American International University-Bangladesh

Mid assignment – Spring 21-22

Course: Computer Vision and Pattern Recognition [MScCS]

Course Teacher: Dr. Debajyoti Karmaker

Submitted by:

Name: Salsabil, Umme Sadia

Id: 21-92098-2

Dataset: cifar10

Method: CNN

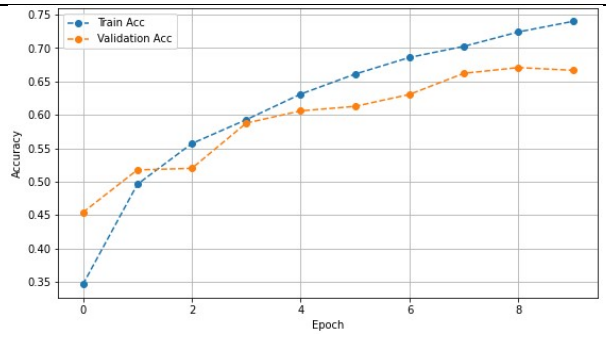
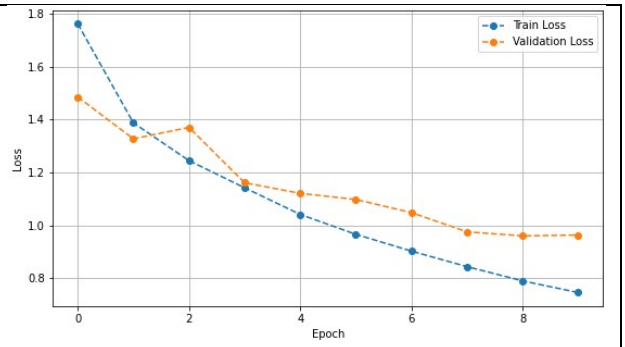
Optimizer used: Adam, SGD, RMSprop

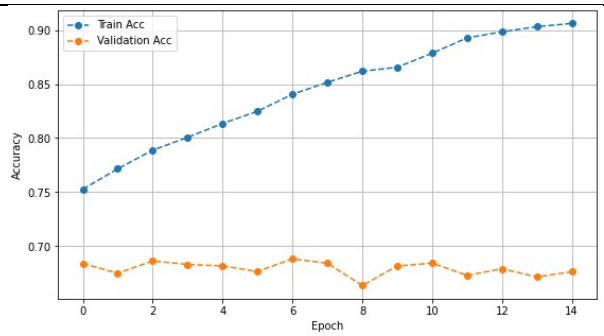
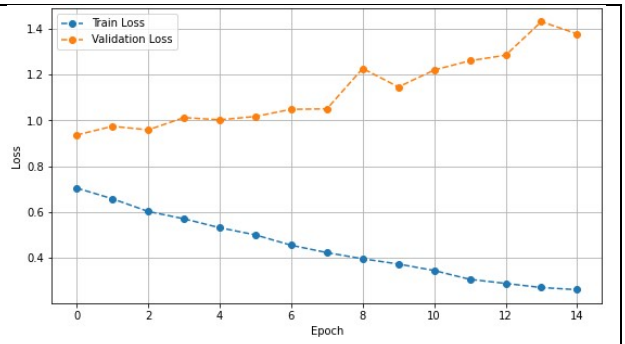
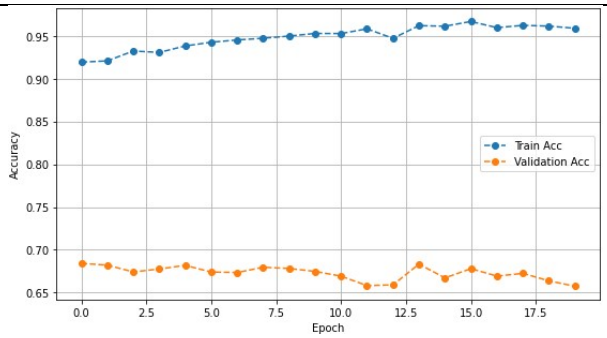
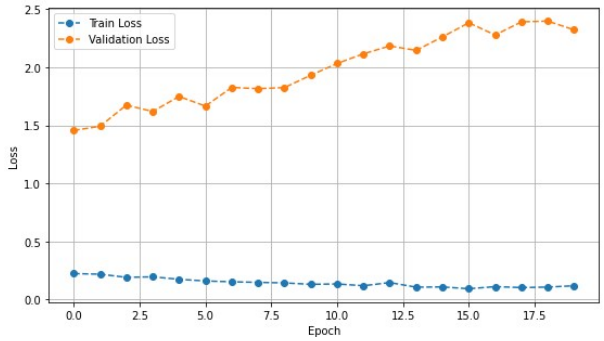
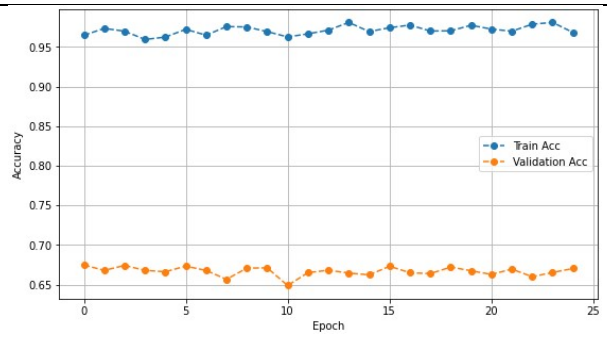
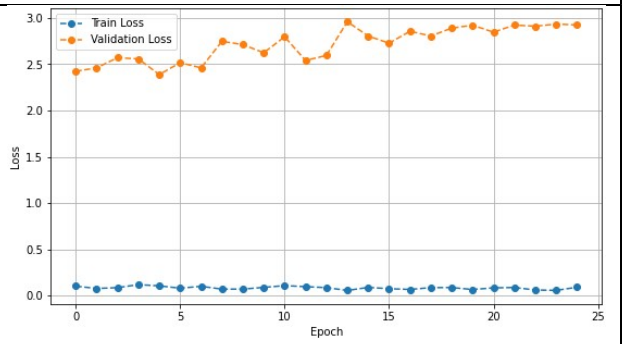
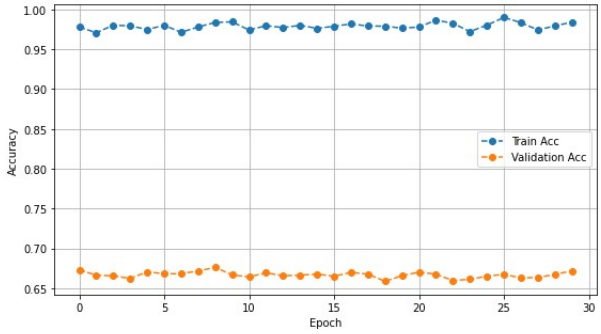
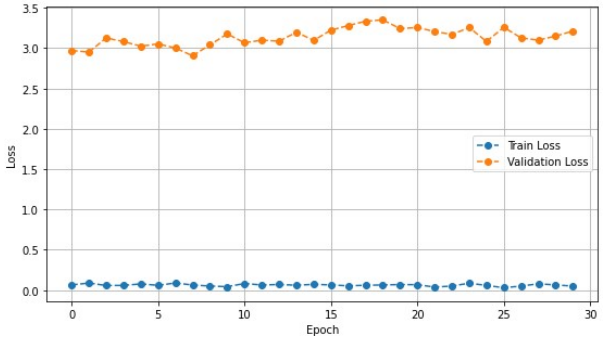
Loss function used: Categorical crossentropy, Sparse categorical crossentropy, Kullback leibler divergence

