CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING
SECTION 02

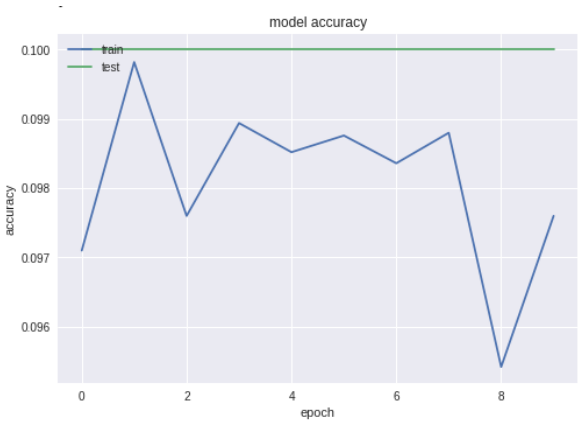
ASSIGNMENT 1

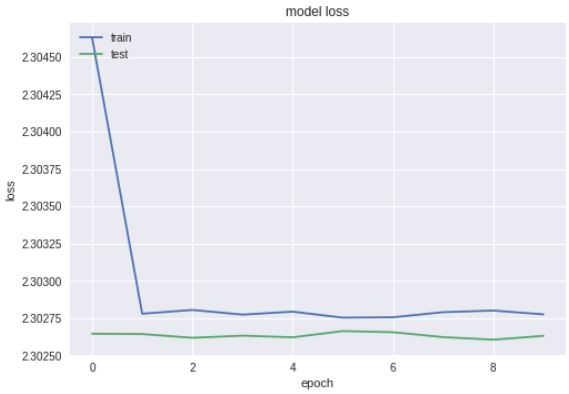
PROBLEM 1 : Implementing CIFAR – 10 using MLP classification

Architecture 1 :

- Input Layer nodes : 784
- Number of hidden layers : 4 (512, 256, 64, 16)
- Output layer nodes : 10
- Activation : Relu, Relu, Relu, Relu, Relu and Softmax
- Dropout layers : 2 (0.2)
- Optimizer : Adam
- Loss : Categorical cross entropy
- Metrics : Accuracy
- Batch size = 32
- Epochs = 10
- Output of this model :
  Training Loss : 2.30 Training accuracy : 0.09
  Testing Loss : 2.30 Testing accuracy : 0.1

This model is not recommended even though the model doesn't overfit because, the accuracy is very low and loss is very high. Thus this architecture will result in a poor model.





Architecture 2 :

- Input Layer nodes : 784
- Number of hidden layers : 2 (256, 64)

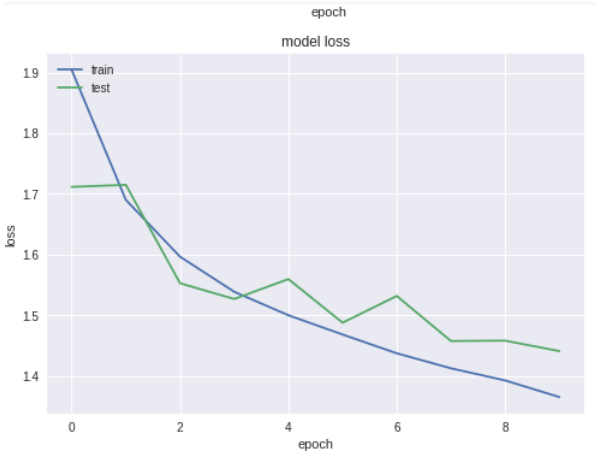
CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING

SECTION 02

ASSIGNMENT 1

- Output layer nodes : 10
- Activation : Relu, Relu, Relu and Softmax
- Dropout layers : 0
- Optimizer : Adam
- Loss : Categorical cross entropy
- Metrics : Accuracy
- Batch size = 32
- Epochs = 10
- Output of this model :
  Training Loss : 1.37 Training accuracy : 0.51
  Testing Loss : 1.44 Testing accuracy : 0.49

This model is not recommended because this one overfits , accuracy is low and loss is very high. Thus this architecture will result in a poor model.



