TIPR ASSIGNMENT 3

CONVOLUTIONAL NETWORK
K-MEANS ALGORITHM
t-SNE VISUALIZATION

SUBMITTED BY:
AKSHAYA A MUKUNDAN
M. Tech DESE
SR No. 15521

Algorithm:

Image properties:

CIFAR-10 = 32 x 32 x 3

Fashion-MNIST = 28 x 28 x 1

Input neuron to the neural network = number of channels = 3 for CIFAR-10

Input neuron to the neural network = number of channels = 1 for Fashion-MNIST

Number of class = 10 for both datasets = final output layer number of neurons.

Rest of the layers are configurable as needed

Arguments from the command prompt are parsed to obtain the train data path, test data path, dataset name, number of neurons in convolutional layer, activation function to be used in convolutional layer. All these parameters are converted to required datatype to be used in the code.

Dataset is extracted and data is reshaped in the form of (Number of data, number of row, number of column, number of channels). Also, both the datasets are normalized to a value between 0 and 1.

Labels are reshaped into the form (Number of data, Number of classes).

Function myconv: Convolutional layer weights are initialized. Convolutional layer for the required configuration is designed using tf.nn.conv2d (tf is tensorflow) and batch normalized as the datas will be processed in batches of 128(random number chosen. Usually preferred to be a number of power 2). Activation functions in tensorflow are defined. Maxpool layers of required kernel size are defined. Then fully connected layers are defined.

In main function, global variables are initialized and data is split into train and validation set. A tensorflow session is written to train, optimize, find cost and test with the validation set and calculate the validation test accuracy.

The model is saved into a Models file with checkpoints.

The graph is restored using the saved model data path. To this model test data is passed which would give the test accuracy of the test data fed and will calculate the accuracy.

To check the number of convolutional layers suitable for a dataset(**task 1**), the number of layers was varied keeping all other parameters fixed. The layer configuration with best performance was chosen for further tasks.

To check number of neurons (**task 2 part A**), all other parameters were fixed and this one was varied and best result were used in further tasks.

To check the best filter size configuration (**task 2 part B**), all other parameters were fixed and the best result was chosen.

To check which activation function suits best (**task 3**), all other parameters was fixed and activation functions- sigmoid, tanh, ReLu and swish were used for various layer configurations and best performance was checked.

For initialization(**task 4**), truncated normal initialization and xavier initializations were used.

For Clustering (**task 5**), using K-Means clustering the embeddings of the validation dataset is clustered into 10 clusters. Based on this prediction, the labels were printed and it was found that the labels were in a specific pattern. ie all the labels were to be added by a number in the range [0,9] to obtain the actual labels. Hence, an algorithm was implemented to perform this mod-addition and then clustering accuracy was measured.

For t-SNE visualization (**task 6**), the embeddings obtained were passed through t-SNE and were reduced to 2 dimensional vectors and was plotted in 2D for all the required specifications.

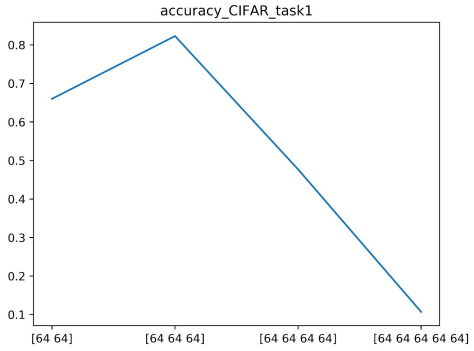
To compare CNN and MLP (**task 7**), MLP was implemented using keras and for the same specifications, accuracy and F1 Scores were measured.

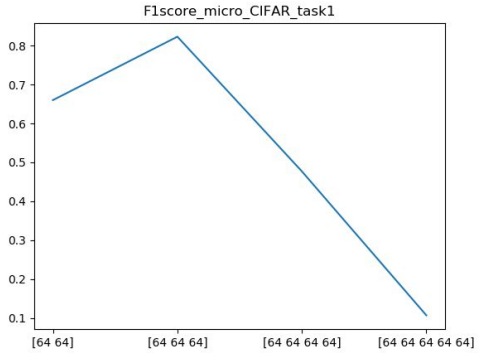
Task - I (2.5 Points)

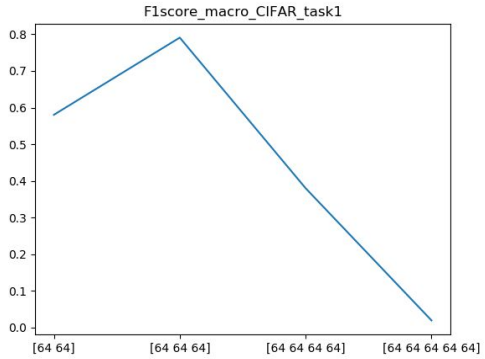
Experiment with different number of layers of your neural network and give a plot of accuracy vs. layer count, macro-F1 vs. layer count and micro-F1 vs. layer count.

CIFAR-10

Here configuration of layer is taken as the number of neurons in that layer



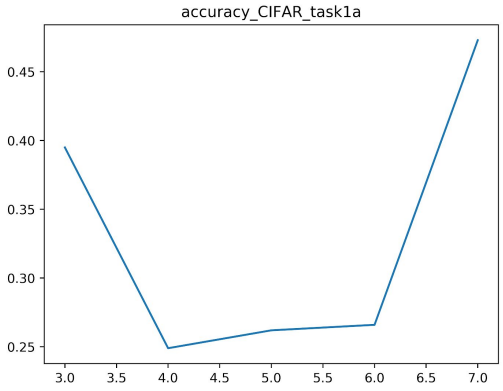


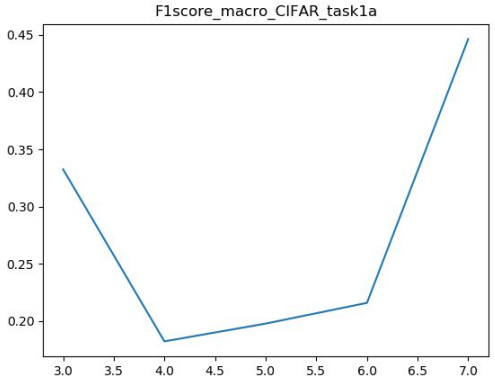


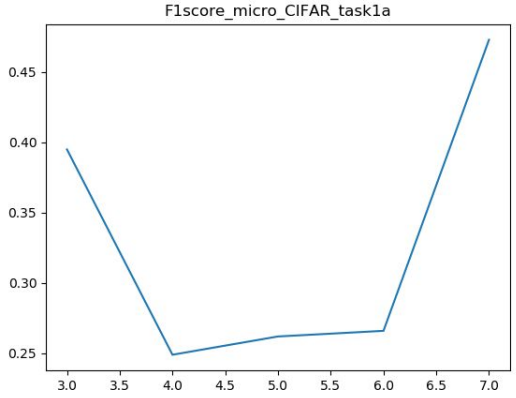
Inference:

For CIFAR dataset, the model seems to overfit as the number of layers is increased beyond 3. Maximum accuracy was obtained with 3 number of layers. An equivalent result was obtained with 2 number of layers. In all layers, the number of neuron is fixed to get accuracy independent of variation of number of neurons. Similarly, activation function chosen for all layers is relu.

With configuration of layer as filter size:





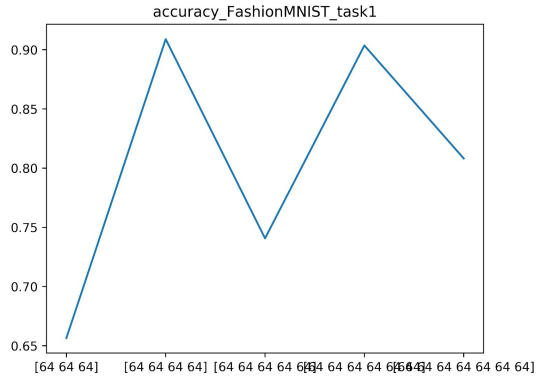


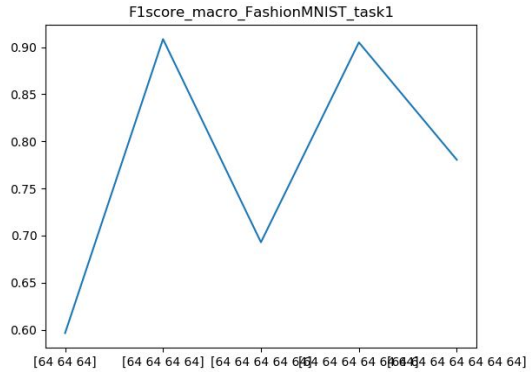
Inference:

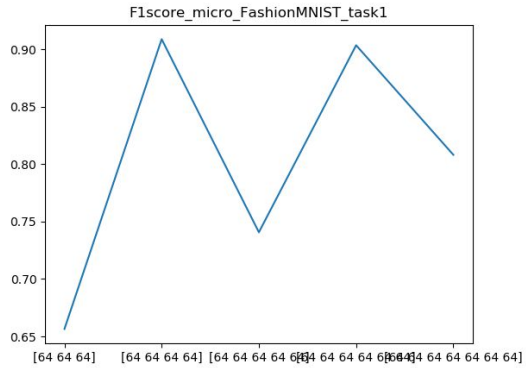
With varying parameter set as the filter-size of convolutional layer, it's observed that as the filter size is less than or equal to 3; or greater than or equal to 7 it gives a better accuracy. Here number of neurons in the layer is chosen as 64 for all layer. 2 fully connected layers are kept at the output of convolutional layer with activation function set as relu. Final output layer of the neural network is also a fully connected layer with activation function set as None. For further tasks, number of layers is fixed as 2 or 3.

Fashion-MNIST

Here configuration of layer is taken as the number of neurons in that layer



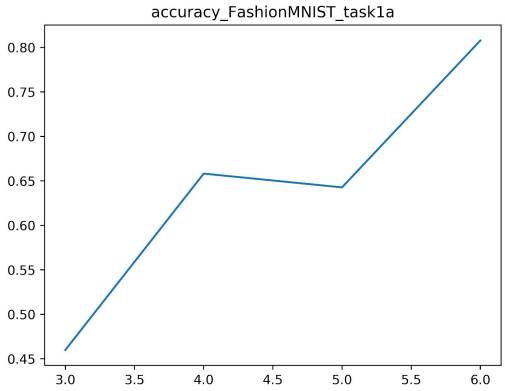


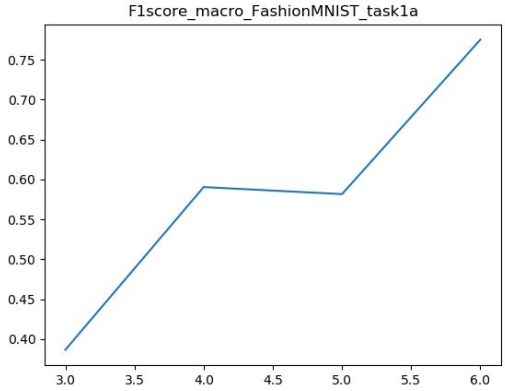


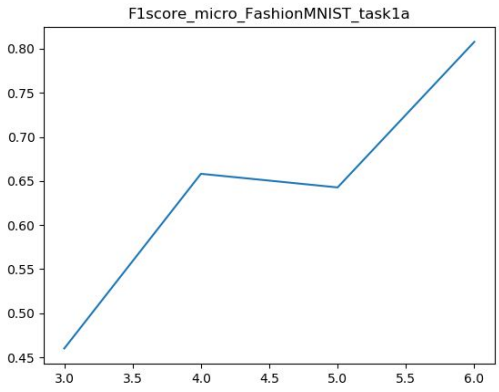
Inference:

Compared to CIFAR-10 dataset, Fashion-MNIST depicts very high accuracy, in the range of 80-95%. Here number of neurons is taken as the filter configuration. All activation functions are set to ReLu. To get no dependence with number of neurons in each layer, it's kept fixed as 64. With layer number variation, we find the accuracy is high as the number of layers increases. Optimal number of layer can be fixed as 4 or 6.