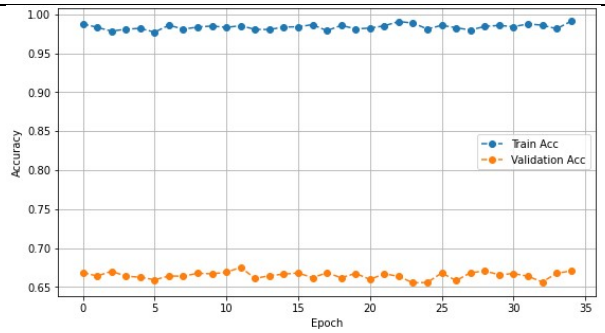
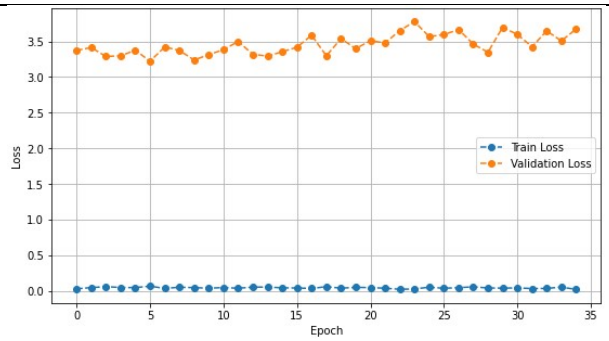
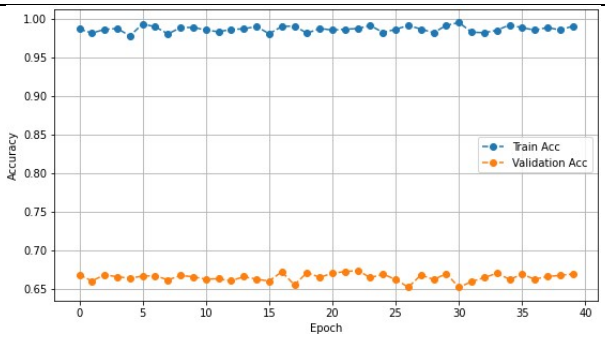
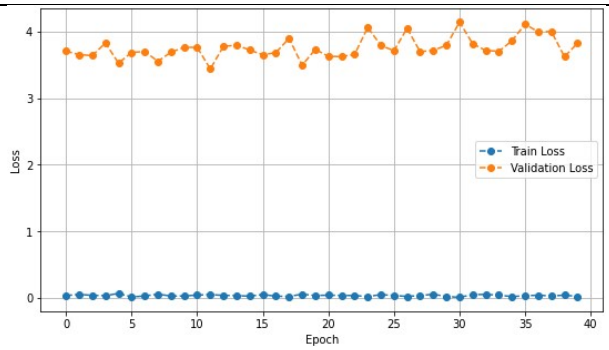
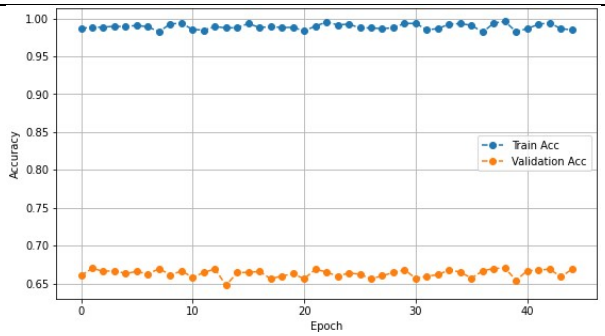
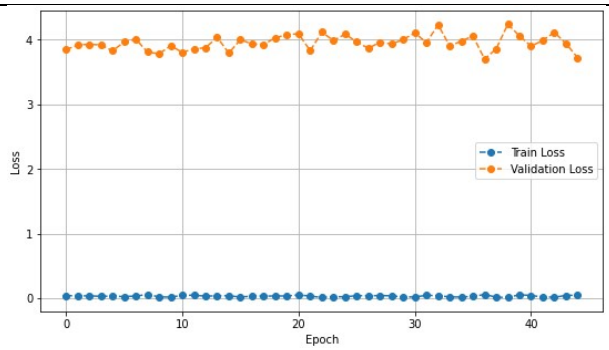
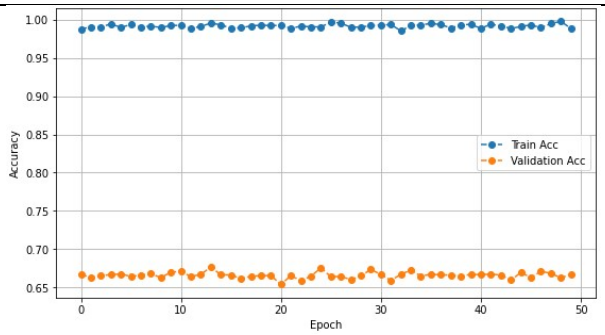
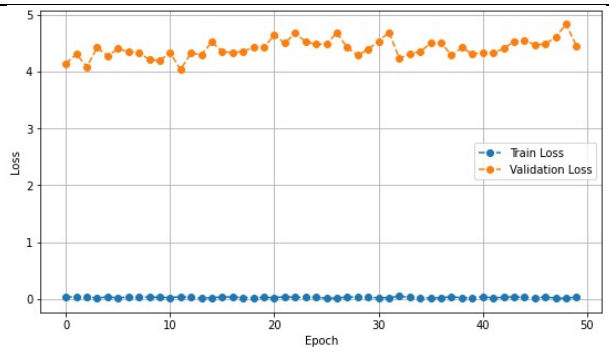
Experiment 1-3: Adam

Optimizer	Adam
Batch size	128
Learning rate	0.001
Validation split	0.3
Epoch	10-50