Architecture 3 :

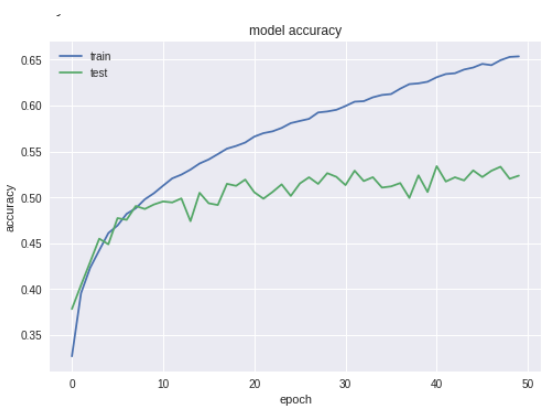
- Input Layer nodes : 256
- Number of hidden layers : 1 (64)
- Output layer nodes : 10
- Activation : Relu, Tanh and Softmax
- Dropout layers : 0
- Optimizer : Rmsprop(learning rate = 0.0001)
- Loss : Categorical cross entropy
- Metrics : Accuracy
- Batch size = 50
- Epochs = 50

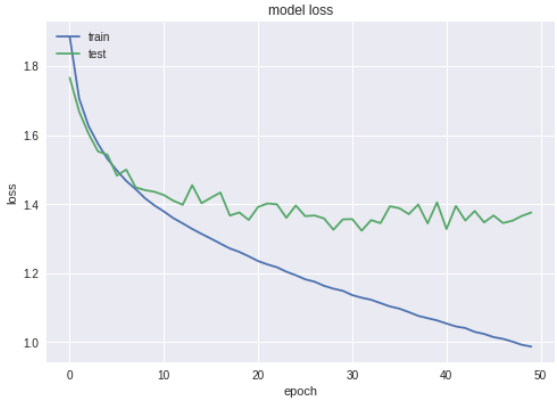
CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING
SECTION 02

ASSIGNMENT 1

- Output of this model :
  Training Loss : 0.95 Training accuracy : 0.66
  Testing Loss : 1.34 Testing accuracy : 0.537

Till now, out of all the architecture this one has better accuracy and loss score.





Architecture 4 :

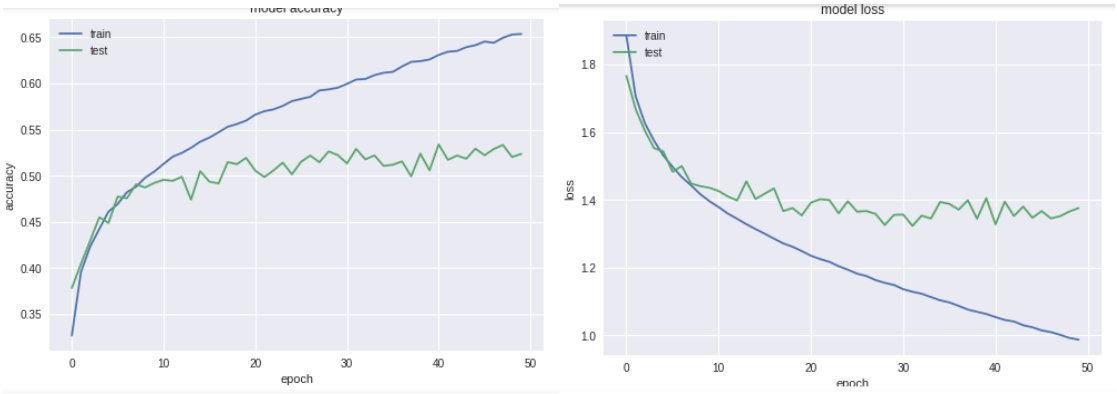
- Input Layer nodes : 256
- Number of hidden layers : 1 (64)
- Output layer nodes : 10
- Activation : Relu, Tanh and Softmax
- Dropout layers : 0
- Optimizer : Rmsprop(learning rate = 0.01)
- Loss : Categorical cross entropy
- Metrics : Accuracy
- Batch size = 50
- Epochs = 50
- Output of this model :
  Training Loss : 2.39 Training accuracy : 0.1008
  Testing Loss : 2.35 Testing accuracy : 0.1

This model is same as the previous model but with a slightly higher learning rate. Even though this model has a higher learning rate, the accuracy is very low and loss score is very high.

