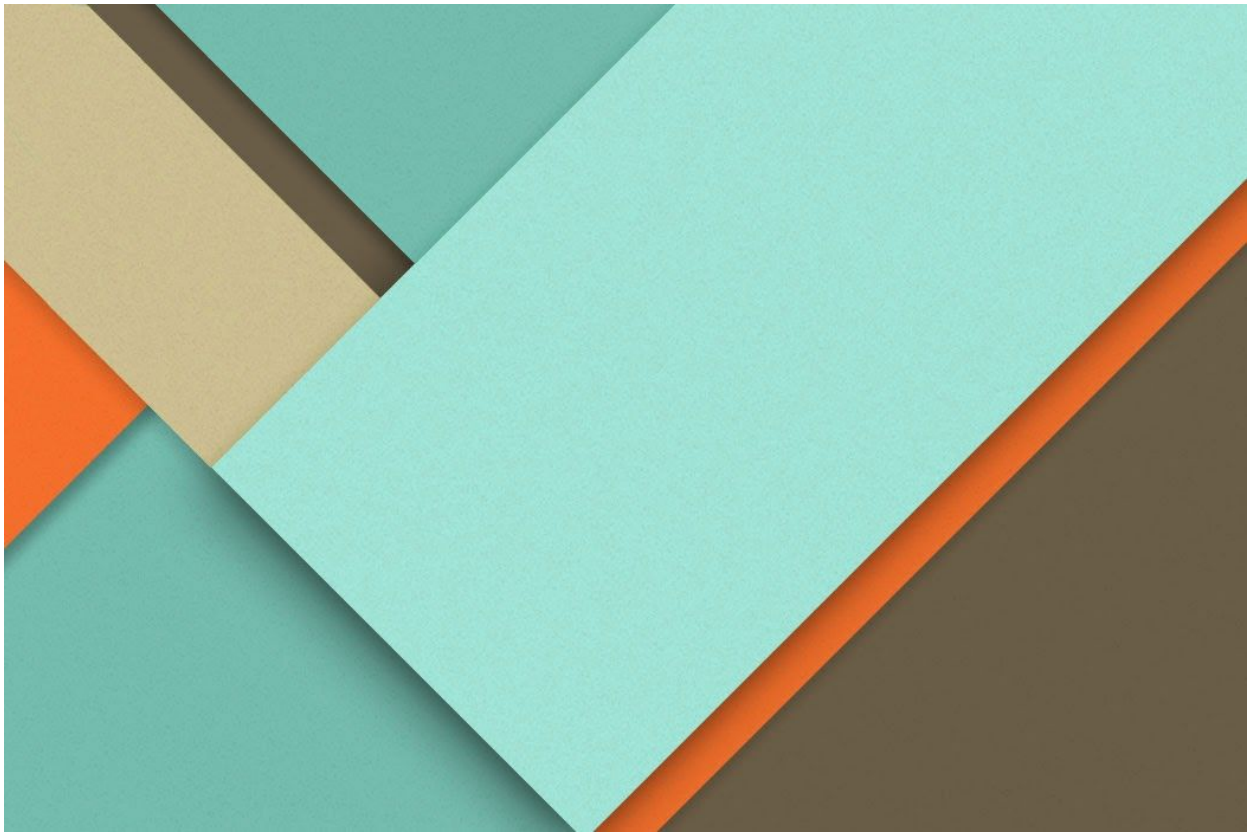


# Deep Learning Multinomial Classification

11.03.2020

Aya Ashraf Saber Mohamed

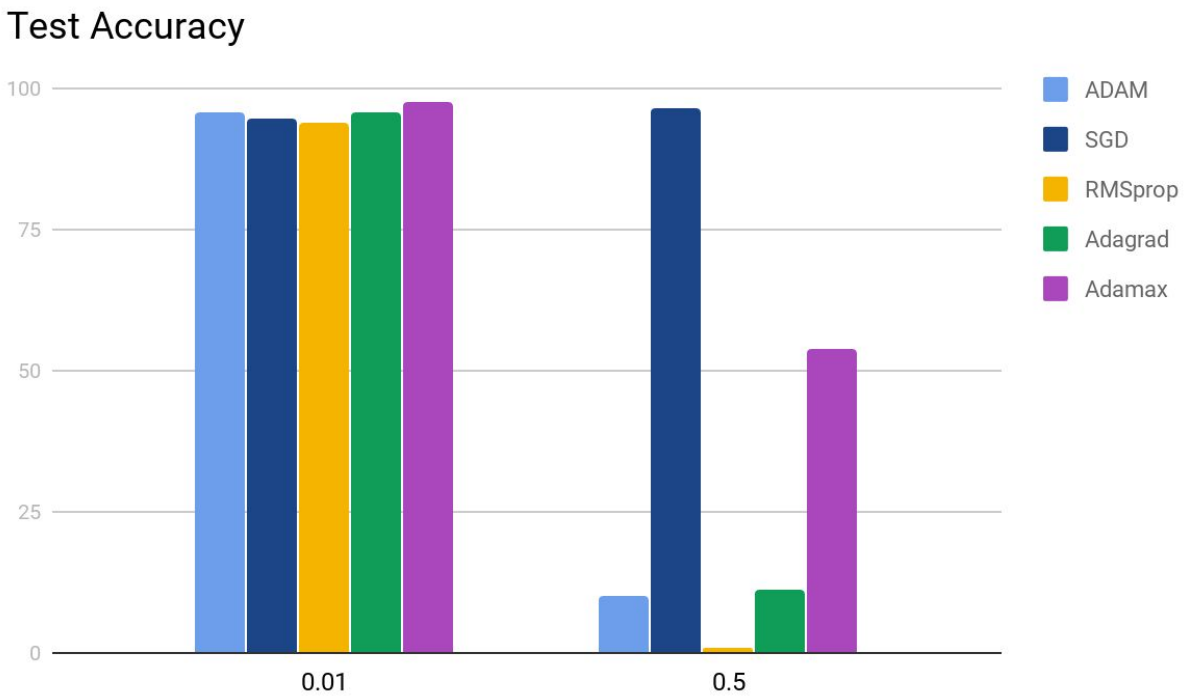
# Fully Connected Layer

## Comparison

Observation the test dataset accuracy and train dataset accuracy are almost the same

| # | Optimizer | Learning Rate | Test Accuracy |
|---|---|---|---|
| 1 | ADAM | 0.01 | &95.91 |
| 2 | | 0.5 | %10.09 |
| 3 | SGD | .01 | %94.62 |
| 4 | | 0.5 | %96.58 |
| 5 | RMSprop | .01 | %93.76 |
| 6 | | .5 | %0.975 |
| 7 | Adamax | .01 | %97.72 |
| 8 | | .5 | %53.92 |
| 9 | Adagrad | .01 | %95.69 |
| 10 | | .5 | %11.35 |

## Evaluation




# CNN

## Running Time , # Multiplication

Architecture 1

```
Model: "sequential_4"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_6 (Conv2D)            (None, 26, 26, 24)        240
_________________________________________________________________
max_pooling2d_5 (MaxPooling2 (None, 13, 13, 24)        0
_________________________________________________________________
conv2d_7 (Conv2D)            (None, 11, 11, 36)        7812
_________________________________________________________________
max_pooling2d_6 (MaxPooling2 (None, 5, 5, 36)          0
_________________________________________________________________
flatten_4 (Flatten)          (None, 900)               0
_________________________________________________________________
dense_9 (Dense)              (None, 128)               115328
_________________________________________________________________
dense_10 (Dense)             (None, 10)                1290
=================================================================
Total params: 124,670
Trainable params: 124,670
Non-trainable params: 0
_________________________________________________________________
```