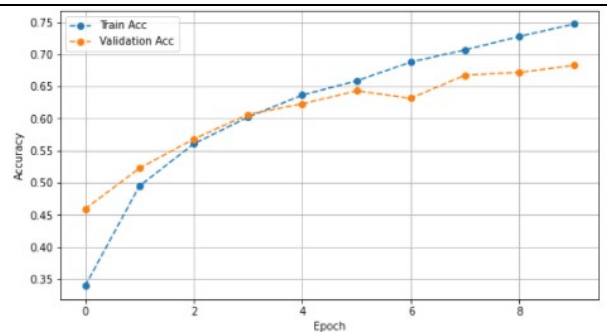
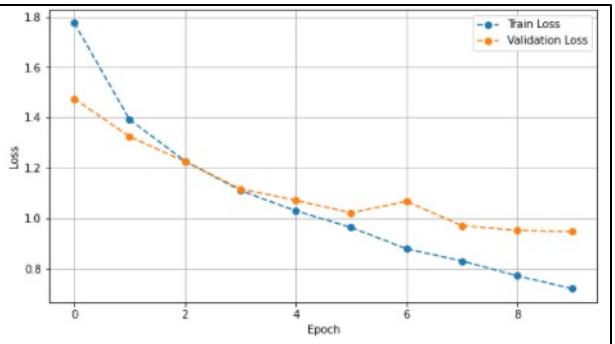
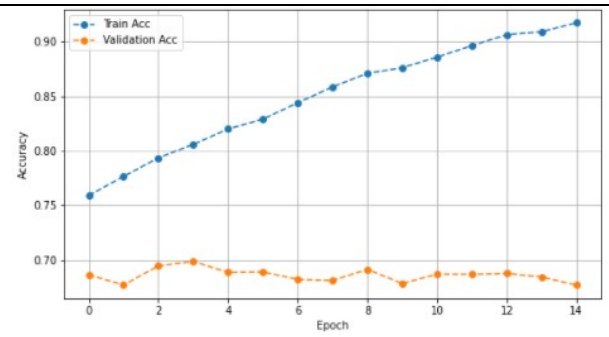
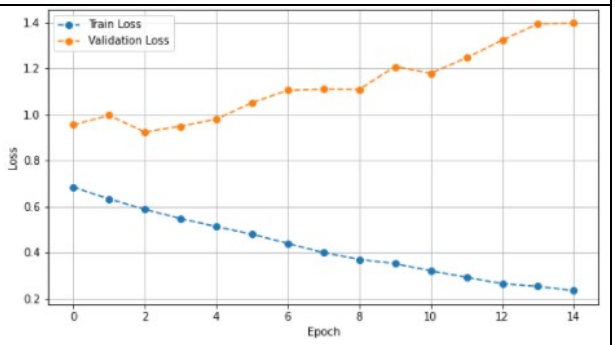
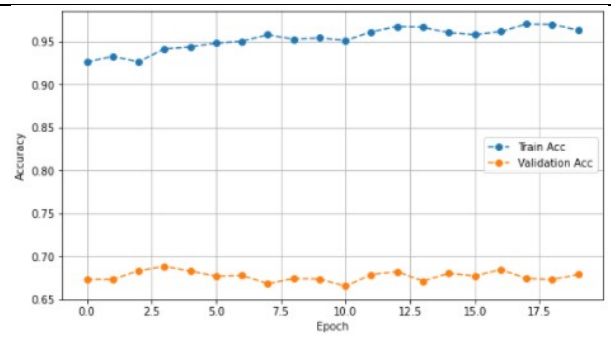
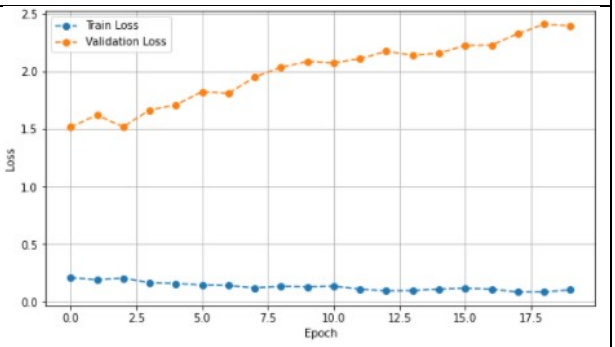
1. Loss function = Categorical crossentropy

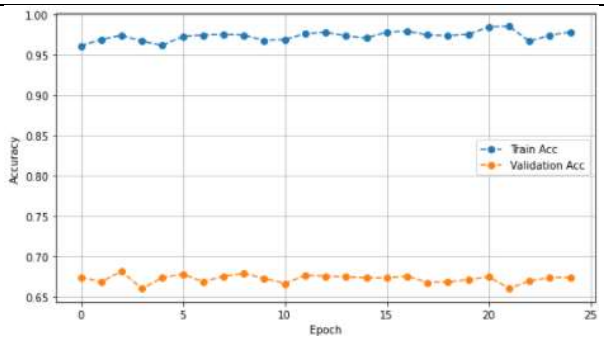
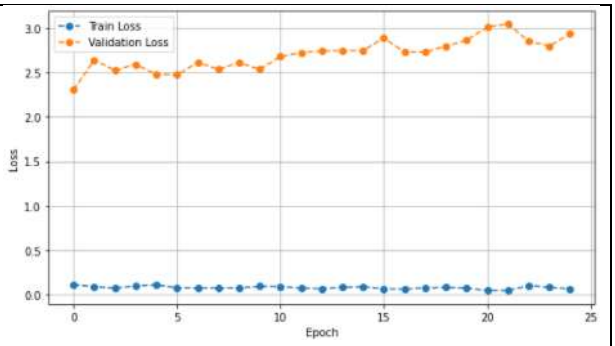
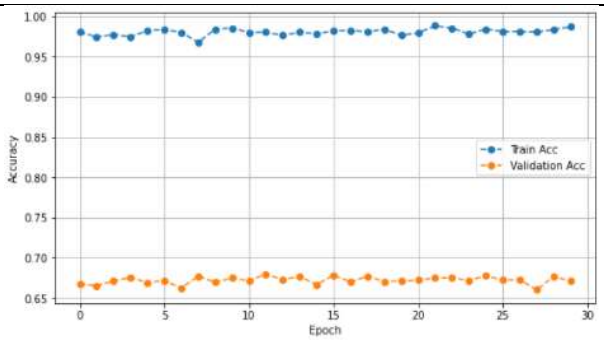
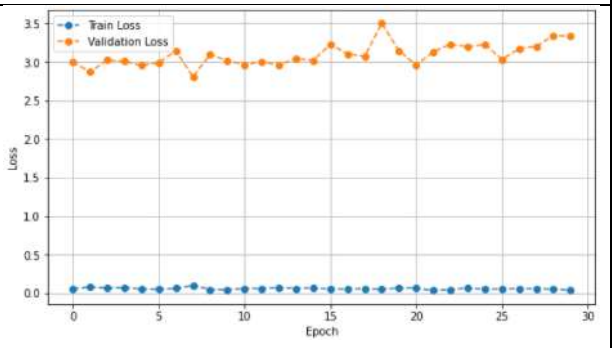
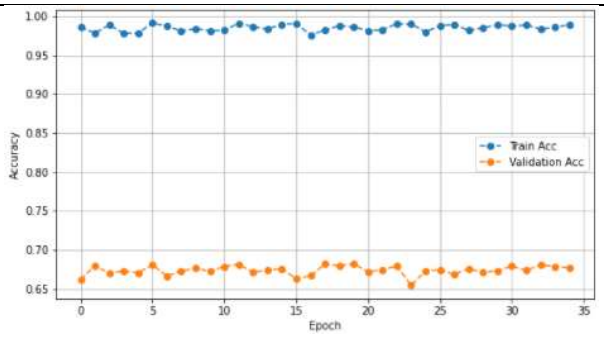
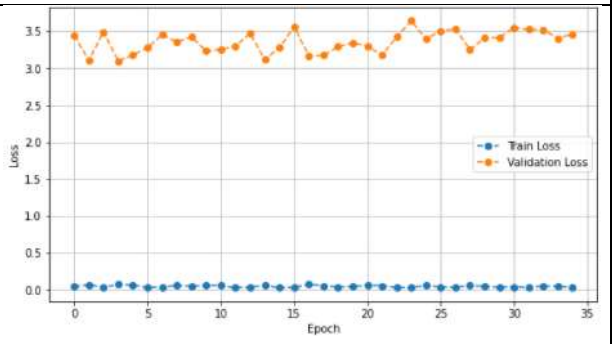
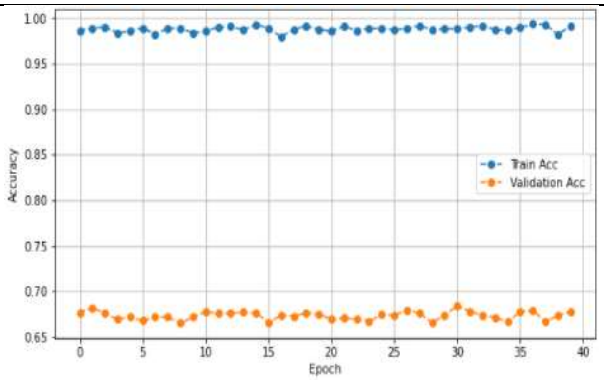
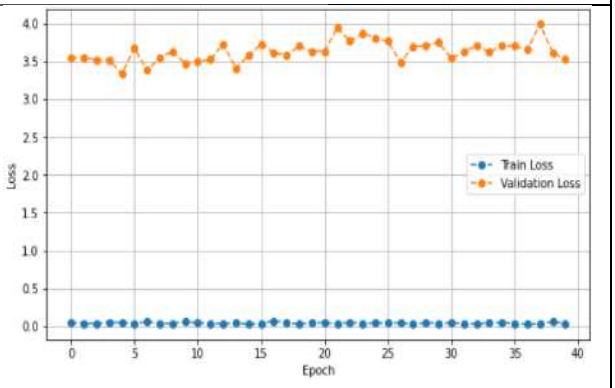
Epoch 10	 Training accuracy: 0.7400 Validation accuracy: 0.6667	 Training loss: 0.7451 Validation loss: 0.9628