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING

SECTION 02

ASSIGNMENT 1



Architecture 5 :

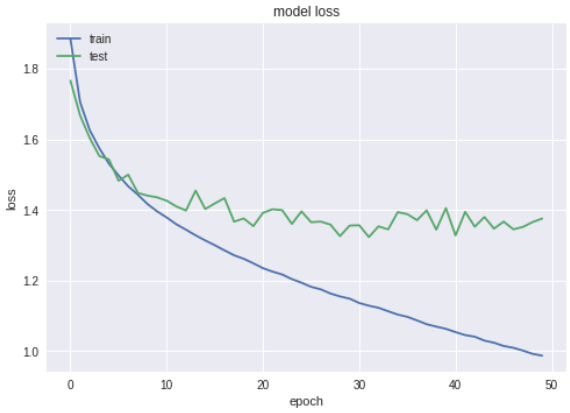
- Input Layer nodes : 784
- Number of hidden layers : 2 (256,64)
- Output layer nodes : 10
- Activation : Tanh, Relu, Tanh and Softmax
- Dropout layers : 0
- Optimizer : Rmsprop(learning rate = 0.01)
- Loss : Categorical cross entropy
- Metrics : Accuracy
- Batch size = 50
- Epochs = 50
- Output of this model :
  Training Loss : 2.39 Training accuracy : 0.09
  Testing Loss : 2.33 Testing accuracy : 0.1

In this model, the number of nodes In the input layer is changed and also the activation. Still, loss is high and accuracy is very low.

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING

SECTION 02

ASSIGNMENT 1



Architecture 6 :

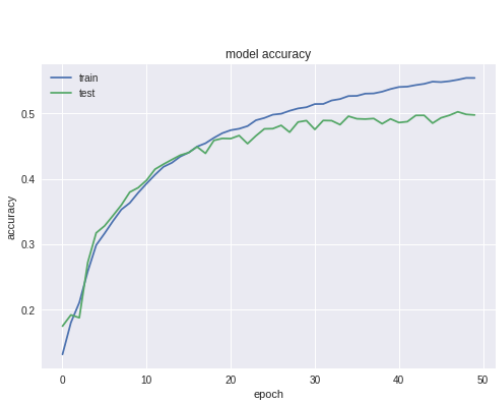
- Input Layer nodes : 784
- Number of hidden layers : 2 (256,64)
- Output layer nodes : 10
- Activation : Relu, Relu, Tanh and Softmax
- Dropout layers : 1(0.2)
- Optimizer : Adam
- Loss : Categorical cross entropy
- Metrics : Accuracy
- Batch size = 50
- Epochs = 50
- Output of this model :
  Training Loss : 1.16 Training accuracy : 0.580
  Testing Loss : 1.41 Testing accuracy : 0.51

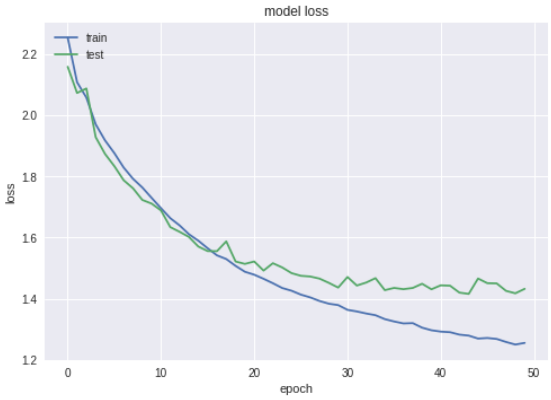
In this model, the activation in the input layer is changed, dropout layer is added and the optimizer is changed. There is an improvement in the accuracy but the loss is still high. A good model compare to other models.

