REPORT

CECS 551 – Assignment 5

Design neural networks to classify handwritten numbers using keras library, however, do not include convolutional layers.

Name: Aishwarya Bhavsar

CSULB ID: 029371509

Date – 15th March, 2022

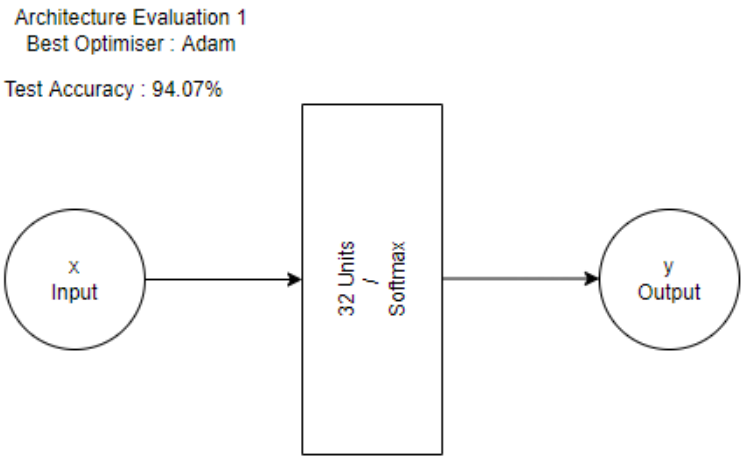
I certify that this submission is my original work – AVB

- **GOAL:** To pick the best optimizer, tune its hyper-parameters and evaluate the best architecture.
- **STEPS/PROCEDURE:**

    1. Train the Model
    2. Compile The Model
    3. Fit The Model
    4. Evaluate The Model
    5. Determine The Accuracy.

- **RESULTS:**

1. **Architecture Evaluation 1**



- The proposed model is made of one MLP layer.

- In Keras, an MLP layer is referred to as Dense, which stands for the densely connected layer.
- The hidden layer contains 32 units each, followed by relu and softmax activation and dropout.
- 128 units are chosen since 256, 512 and 1,024 units have lower performance metrics. At 128 units, the network converges quickly, but has a lower test accuracy. The added number units for 512 or 1,024 does not increase the test accuracy significantly.
- The main data structure in Keras is the Sequential class, which allows the creation of a basic neural network.
- The Sequential class of the Keras library is a wrapper for the sequential neural network model that Keras offers and can be created in the following way:
- from keras.models import Sequential
- model = Sequential()
- The model in Keras is considered as a sequence of layers and each of them gradually “distills” the input data to obtain the desired output.
- In Keras, we can add the required types of layers through the add() method.

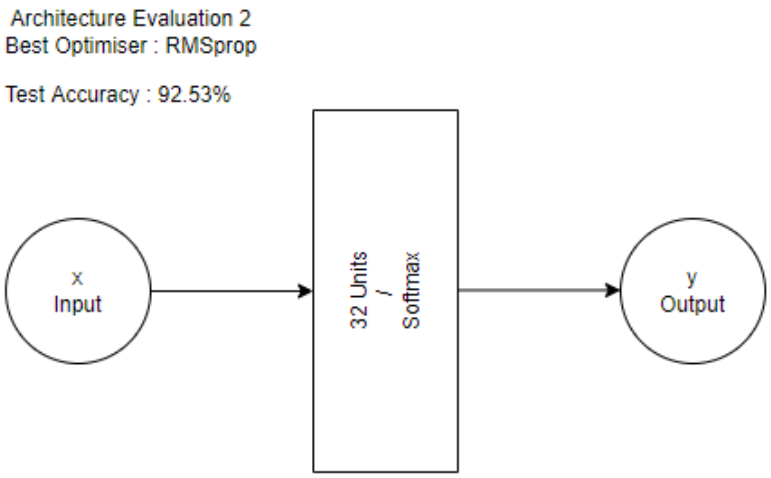
Optimisers	Training Accuracy	Testing Accuracy
Adam	95.97%	94.07%
Adagrad	51.80%	51.58%
SGD	56.98%	57.49%
RMSprop	9.87%	7.61%

- Thus, by comparing all the 4 optimisers for model architecture 1, we conclude that Adam works best when, the hyperparameters are tuned in the following way:

Hyper parameters used for Adam

1. learning_rate=0.01,
2. beta_1=0.9,
3. beta_2=0.999,
4. epsilon=1e-07,
5. amsgrad=False,
6. Activation Function : relu & softmax
7. Number Of Hidden Units : 32
8. Loss : Catergorical_crossentropy
9. Metrics: Accuracy