With configuration of layer as filter size:







Inference:

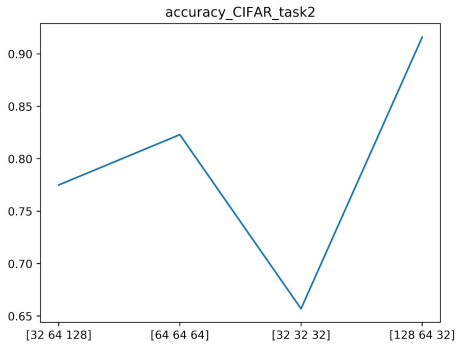
As the number of layers increases, the metrics(accuracy and F1 scores) are found to be increasing. Hence, for Fashion-MNIST dataset, high number of layers is desired. In all layers, the number of neurons is fixed as 64, and activation function chosen is ReLu. Hence, for this dataset number of layers is fixed as 6 for further tasks.

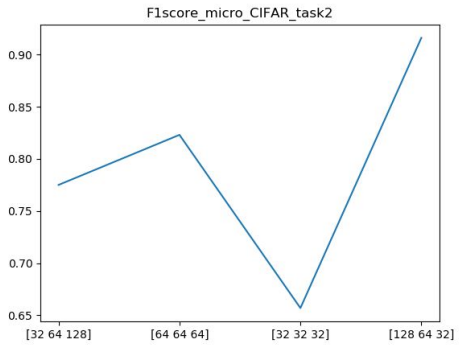
Task - II (2.5 Points)

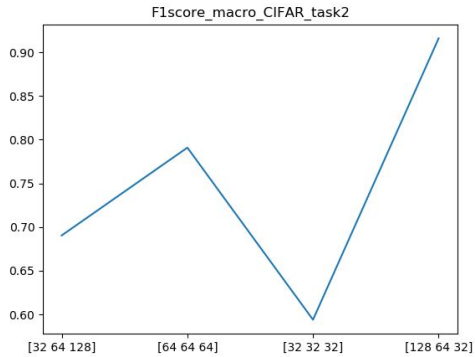
Try out various configurations of number of neurons in each layer and nd the best one. Provide a suitable plot of the same.

CIFAR-10

Here configuration of layer is taken as the number of neurons in that layer



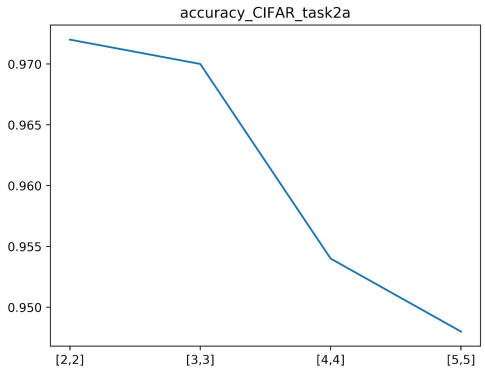


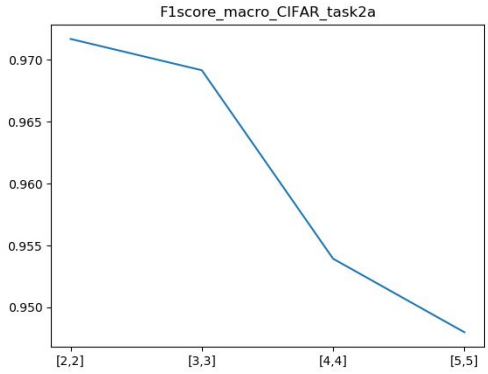


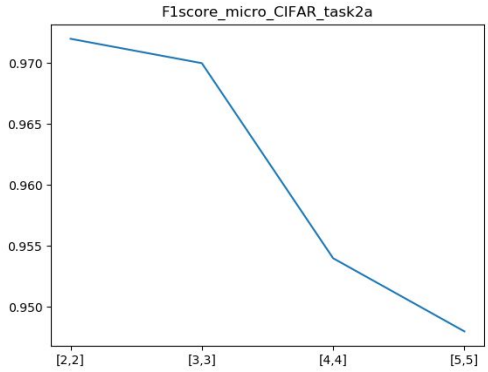
Inference:

For CIFAR-10 dataset, low number of layers used to give high accuracy. Hence, the number of layers is fixed as 3. Activation used is also kept same as that in task 1(ReLu) to make the resultant metrics independent of these parameters. Metrics are found to improve as the number of neurons in input layer is high. Although, an equivalent grade of accuracy is obtained by setting all layer number of neurons as 64.

With configuration of layer as filter size:





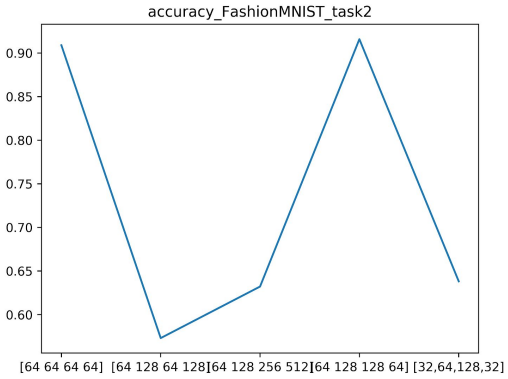


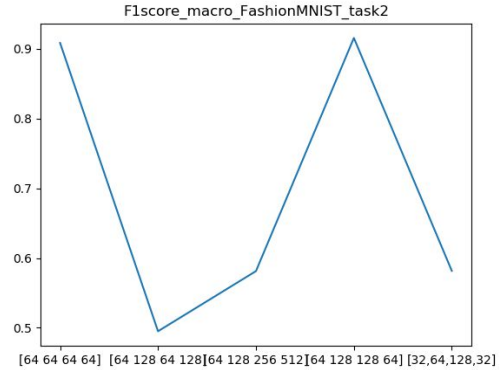
Inference:

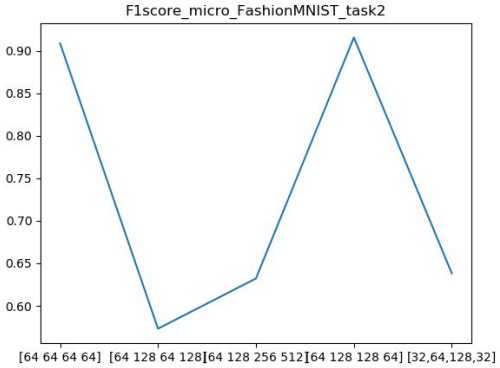
From earlier results, layer number was chosen as 2(high accuracy in task 1), number of neurons in each layer fixed as 64(as obtained from the previous result) and the activation function used is ReLu itself. With the filter-size as the varying parameter, it was found that the metrics had good values with low filter sizes. With 97.2% accuracy with filter size as 2 in both the convolutional layers.