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING

SECTION 02

ASSIGNMENT 1





Architecture 7 :

- Input Layer nodes : 784
- Number of hidden layers : 2 (256,64)
- Output layer nodes : 10
- Activation : Relu, Relu, Relu and Softmax
- Dropout layers : 1(0.5)
- Optimizer : Adam
- Loss : Categorical cross entropy
- Metrics : Accuracy
- Batch size = 32
- Epochs = 50
- Output of this model :
  Training Loss : 1.45 Training accuracy : 0.47
  Testing Loss : 1.51 Testing accuracy : 0.46

In this model, the all the activation is changed to Relu and dropout has been increased. The batch size is reduced to 32. The loss is high but the accuracy is not bad for a mlp classifier

---

Architecture 8 :

- Input Layer nodes : 784
- Number of hidden layers : 2 (256,64)
- Output layer nodes : 10
- Activation : Relu, Relu, Relu and Softmax
- Dropout layers : 0
- Optimizer : Adam
- Loss : Categorical cross entropy

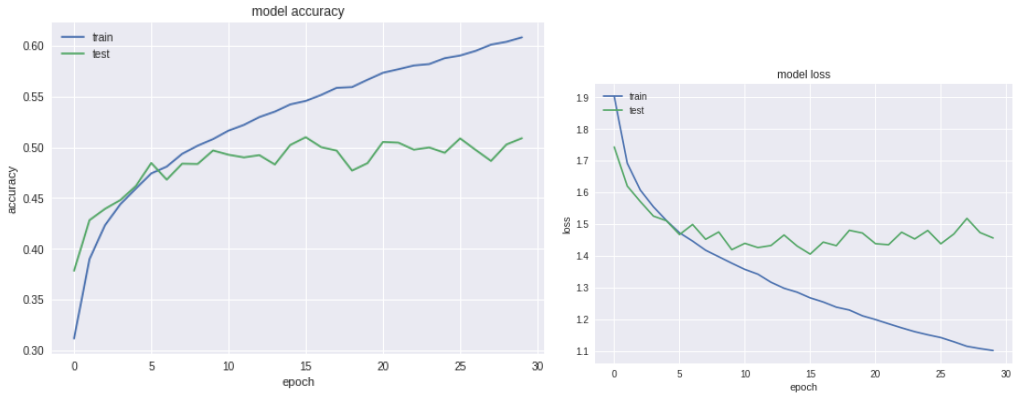
CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING

SECTION 02

ASSIGNMENT 1

- Metrics : Accuracy
- Batch size = 32
- Epochs = 30
- Output of this model :
  Training Loss : 1.10 Training accuracy : 0.60
  Testing Loss : 1.47 Testing accuracy : 0.50

Dropout layer is not added and the number of epochs is reduced in this model. Better accuracy but poor loss score.



Architecture 9 :

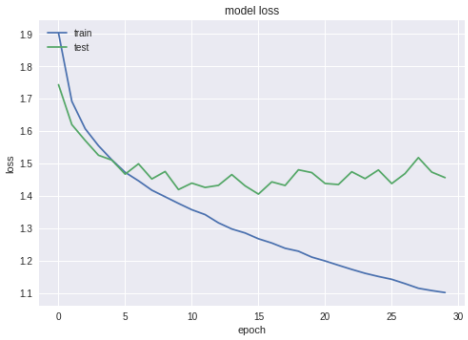
- Input Layer nodes : 256
- Number of hidden layers : 1 (64)
- Output layer nodes : 10
- Activation : Relu, Relu, and Softmax
- Dropout layers : 0
- Optimizer : Adam
- Loss : Categorical cross entropy
- Metrics : Accuracy
- Batch size = 24
- Epochs = 30
- Output of this model :
  Training Loss : 1.28 Training accuracy : 0.53
  Testing Loss : 1.45 Testing accuracy : 0.49

In this model, the number of nodes in the input layer is decreased to 256, hidden layer is decreased to 1 and batch size is decreased to 24. This model is overfitting.