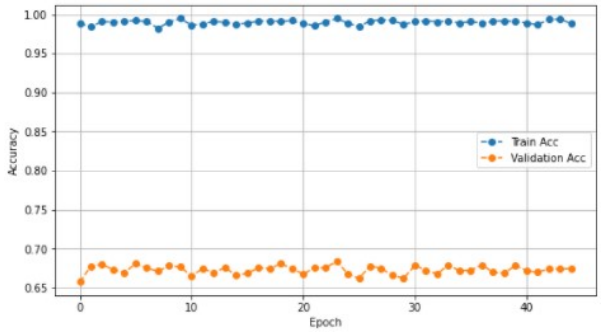
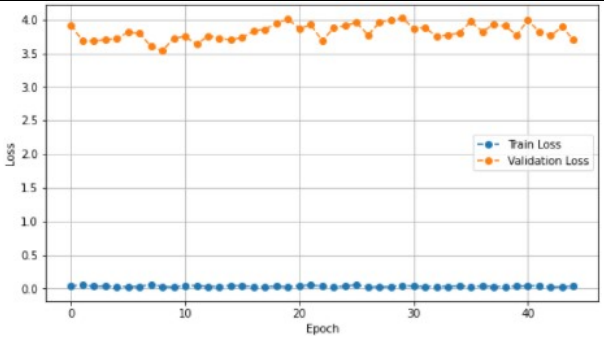
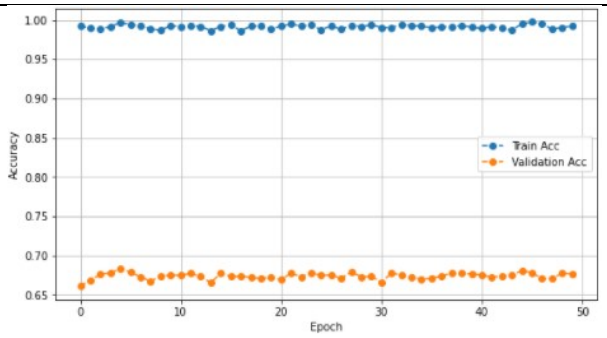
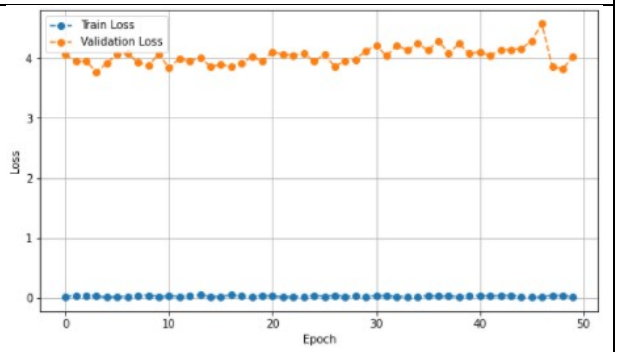
Epoch 15	 Training accuracy: 0.9062 Validation accuracy: 0.6763	 Training loss: 0.2603 Validation loss: 1.3779
Epoch 20	 Training accuracy: 0.9591 Validation accuracy: 0.6573	 Training loss: 0.1187 Validation loss: 2.3233
Epoch 25	 Training accuracy: 0.9684 Validation accuracy: 0.6703	 Training loss: 0.0897 Validation loss: 2.9236
Epoch 30	 Training accuracy: 0.9839 Validation accuracy: 0.6717	 Training loss: 0.0500 Validation loss: 3.2131

Epoch 35	 Training accuracy: 0.9912 Validation accuracy: 0.6706	 Training loss: 0.0256 Validation loss: 3.6732
Epoch 40	 Training accuracy: 0.9907 Validation accuracy: 0.6692	 Training loss: 0.0287 Validation loss: 3.8240
Epoch 45	 Training accuracy: 0.9849 Validation accuracy: 0.6687	 Training loss: 0.0506 Validation loss: 3.7251
Epoch 50	 Training accuracy: 0.9885 Validation accuracy: 0.6667	 Training loss: 0.0401 Validation loss: 4.4480

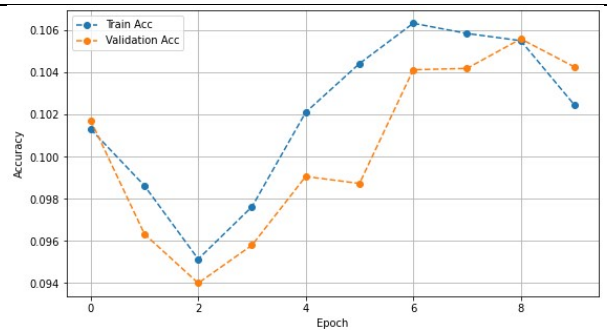
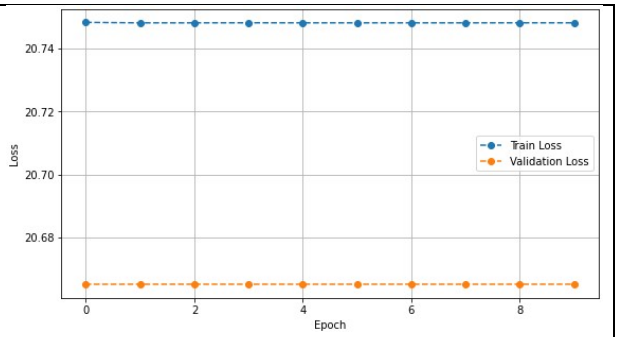
2. Loss function = Sparse categorical crossentropy