Fashion-MNIST

Here configuration of layer is taken as the number of neurons in that layer



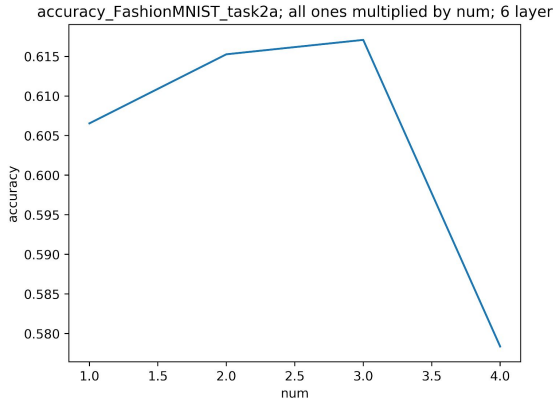


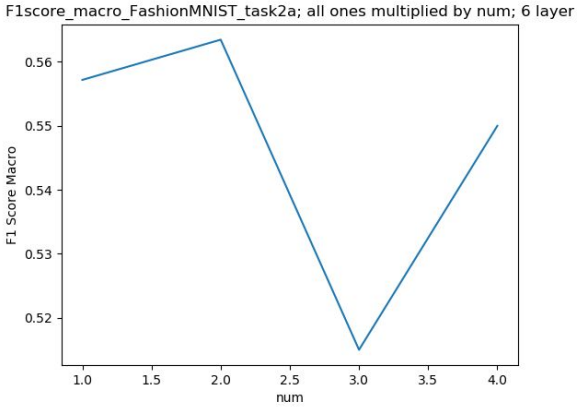


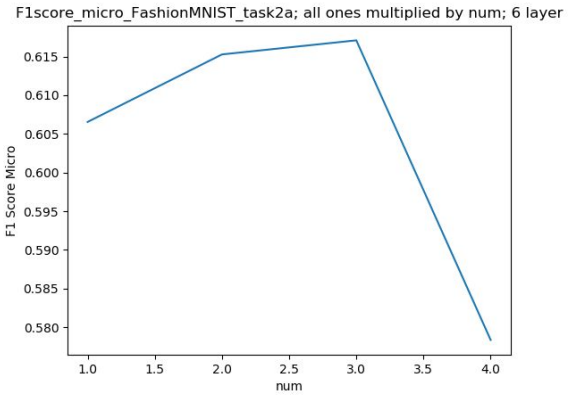
Inference:

For Fashion-MNIST dataset, the metrics evaluated gave good performance when the number of neurons in each layer is kept same as 64. It also gave an equivalent performance for a symmetric layer configuration of [64 128 128 64]. To generalize for both the datasets, the number of neurons in each layer is fixed as 64.

With configuration of layer as filter size:







Inference:

All the layers have 64 number of neurons(from previous result) with activation function chosen as ReLu.

As the filter size increases, the accuracy is found to drop rapidly for Fashion MNIST also.

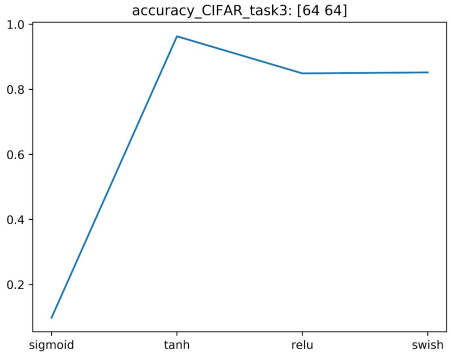
Hence, for further tasks the filter size is chosen in the range [2,3].

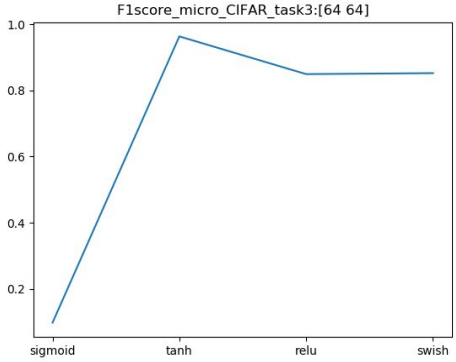
Task - III (2.5 Points)

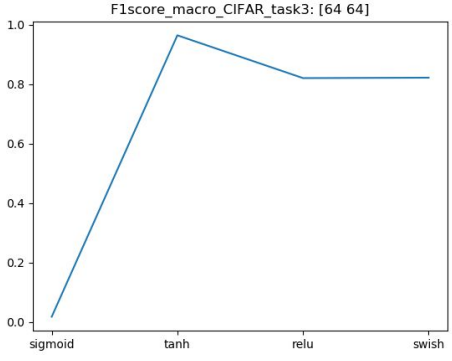
Fix the number of layers and neurons in each layer and feed the hidden layer with the following activation functions - sigmoid, tanh, relu and swish. Evaluate the performance of your model. Do it for different configurations of your neural net. Provide a suitable plot of the results obtained.

CIFAR-10

a) 2 convolutional layers; [64 64] layer configuration; [3 3] filter size



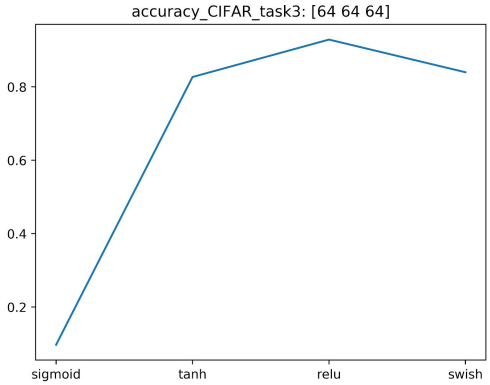


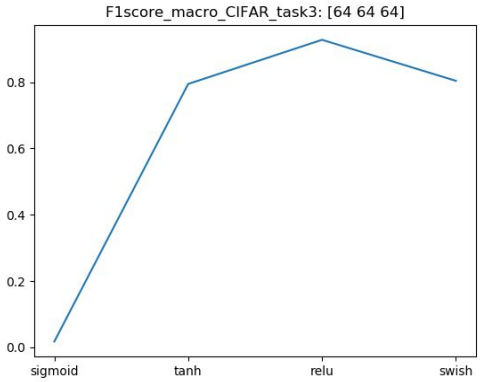


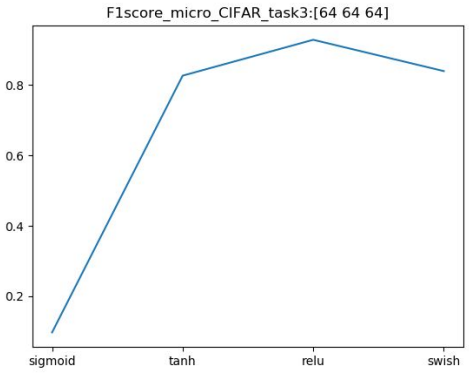
Inference:

With the above layer configurations(fixed based on results obtained in previous tasks), tanh function is found to provide highest accuracy.

b) 3 layers; [64 64 64] layer configuration; [3 3 3] filter size







Inference:

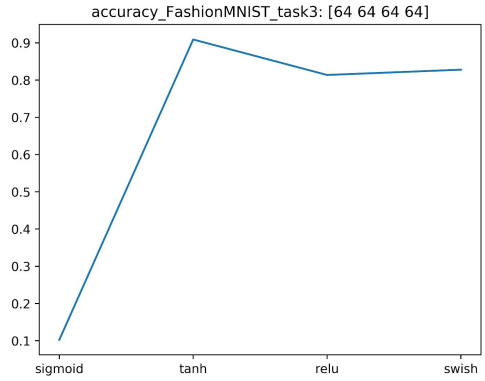
Since, CIFAR gave an equivalently better performance with 3 number of convolutional layers, activation function was varied for this configuration too.

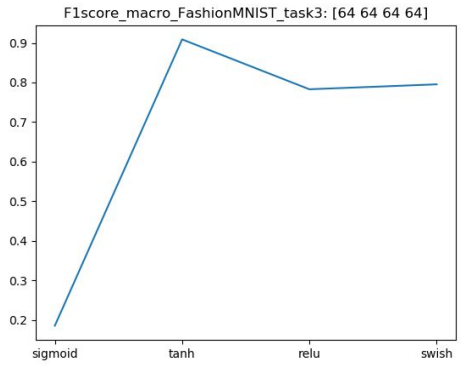
With the above layer configurations (fixed based on results obtained in previous tasks), ReLu function is found to provide highest accuracy.

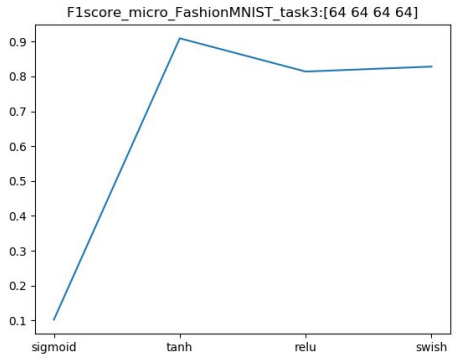
Hence, both tanh and ReLu gave good performance for CIFAR-10 dataset.

Fashion MNIST

a) 4 layers; [64 64 64 64] layer configuration; [3 3 3 3] filter size



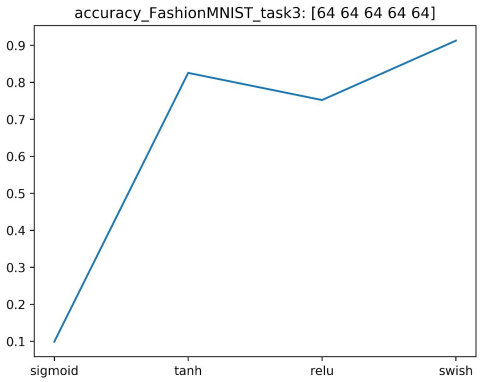


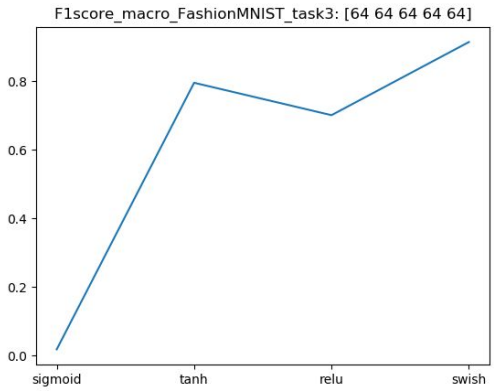


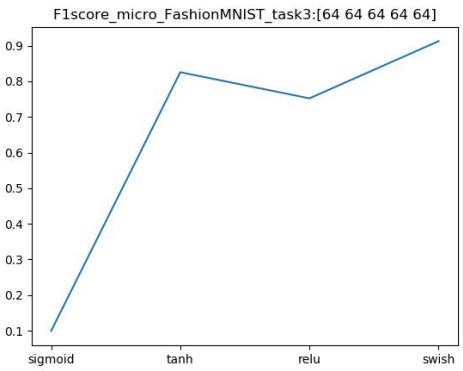
Inference:

For Fashion-MNIST dataset, higher accuracy were obtained when the number of convolutional layers were more than 4. Hence, for a 4 layer configuration, different activation functions were tried. It was found that as in CIFAR-10, metrics values were very poor for sigmoid activation function. For all other activation functions, metrics values were really good.