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING

SECTION 02

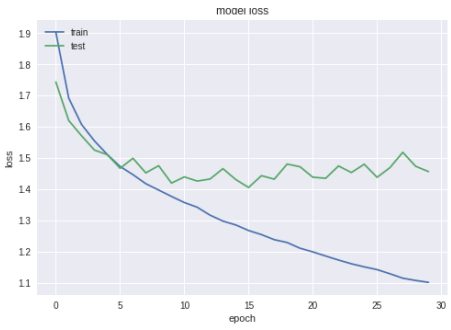
ASSIGNMENT 1



Architecture 10 :

- Input Layer nodes : 512
- Number of hidden layers : 2 (256,64)
- Output layer nodes : 10
- Activation : Relu, Relu, Relu and Softmax
- Dropout layers : 0
- Optimizer : Adam
- Loss : Categorical cross entropy
- Metrics : Accuracy
- Batch size = 32
- Epochs = 30
- Output of this model :
  Training Loss : 1.12 Training accuracy : 0.59
  Testing Loss : 1.48 Testing accuracy : 0.49

In this model, the number of input nodes is increased to 512, batch size increased to 32. The model still overfits and the testing loss is very high.



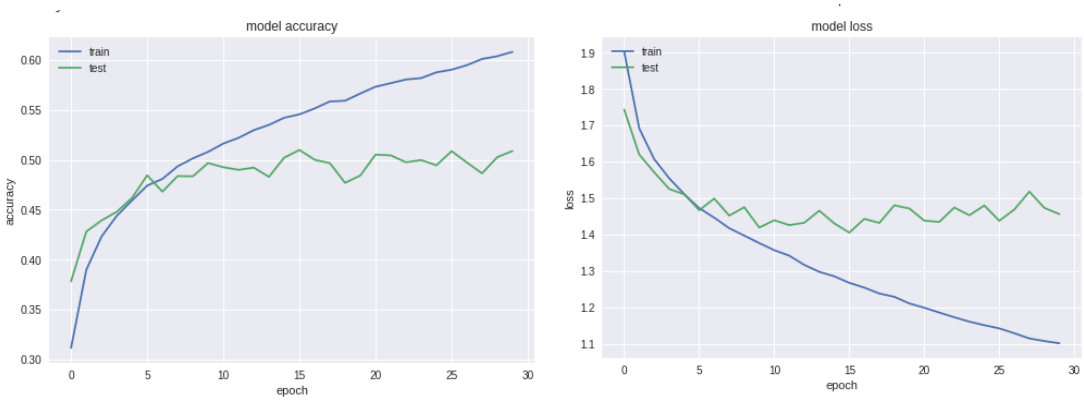
CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING
SECTION 02

ASSIGNMENT 1

Architecture 11 :

- Input Layer nodes : 2048
- Number of hidden layers : 2 (1024,512)
- Output layer nodes : 10
- Activation : Relu, Relu, Relu and Softmax
- Dropout layers : 3(0.2)
- Optimizer : AdaDelta
- Loss : Categorical cross entropy
- Metrics : Accuracy
- Batch size = 32
- Epochs = 10
- Output of this model :
  Training Loss : 1.63 Training accuracy : 0.42
  Testing Loss : 1.62 Testing accuracy : 0.42

In this model, the base architecture is changed. The input layer has 2048 nodes and has 2 hidden layers with 1024 and 512 nodes with relu activation. After each layer a dropout of 20% is added and the optimizer used is Adadelta. Number of epochs is reduced to 10. The model has performed well without overfitting but with a high loss rate and low accuracy.



Which model and why? Does it overfit?

Architectures 2,3,6,7,8,9,10,11 has performed better compared to other models. Out of these models though architecture 3 has a good testing accuracy of 0.53 it overfits. So, it is not a good model. Architectures 10 and 6 are eliminated because they overfit. Architecture 2 performed well but has a high

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING
SECTION 02

ASSIGNMENT 1

test and train loss rate compared to other models. This leaves us with architectures7,8,9 and 11. Out of these 4 models 7 and 11 have performed well without overfitting but their loss rates are very high compared to other two models that is under consideration. Architectures 8 ad 9 slightly overfit, but comparing their loss rates, **I recommend Architecture 8.**

To improve the model :

Usually, number of epochs, deepening of the neural network will improve the accuracy and decrease the moss but in this case after trying both it gave unsatisfactory results.