## Epochs = 10

```
_________________________________________________________________
None
Train on 60000 samples, validate on 10000 samples
Epoch 1/10
60000/60000 [==============================] - 37s 609us/sample - loss: 2.2513 - accuracy: 0.1868 - val_loss: 2.2085 - val_accuracy: 0.2560
Epoch 2/10
60000/60000 [==============================] - 36s 596us/sample - loss: 2.1596 - accuracy: 0.3474 - val_loss: 2.1116 - val_accuracy: 0.4236
Epoch 3/10
60000/60000 [==============================] - 36s 594us/sample - loss: 2.0591 - accuracy: 0.4960 - val_loss: 2.0038 - val_accuracy: 0.5613
Epoch 4/10
60000/60000 [==============================] - 35s 591us/sample - loss: 1.9459 - accuracy: 0.6073 - val_loss: 1.8816 - val_accuracy: 0.6511
Epoch 5/10
60000/60000 [==============================] - 36s 602us/sample - loss: 1.8183 - accuracy: 0.6769 - val_loss: 1.7453 - val_accuracy: 0.7018
Epoch 6/10
60000/60000 [==============================] - 36s 605us/sample - loss: 1.6785 - accuracy: 0.7167 - val_loss: 1.5990 - val_accuracy: 0.7366
Epoch 7/10
60000/60000 [==============================] - 36s 602us/sample - loss: 1.5325 - accuracy: 0.7407 - val_loss: 1.4503 - val_accuracy: 0.7585
Epoch 8/10
60000/60000 [==============================] - 37s 610us/sample - loss: 1.3884 - accuracy: 0.7581 - val_loss: 1.3075 - val_accuracy: 0.7753
Epoch 9/10
60000/60000 [==============================] - 37s 623us/sample - loss: 1.2537 - accuracy: 0.7717 - val_loss: 1.1776 - val_accuracy: 0.7875
Epoch 10/10
60000/60000 [==============================] - 37s 624us/sample - loss: 1.1337 - accuracy: 0.7832 - val_loss: 1.0639 - val_accuracy: 0.8004
Test accuracy: 0.8004
```

Architecture 2

```
Model: "sequential_5"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_8 (Conv2D)            (None, 26, 26, 32)        320
_________________________________________________________________
conv2d_9 (Conv2D)            (None, 24, 24, 64)        18496
_________________________________________________________________
max_pooling2d_7 (MaxPooling2 (None, 12, 12, 64)        0
_________________________________________________________________
dropout_2 (Dropout)          (None, 12, 12, 64)        0
_________________________________________________________________
flatten_5 (Flatten)          (None, 9216)              0
_________________________________________________________________
dense_11 (Dense)             (None, 128)               1179776
_________________________________________________________________
dropout_3 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_12 (Dense)             (None, 10)                1290
=================================================================
Total params: 1,199,882
Trainable params: 1,199,882
Non-trainable params: 0
_________________________________________________________________
```

## Epochs = 10

```
Train on 60000 samples, validate on 10000 samples
Epoch 1/10
60000/60000 [==============================] - 148s 2ms/sample - loss: 2.2661 - accuracy: 0.1851 - val_loss: 2.2160 - val_accuracy: 0.4008
Epoch 2/10
60000/60000 [==============================] - 148s 2ms/sample - loss: 2.1752 - accuracy: 0.3410 - val_loss: 2.0913 - val_accuracy: 0.5646
Epoch 3/10
60000/60000 [==============================] - 147s 2ms/sample - loss: 2.0326 - accuracy: 0.4647 - val_loss: 1.8940 - val_accuracy: 0.6792
Epoch 4/10
60000/60000 [==============================] - 148s 2ms/sample - loss: 1.8242 - accuracy: 0.5470 - val_loss: 1.6204 - val_accuracy: 0.7507
Epoch 5/10
60000/60000 [==============================] - 147s 2ms/sample - loss: 1.5680 - accuracy: 0.6087 - val_loss: 1.3128 - val_accuracy: 0.7947
Epoch 6/10
60000/60000 [==============================] - 147s 2ms/sample - loss: 1.3280 - accuracy: 0.6496 - val_loss: 1.0522 - val_accuracy: 0.8178
Epoch 7/10
60000/60000 [==============================] - 147s 2ms/sample - loss: 1.1386 - accuracy: 0.6856 - val_loss: 0.8649 - val_accuracy: 0.8378
Epoch 8/10
60000/60000 [==============================] - 149s 2ms/sample - loss: 1.0077 - accuracy: 0.7088 - val_loss: 0.7395 - val_accuracy: 0.8503
Epoch 9/10
60000/60000 [==============================] - 150s 2ms/sample - loss: 0.9089 - accuracy: 0.7313 - val_loss: 0.6524 - val_accuracy: 0.8599
Epoch 10/10
60000/60000 [==============================] - 148s 2ms/sample - loss: 0.8424 - accuracy: 0.7470 - val_loss: 0.5910 - val_accuracy: 0.8652
Test accuracy: 0.8652
```