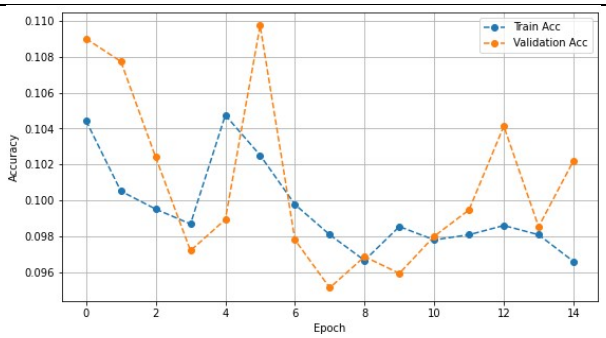
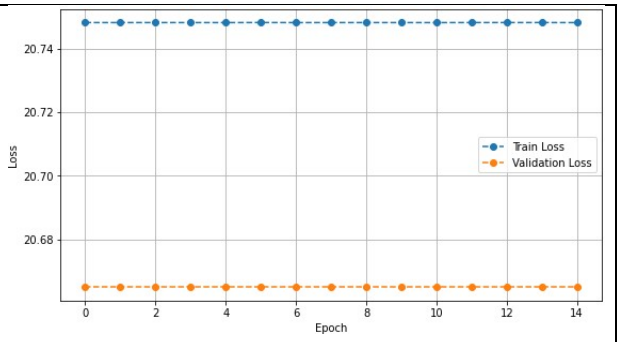
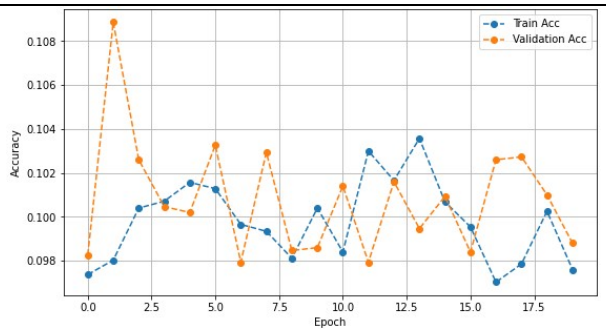
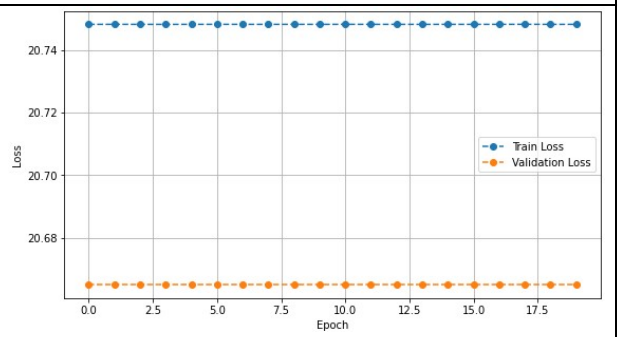
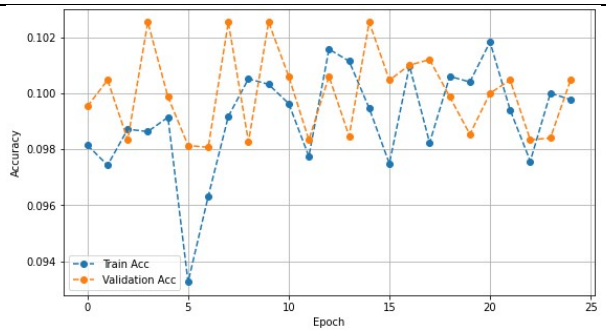
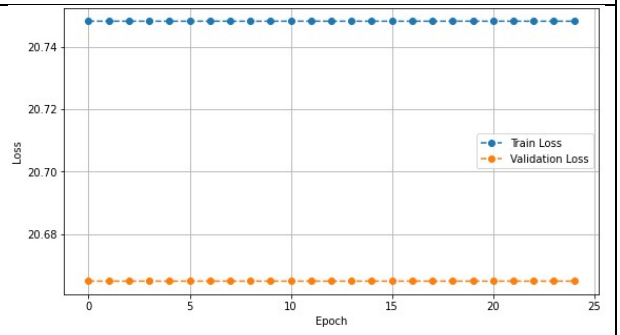
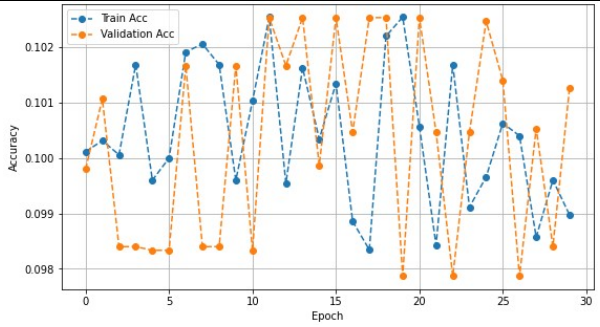
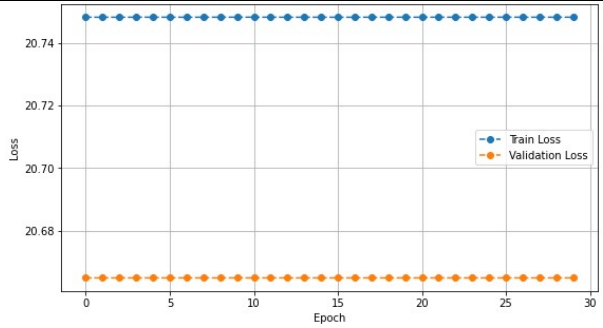
Epoch 10	 Training accuracy: 0.7469 Validation accuracy: 0.6825	 Training loss: 0.7201 Validation loss: 0.9473
Epoch 15	 Training accuracy: 0.9173 Validation accuracy: 0.6768	 Training loss: 0.2341 Validation loss: 1.3975
Epoch 20	 Training accuracy: 0.9628 Validation accuracy: 0.6786	 Training loss: 0.1049 Validation loss: 2.3946