If possible regularization can be added because that helps in solving overfitting problems. CNN architecture with regularization should be designed to improve the accuracy of the problem set.

Another way could be to use data augmentation which will result in more data on which the model will be trained.

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING
SECTION 02

ASSIGNMENT 1

Problem 2 :

Part 1 : CNN using Keras :

No : 11 :- Functional API :

As the name suggest, Functional API is not a hyperparameter, it is simply another method to build a model. Usually we use Sequential model, here in Functional API we simply give the input dimensions as input tensor which will be passed through different layers like the normal model and produce an output tensor. These tensors can be called as layers too.

Design rationale :

For this part of the assignment, I have designed a CNN architecture for the CIFAR-10 dataset using Keras Functional API.

The input is defined as a input tensor with input parameters 32,32,3 where 32,32 indicates the 32x32 dimension of the image and 3 indicates that it is a colored image(RGB).

This input tensor is then multiplied(dot product) with the first convolution layer and the output tensor is stored in conv1 as tensor. This tensor is then multiplied with the next layer and so on.

Finally, the model is defined with the input tensor and the output tensor(after flattening and dense layer). The test and the train data is then fitted to this model after the compile method with Adam as the optimizer, Categorical cross-entropy as the loss and accuracy as metrics. Batch size is 32 and the number of epochs is 10.

The output is :

Training loss : 0.712
Training accuracy : 0.750
Test loss : 0.71
Testing accuracy : 0.759

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING
SECTION 02

ASSIGNMENT 1

As you can see, this model has performed better than the one that was designed for problem 1 where we used MLP to classify the CIFAR-10 dataset. Moreover, this model doesn't overfit too.

Though this model doesn't use the sequential model from the Keras models, it has performed well.

Part B : Configuration file :

1st Configuration :

1. batch_size = 32
2. num_classes = 10
3. epochs = 10
4. nodes_1 = 32
5. nodes_2 = 64
6. filters = 3
7. dropout_1 = 0.25
8. dropout_2 = 0.5
9. dropout_3 = 0
10. data_augmentation = 'True'
11. max_pooling = 2
12. activation_1 = 'relu'
13. activation_2 = 'relu'
14. activation_final_layer = 'softmax'
15. dense = 512
16. optimizer = 'adam'
17. num_predictions = 20

Output :

Training loss : 0.85 Training accuracy : 0.70

Testing loss : 0.76 Testing accuracy : 0.73

This model has a good accuracy with high loss.

2nd configuration :

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING

SECTION 02

ASSIGNMENT 1

1. batch_size = 32
2. num_classes = 10
3. epochs = 20
4. num_predictions = 20
5. nodes_1 = 32
6. nodes_2 = 64
7. filters = 3
8. dropout_1 = 0.2
9. dropout_2 = 0.5
10. dropout_3 = 0
11. data_augmentation = 'True'
12. max_pooling = 2
13. activation_1 = 'relu'
14. activation_2 = 'relu'
15. activation_final_layer = 'softmax'
16. dense = 512
17. optimizer = 'adam'

Output :

Training loss : 0.76 Training accuracy : 0.73

Testing loss : 0.69 Testing accuracy : 0.76

This model has a very good accuracy and moderate loss value.

3rd Configuration :

1. batch_size = 24
2. num_classes = 10
3. epochs = 20
4. num_predictions = 20
5. nodes_1 = 32
6. nodes_2 = 128
7. filters = 3
8. dropout_1 = 0.2
9. dropout_2 = 0
10. dropout_3 = 0
11. data_augmentation = 'True'
12. max_pooling = 2
13. activation_1 = 'relu'

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING

SECTION 02

ASSIGNMENT 1

14. activation_2 = 'relu'
15. activation_final_layer = 'softmax'
16. dense = 128
17. optimizer = 'sgd'

Output :

Training loss : 0.78 Training accuracy : 0.72
Testing loss : 0.76 Testing accuracy : 0.74

A good model with high loss score.