# Comparison between CNN and Fully Connected Layer:

## Given number of Epochs = 5

| Comparison | CNN | Fully Connected Layer |
|---|---|---|
| Training Accuracy | %62.33 | %97.16 |
| Test Accuracy | %77.76 | %97.16 |

## Relu and Tanh activation Function: Epoch = 5

| Comparison | CNN | Fully Connected Layer |
|---|---|---|
| Training Accuracy (Relu ) | %62.33 | %97.16 |
| Test Accuracy (Relu) | %77.76 | %97.16 |
| Training Accuracy (Tanh) | %64.09 | %97.42 |
| Test Accuracy (Tanh) | %66.52 | %97.42 |

## Hyper Parameter

| Filter Size | Test Accuracy |
|---|---|
| 32 | %77.76 |
| 16 | %68.01 |
| 8 | %53.05 |

| Stride Length | Test Accuracy |
|---|---|
| 1 | %77.76 |
| 2 | %70.56 |
| 3 | %65.50 |

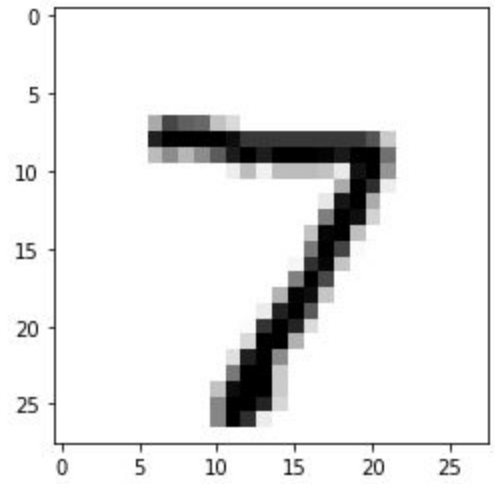
| Pool layers | Test Accuracy |
|---|---|

| | |
|---|---|
| 1 | %80.04 |
| 2 | %86.52 |

## CNN Model Prediction:

### 1- Model structure in the Assignment

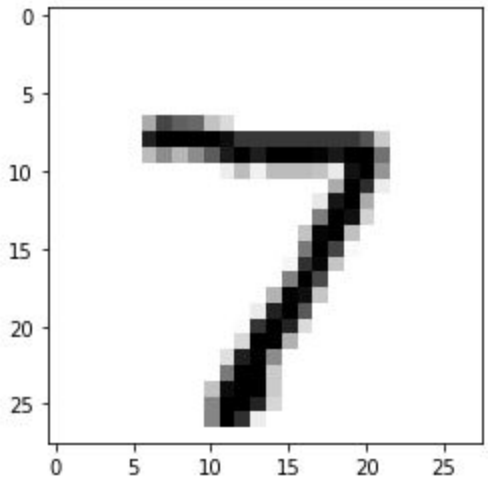
```
[0.08887121 0.04857256 0.08950614 0.06493146 0.11228155 0.0801947
 0.06087552 0.23661964 0.09799125 0.12015595]
arg max 7
7
```




### 2- Update  CNN Architecture with # Epoch = 5

## Architecture:

```
[0.04406603 0.01642291 0.02297809 0.03457483 0.06798876 0.0291586
 0.02277647 0.6455393  0.03967604 0.07681904]
arg max 7
7
```




## Resources:-

[Google Colab code]