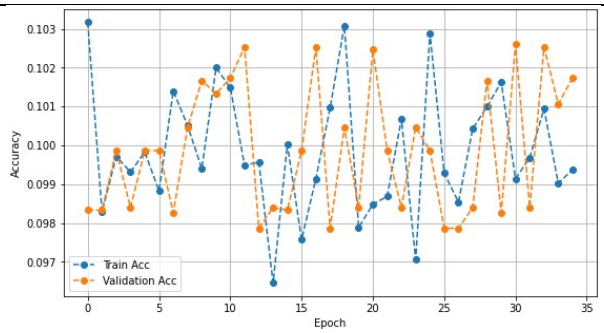
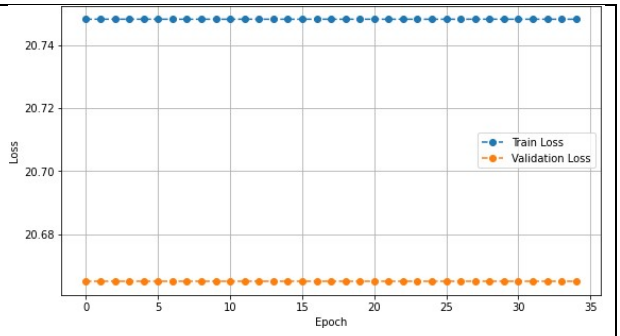
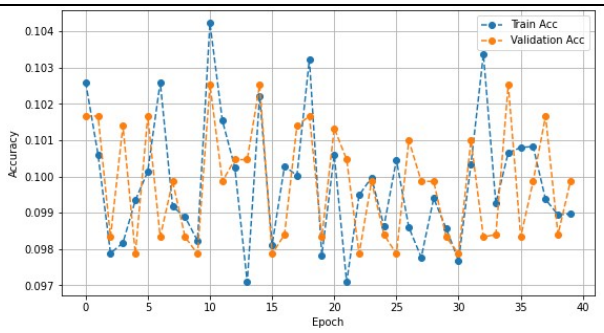
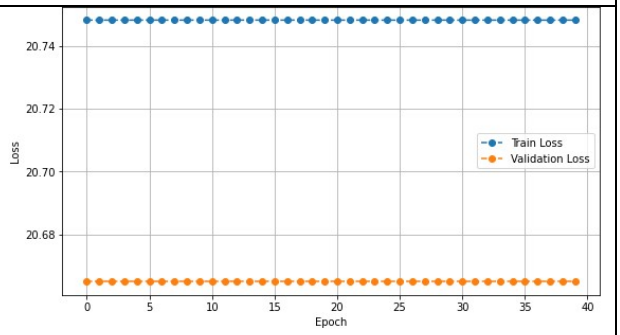
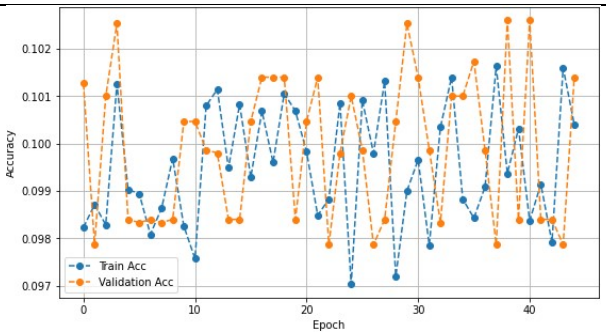
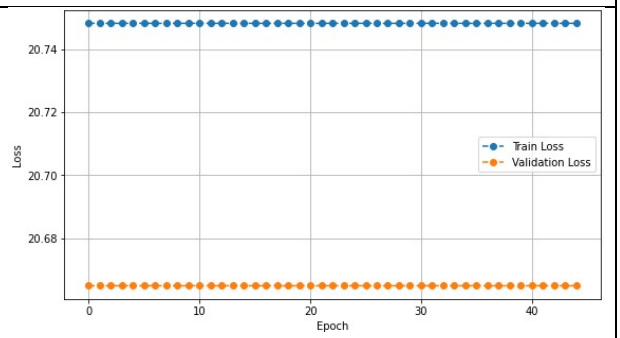
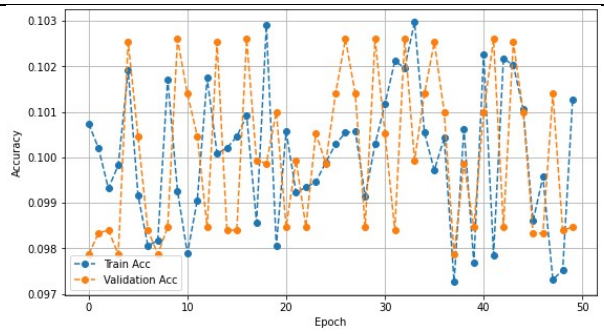
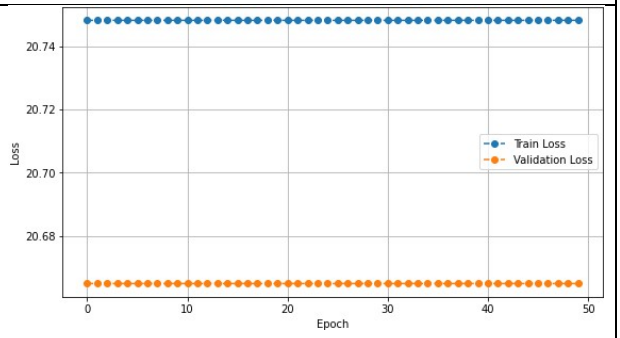
Epoch 25	 Training accuracy: 0.9784 Validation accuracy: 0.6741	 Training loss: 0.0598 Validation loss: 2.9373
Epoch 30	 Training accuracy: 0.9871 Validation accuracy: 0.6711	 Training loss: 0.0392 Validation loss: 3.3322
Epoch 35	 Training accuracy: 0.9890 Validation accuracy: 0.6773	 Training loss: 0.0336 Validation loss: 3.4659
Epoch 40	 Training accuracy: 0.9910	 Training loss: 0.0280

	Validation accuracy: 0.6773	Validation loss: 3.5282
Epoch 45	 Training accuracy: 0.9883 Validation accuracy: 0.6744	 Training loss: 0.0372 Validation loss: 3.6974
Epoch 50	 Training accuracy: 0.9920 Validation accuracy: 0.6767	 Training loss: 0.0231 Validation loss: 4.0230

3. Loss function = Kullback leibler divergence

Epoch 10	 Training accuracy: 0.1025 Validation accuracy: 0.1043	 Training loss: 20.7483 Validation loss: 20.6649

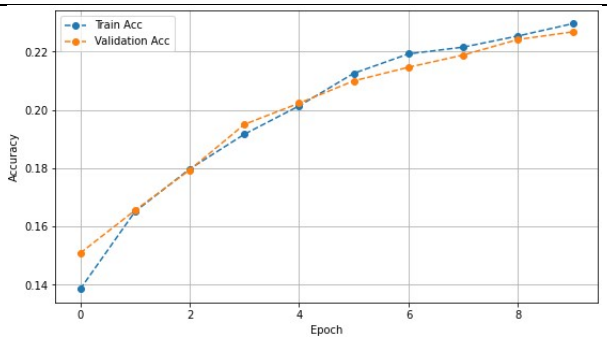
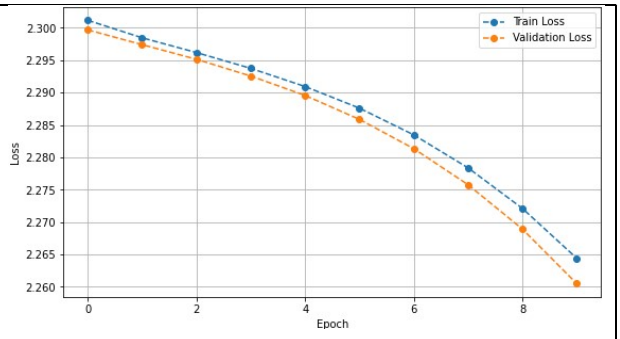
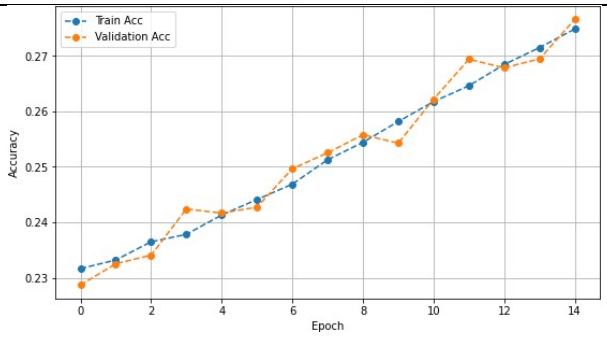
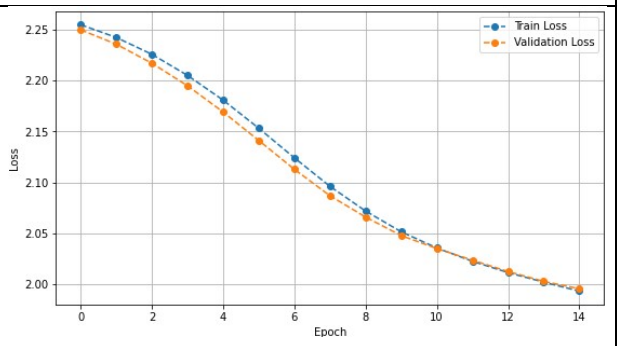
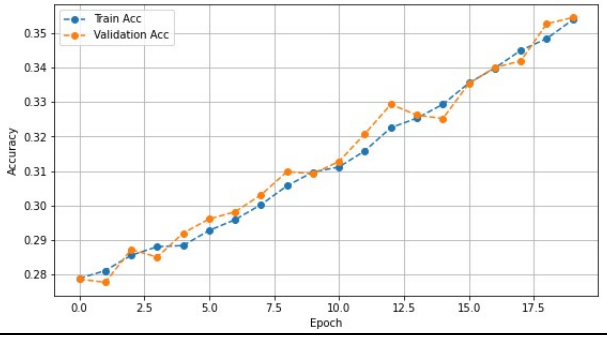
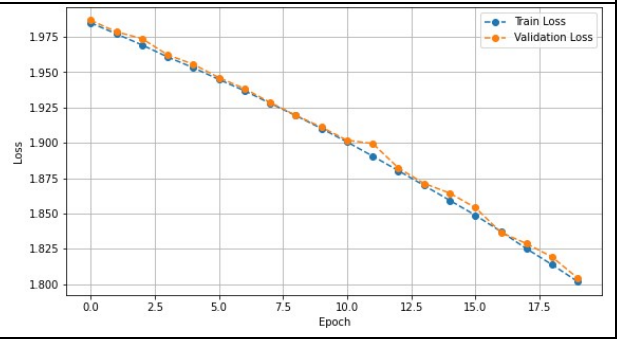
Epoch 15	 Training accuracy: 0.0966 Validation accuracy: 0.1022	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 20	 Training accuracy: 0.0976 Validation accuracy: 0.0988	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 25	 Training accuracy: 0.0998 Validation accuracy: 0.1005	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 30	 Training accuracy: 0.0990 Validation accuracy: 0.1013	 Training loss: 20.7483 Validation loss: 20.6649