4th Configuration :

1. batch_size = 32
2. num_classes = 10
3. epochs = 20
4. num_predictions = 20
5. nodes_1 = 64
6. nodes_2 = 128
7. filters = 3
8. dropout_1 = 0.2
9. dropout_2 = 0
10. dropout_3 = 0
11. data_augmentation = 'True'
12. max_pooling = 2
13. activation_1 = 'relu'
14. activation_2 = 'tanh'
15. activation_final_layer = 'softmax'
16. dense = 1024
17. optimizer = 'adam'

Output :

Training loss : 0.79 Training accuracy : 0.74
Testing loss : 0.78 Testing accuracy : 0.73

Though the model has high loss score, the accuracy is good and it doesn't overfit too.

5th Configuration :

1. batch_size = 32

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING

SECTION 02

ASSIGNMENT 1

2. num_classes = 10
3. epochs = 20
4. num_predictions = 20
5. nodes_1 = 32
6. nodes_2 = 64
7. filters = 3
8. dropout_1 = 0.2
9. dropout_2 = 0.5
10. dropout_3 = 0
11. data_augmentation = 'True'
12. max_pooling = 2
13. activation_1 = 'relu'
14. activation_2 = 'tanh'
15. activation_final_layer = 'softmax'
16. dense = 1024
17. optimizer = 'adam'

Output :

Training loss : 1.9 Training accuracy : 0.45

Testing loss : 1.7 Testing accuracy : 0.51

This model has very high loss score and also poor accuracy value.

6th Configuration :

1. batch_size = 32
2. num_classes = 10
3. epochs = 20
4. num_predictions = 20
5. nodes_1 = 32
6. nodes_2 = 64
7. filters = 3
8. dropout_1 = 0.2
9. dropout_2 = 0
10. dropout_3 = 0
11. data_augmentation = 'True'
12. max_pooling = 2
13. activation_1 = 'relu'
14. activation_2 = 'relu'
15. activation_final_layer = 'softmax'
16. dense = 1024
17. optimizer = 'adadelta'

Output :

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING

SECTION 02

ASSIGNMENT 1

Training loss : 0.69 Training accuracy : 0.76
Testing loss : 0.71 Testing accuracy : 0.77

Good accuracy score but high loss score.

7th Configuration :

1. hidden_layers = 2
2. batch_size = 24
3. num_classes = 10
4. epochs = 10
5. nodes_1 = 32
6. nodes_2 = 64
7. filters = 3
8. dropout_1 = 0.25
9. dropout_2 = 0.5
10. dropout_3 = 0
11. data_augmentation = 'False'
12. max_pooling = 2
13. activation_1 = 'relu'
14. activation_2 = 'relu'
15. activation_final_layer = 'softmax'
16. dense = 512
17. optimizer = 'adam'
18. num_predictions = 20

Output :

Training loss :0.95 Training accuracy :0.66

Testing loss : 0.85 Testing accuracy :0.70

Poor model because low accuracy score and high loss rate.

8th Configuration :

1. hidden_layers = 2

CSYE 7374 – SPECIAL TOPICS IN COMPUTER SYSTEMS ENGINEERING

SECTION 02

ASSIGNMENT 1

2. batch_size = 64
3. num_classes = 10
4. epochs = 25
5. nodes_1 = 32
6. nodes_2 = 64
7. filters = 3
8. dropout_1 = 0.2
9. dropout_2 = 0
10. dropout_3 = 0
11. data_augmentation = 'False'
12. max_pooling = 2
13. activation_1 = 'relu'
14. activation_2 = 'relu'
15. activation_final_layer = 'softmax'
16. dense = 1024
17. optimizer = 'sgd'
18. num_predictions = 20

Output :

Training loss : 0.86 Training accuracy :0.70

Testing loss : 0.91 Testing accuracy :0.69

This model has a poor accuracy score and high loss rate.

Architecture 7 is a good model compared to all other architectures. It has a high accuracy score and low loss rate compared to other models. I think it's the batch size that influenced the accuracy and loss.