b) 5 layers; [64 64 64 64 64] layer configuration; [3 3 3 3 3] filter size



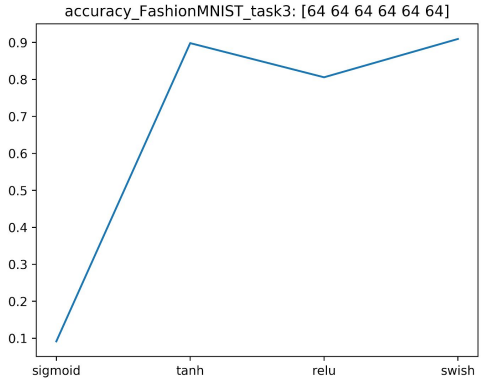


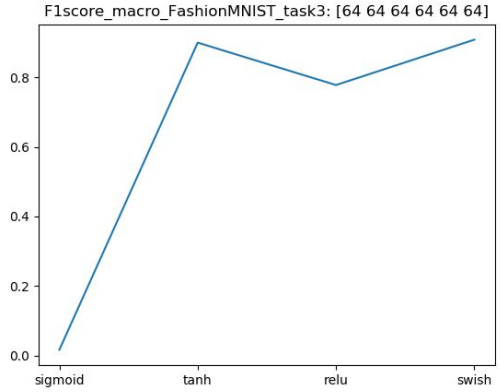


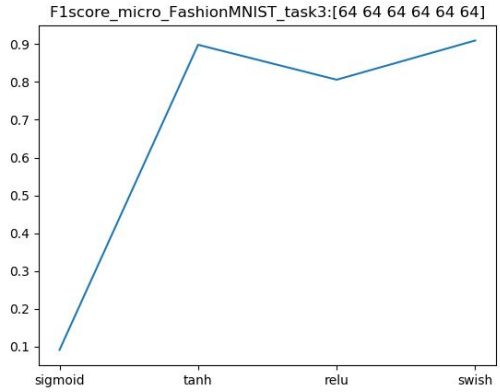
Inference:

With a 5 layer configuration also, metric parameters were good for all activation functions, except sigmoid.

c) 6 layers; [64 64 64 64 64 64] layer configuration; [3 3 3 3 3 3] filter size







Inference:

With a 6 layer configuration also, metric parameters were good for all activation functions, except sigmoid. With best result for tanh activation function.

Task - IV (2.5 Points)

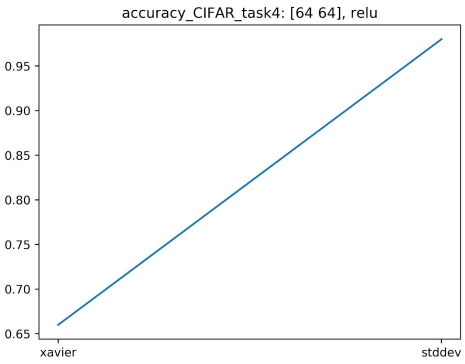
Try out various initialization techniques and observe the results.

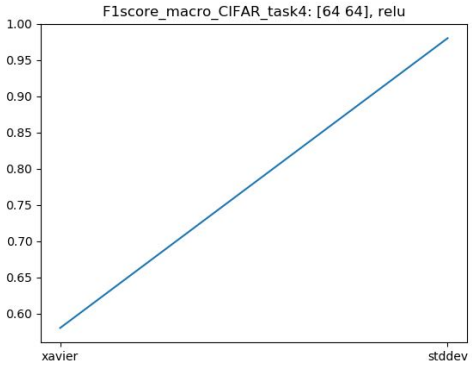
CIFAR-10

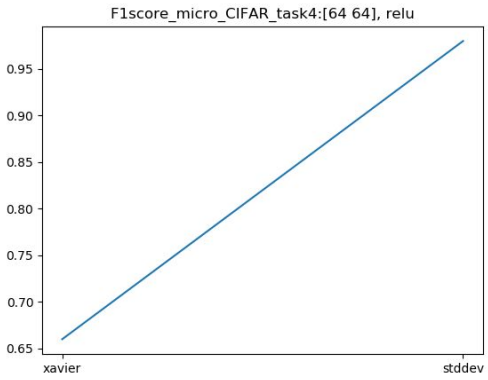
2 layer, [64 64], relu activation function, learning rate = 0.001

Xavier Initialization:
Accuracy = 66%
F1 Score Macro = 0.58035
F1 Score Micro = 0.66

Truncated normal with Standard deviation = 0.1
Accuracy = 98%
F1 Score Macro = 0.98008
F1 Score Micro = 0.98







Inference:

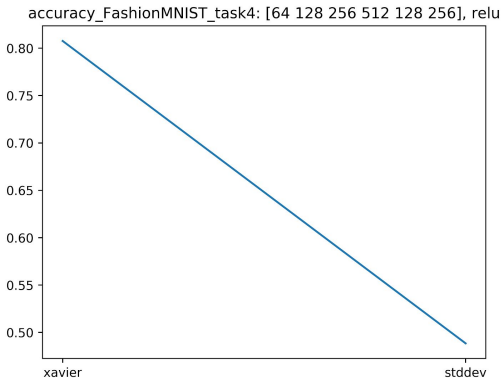
For CIFAR-10 dataset, initialization with a truncated normal function of mean zero and standard deviation of 0.08 was found to give the best result. Hence, when the dataset is CIFAR-10, weights are chosen using truncated normal.

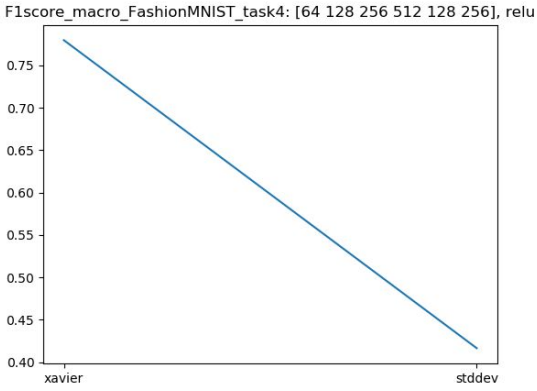
Fashion MNIST

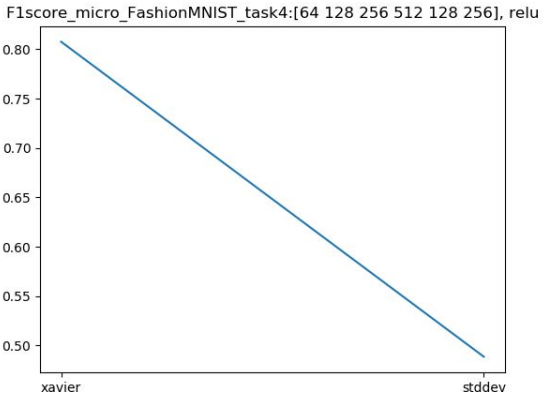
6 layer; [64 128 256 512 128 256]; relu activation function; learning rate = 0.001

Xavier Initialization:
Accuracy = 80.764%
F1 Score Macro = 0.77964
F1 Score Micro = 0.80764