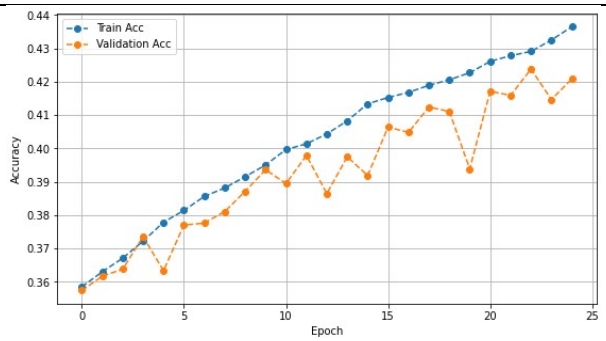
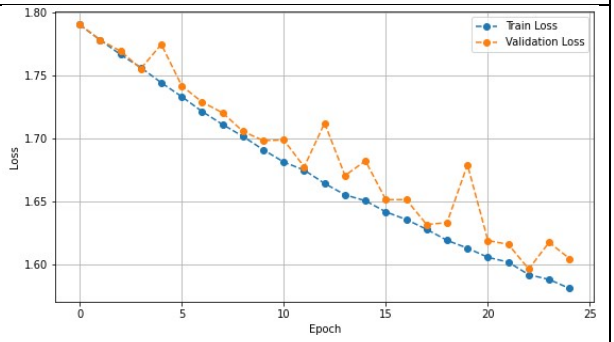
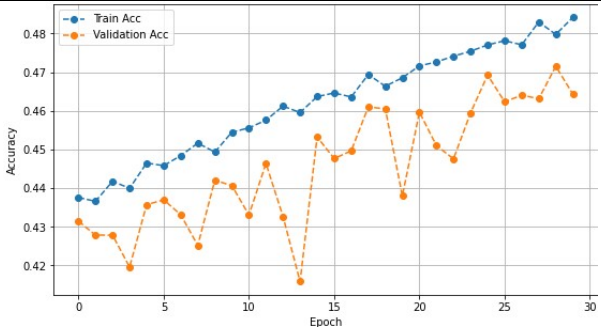
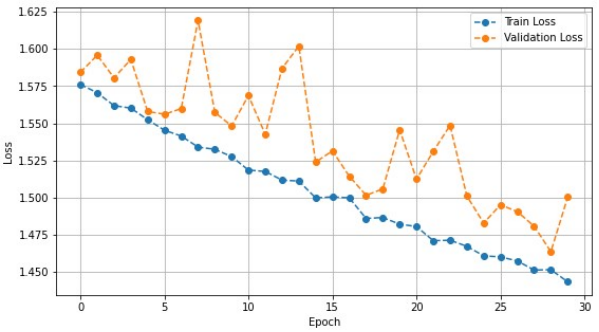
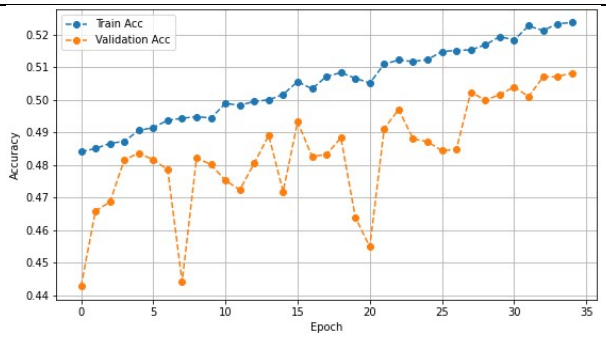
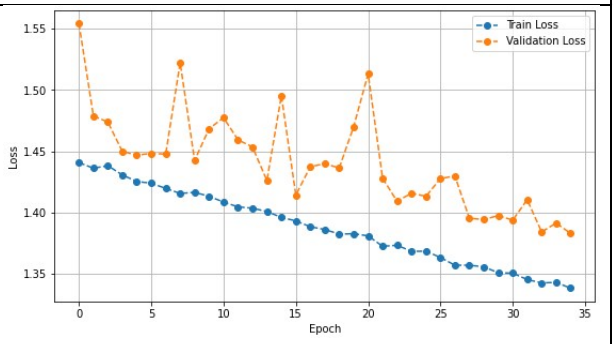
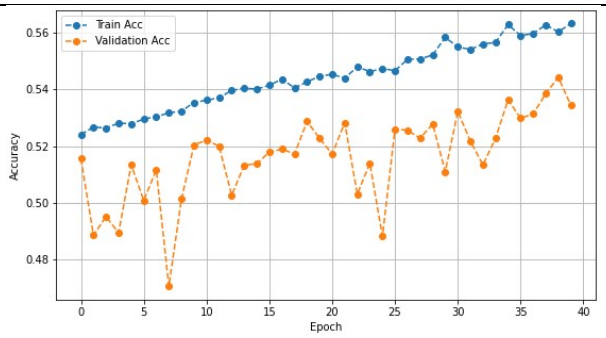
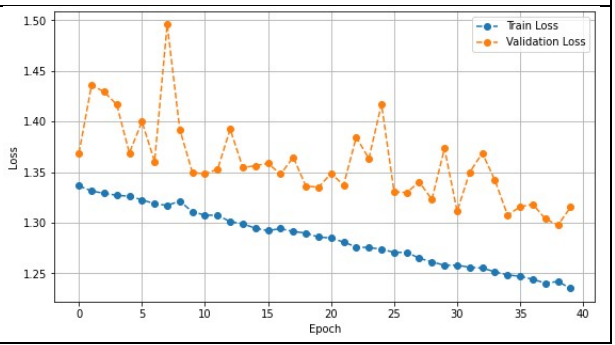
Epoch 35	 Training accuracy: 0.0994 Validation accuracy: 0.1017	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 40	 Training accuracy: 0.0990 Validation accuracy: 0.0999	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 45	 Training accuracy: 0.1004 Validation accuracy: 0.1014	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 50	 Training accuracy: 0.1013 Validation accuracy: 0.0985	 Training loss: 20.7483 Validation loss: 20.6649