2. Architecture Evaluation 2



- The proposed model is made of one MLP layer.
- In Keras, an MLP layer is referred to as Dense, which stands for the densely connected layer.
- The hidden layer contains 32 units each, followed by relu, softmax, softmax activation and dropout.
- 540 units are chosen since 256, 512 and 1,024 units have lower performance metrics. At 128 units, the network converges quickly, but has a lower test accuracy. The added number units for 512 or 1,024 does not increase the test accuracy significantly.
- The main data structure in Keras is the Sequential class, which allows the creation of a basic neural network.
- The Sequential class of the Keras library is a wrapper for the sequential neural network model that Keras offers and can be created in the following way:
- from keras.models import Sequential
- model = Sequential()
- The model in Keras is considered as a sequence of layers and each of them gradually “distills” the input data to obtain the desired output.
- In Keras, we can add the required types of layers through the add() method.

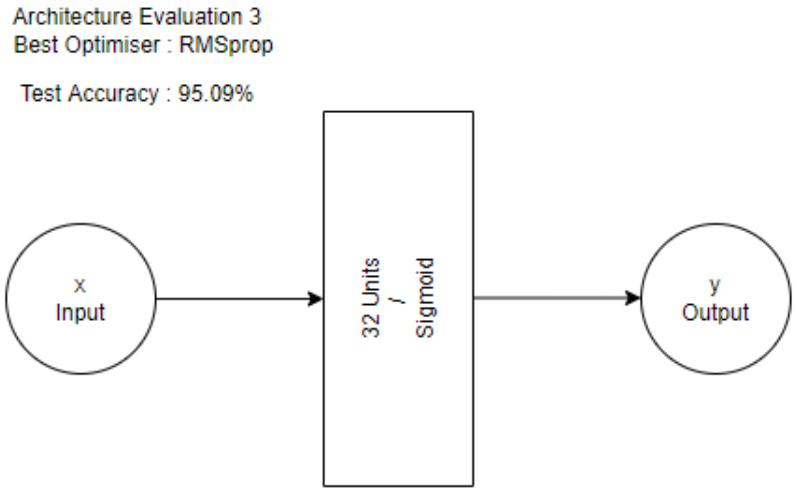
Optimisers	Training Accuracy	Testing Accuracy
Adam	90.68%	90.08%
Adagrad	82.50%	83.02%
SGD	87.84%	88.46%
RMSprop	92.53%	92.32%

- Thus, by comparing all the 4 optimisers for model architecture 1, we conclude that RMSprop works best when, the hyperparameters are tuned in the following way:

Hyper parameters used for RMSprop

1.learning_rate=0.005,
2.rho=0.2,
3.momentum=0.9,
4.epsilon=1e-07,
5.centered=False
6.Batch Size : 540
7.Dropout: 0.3
8.Activation Function : Softmax

3. Architecture Evaluation 3



- The proposed model is made of one MLP layer.
- In Keras, an MLP layer is referred to as Dense, which stands for the densely connected layer.
- The hidden layer contains 32 units each, followed by relu, sigmoid, sigmoid activation and dropout.
- 32 units are chosen since 256, 512 and 1,024 units have lower performance metrics. At 128 units, the network converges quickly, but has a lower test accuracy. The added number units for 512 or 1,024 does not increase the test accuracy significantly.
- The main data structure in Keras is the Sequential class, which allows the creation of a basic neural network.
- The Sequential class of the Keras library is a wrapper for the sequential neural network model that Keras offers and can be created in the following way:
- from keras.models import Sequential
- model = Sequential()

- The model in Keras is considered as a sequence of layers and each of them gradually “distills” the input data to obtain the desired output.
- In Keras, we can add the required types of layers through the add() method.

Optimisers	Training Accuracy	Testing Accuracy
Adam	94.88%	94.53%
Adagrad	75.83%	76.32%
SGD	92.92%	92.92%
RMSprop	95.09%	94.75%

- Thus, by comparing all the 4 optimisers for model architecture 1, we conclude that RMSprop works best when, the hyperparameters are tuned in the following way:
- Hyper parameters used for RMSprop
  1.learning_rate=0.005,
  2.rho=0.2,
  3.momentum=0.9,
  4.epsilon=1e-07,
  5.centered=False
  6.Batch Size : 32
  7.Dropout: 0.3
  8.Activation Function : Sigmoid

- **<u>CONCLUSION:</u>**
  - RMSprop optimiser is good optimiser as compared to other optimisers.
  - It gives highest accuracy when combined with sigmoid activation function.
  - Adam optimiser gives best accuracy when combined with relu and softmax activation functions.
  - The accuracy of any model depends on the activation function and not on the number of layers.
  - Therefore, choosing correct optimisers and activation functions is important and will yield fruitful results.