Truncated normal with standard deviation = 0.1
Accuracy = 48.85%
F1 Score Macro = 0.4166
F1 Score Micro = 0.4885







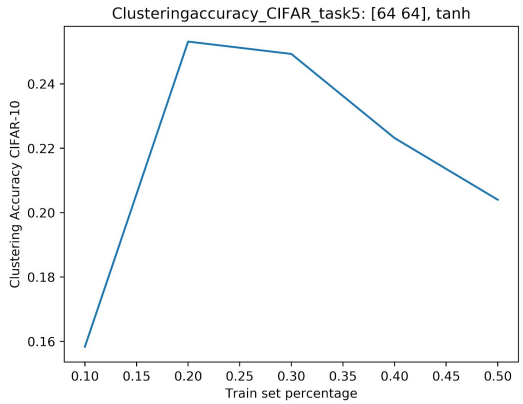
Inference:

For Fashion-MNIST dataset xavier initialization was found to give much better results, hence weights were initialized in further tasks using xavier initialization technique for Fashion-MNIST dataset.

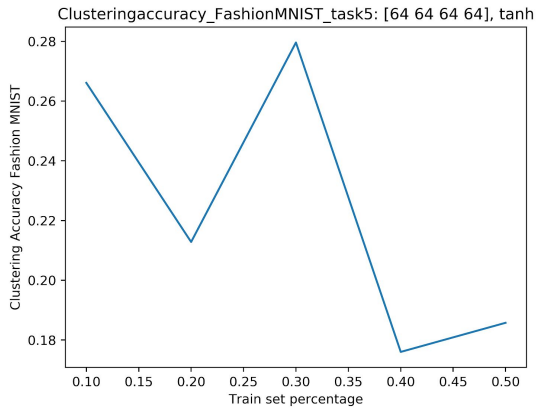
Task-V (5 Points)

Use the following semi-supervised approach to generate embeddings(vectorial representation of data samples) and evaluate clustering accuracy. Take 10-50% of the entire dataset in steps of 10% for training purpose and generate the embeddings (output of the penultimate layer of your CNN) on the remaining data. Fix embedding-dimension of your choice(which gives best result). To check the definition of clustering accuracy

CIFAR-10



Fashion MNIST



Inference:

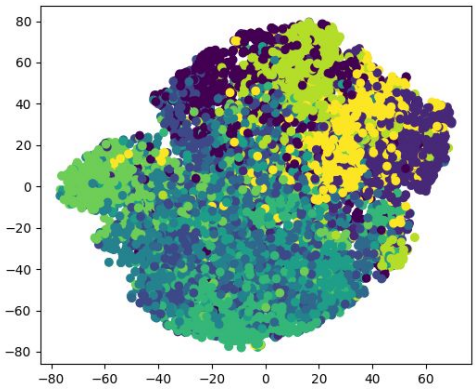
The clustering accuracy obtained for both the datasets are low. Fashion-MNIST shows a much better result. K-Means clustering was performed using sklearn.cluster package and based on these 10 clusters formed, clustering accuracy was predicted for the testset. From the whole dataset 10-50% in steps of 10% was used as training set. In above plot, the x axis shows the amount of training percent used. Y-axis represents clustering accuracy for CIFAR-10 and Fashion-MNIST datasets respectively. Clustering accuracy is found to be in the range of 15-28% for CIFAR-10 and 20-30% for Fashion-MNIST. The low clustering accuracy might be due to the low train percent of data.

Task-VI (5 Points)

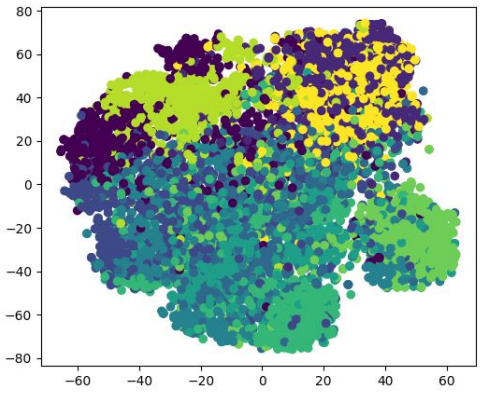
Plot a t-SNE visualization in 2-dimensional space on the obtained embedding in the previous task and observe if they are clustered into distinct groups as Expected.

CIFAR-10

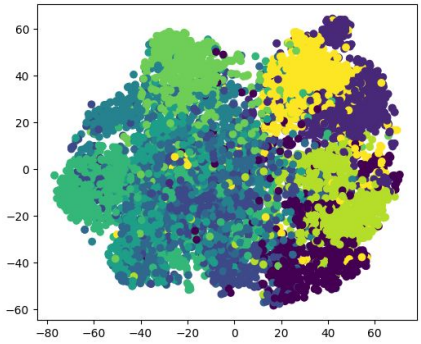
Training - 10%



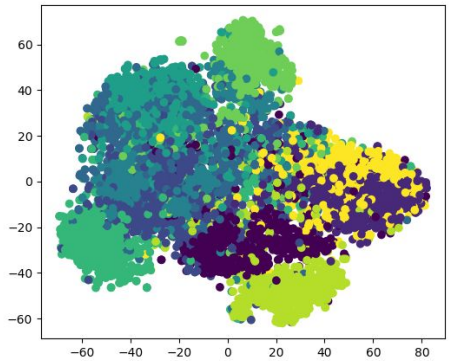
Training - 20%



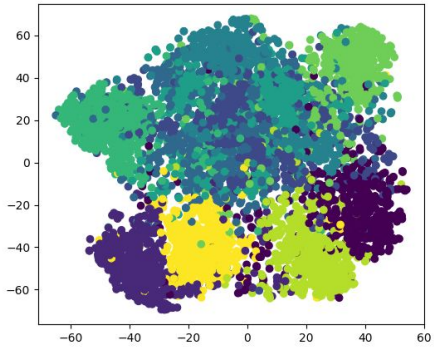
Training - 30%



Training - 40%



Training - 50%

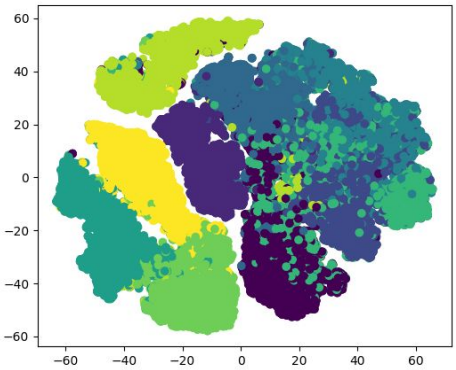


Inference:

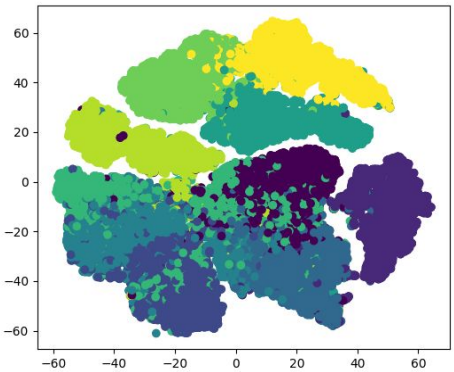
It is found that as the train data percent increases, definite clusters are formed. With 50% train data more than 6 definite clusters are formed. Rest of them are not clusterized properly as we saw that the clustering accuracy was low for CIFAR-10

Fashion-MNIST

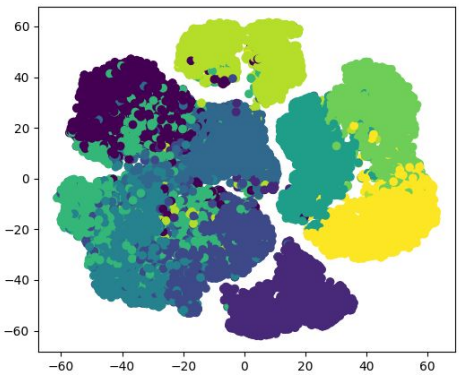
Training - 10%



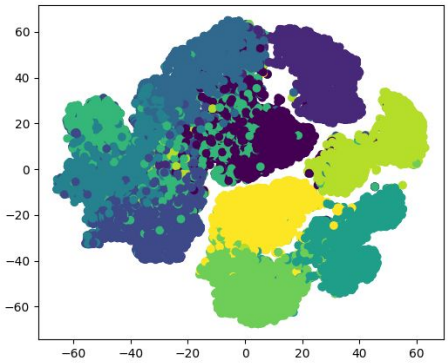
Training - 20%



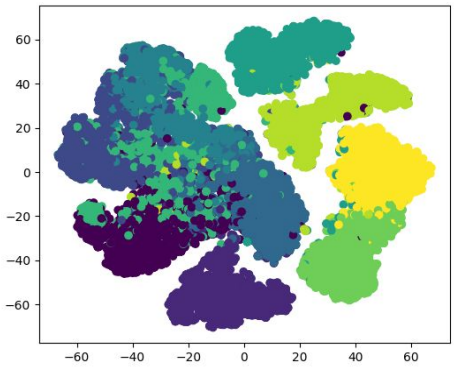
Training - 30%



Training - 40%



Training - 50%



Inference:

With 10% data for training itself, almost 8 definite clusters were formed. As the training percent increases, the clusters become more separate and clear. Hence, Fashion-MNIST shows better clustering performance as was observed with the clustering accuracy obtained for the dataset.

Task-VII (5 Points)

Compare your CNN results with MLP(either use keras or the one you designed in the previous assignment).

CIFAR-10

CNN: 92.9% (result already shown as plots in previous tasks)

MLP: 41.1%

MLP result:

```
Epoch 1/5
50000/50000 [==============================] - 39s 784us/sample - loss: 1.9715 - acc: 0.2899
Epoch 2/5
50000/50000 [==============================] - 39s 771us/sample - loss: 1.8339 - acc: 0.3334
Epoch 3/5
50000/50000 [==============================] - 38s 756us/sample - loss: 1.7961 - acc: 0.3490
Epoch 4/5
50000/50000 [==============================] - 38s 762us/sample - loss: 1.7745 - acc: 0.3573
Epoch 5/5
50000/50000 [==============================] - 41s 814us/sample - loss: 1.7510 - acc: 0.3679
10000/10000 [==============================] - 1s 147us/sample - loss: 1.6499 - acc: 0.4110
[1.64989410572052, 0.411]
```

Fashion-MNIST

CNN: 90.89% (result already shown as plots in previous tasks)

MLP: 86.8 %

MLP result:

```
Epoch 1/5
60000/60000 [==============================] - 15s 257us/sample - loss: 0.4975 - acc: 0.8211
Epoch 2/5
60000/60000 [==============================] - 14s 233us/sample - loss: 0.3811 - acc: 0.8621
Epoch 3/5
60000/60000 [==============================] - 14s 237us/sample - loss: 0.3516 - acc: 0.8718
Epoch 4/5
60000/60000 [==============================] - 15s 253us/sample - loss: 0.3271 - acc: 0.8785
Epoch 5/5
60000/60000 [==============================] - 15s 243us/sample - loss: 0.3123 - acc: 0.8837
10000/10000 [==============================] - 1s 59us/sample - loss: 0.3586 - acc: 0.8680
[0.3586395553588867, 0.868]
```

Conclusion:

Of the 2 dataset, Fashion-MNIST is found to provide better result with the code implemented for majority of the cases. Although CIFAR-10 shows better accuracy at certain configurations.

The best results using train and validation sets obtained are:

CIFAR-10

Number of layers = 2
Convolutional layer number of neurons = [64 64]
Convolutional layer filter size = [2 2]
Activation function = ReLu
Truncated normal initialization with mean zero and standard deviation as 0.08
Fully connected layers with number of neurons = [128 64 10]
First 2 fully connected layer activation = ReLu
Final fully connected layer activation = None

Accuracy = 98%
F1 Score Macro = 98.008%
F1 Score Micro = 98%

Fashion-MNIST

Number of layers = 4
Convolutional layer number of neurons = [64 64 64 64]
Convolutional layer filter size = [2 2]
Activation function = swish
Xavier initialization
Fully connected layers with number of neurons = [128 64 10]
First 2 fully connected layer activation = ReLu
Final fully connected layer activation = None

Accuracy = 91.2727%
F1 Score Macro = 91.3579%
F1 Score Micro = 91.2727%

Clustering accuracy and t-SNE plots showed good results for Fashion-MNIST when compared to the performance of CIFAR-10
Also, CNN shows far better results for the 2 datasets than MLP using keras.