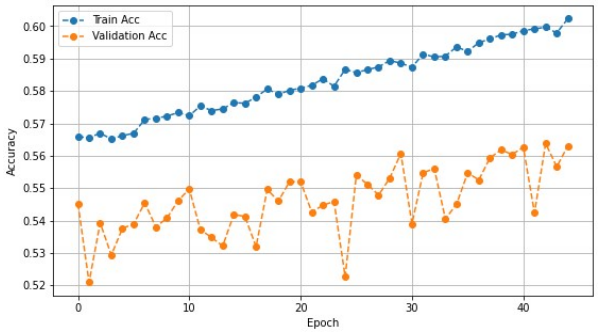
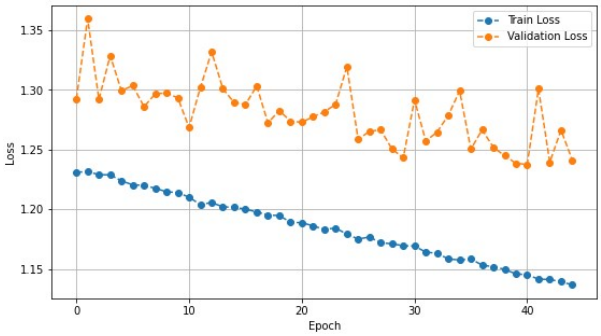
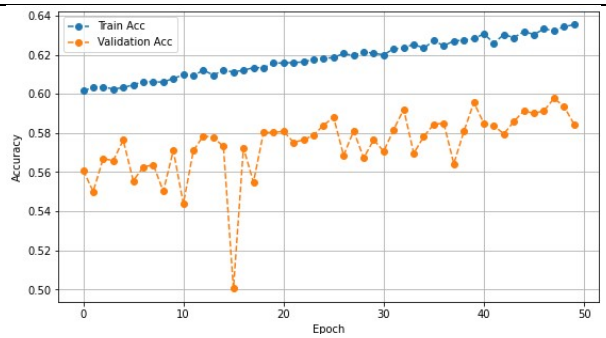
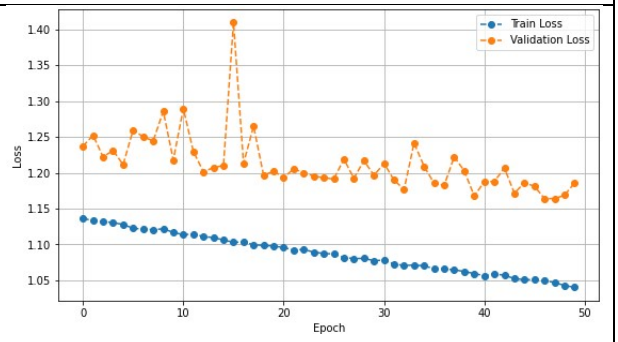
Experiment4-6: SGD

Optimizer	SGD
Batch size	128
Learning rate	0.001
Validation split	0.3
Epoch	10-50

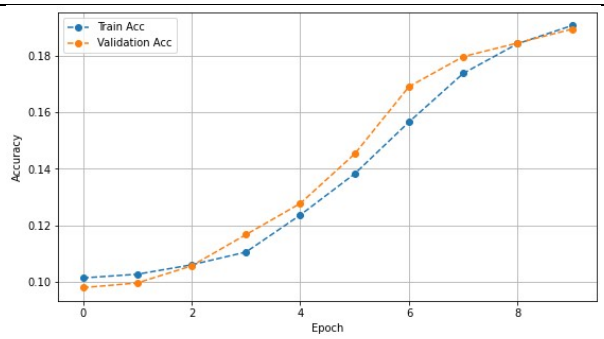
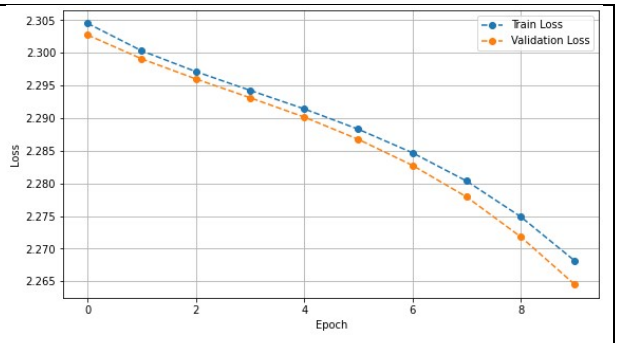
4. Loss function = Categorical crossentropy

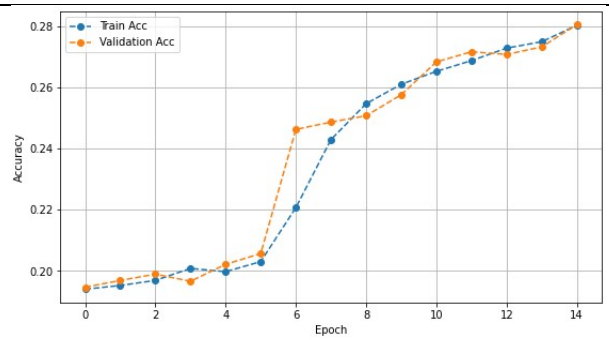
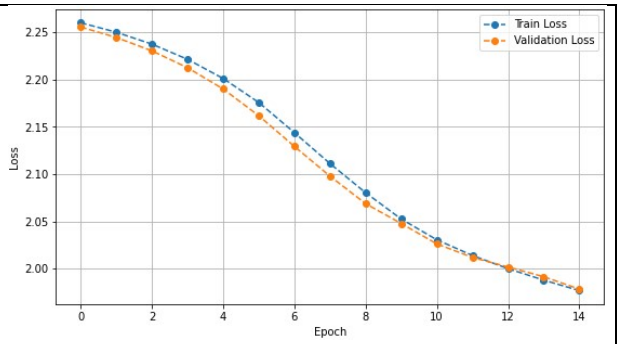
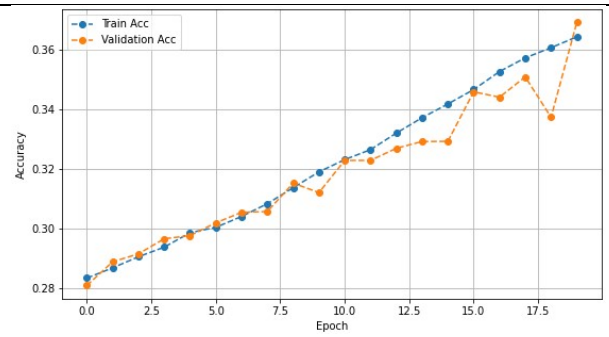
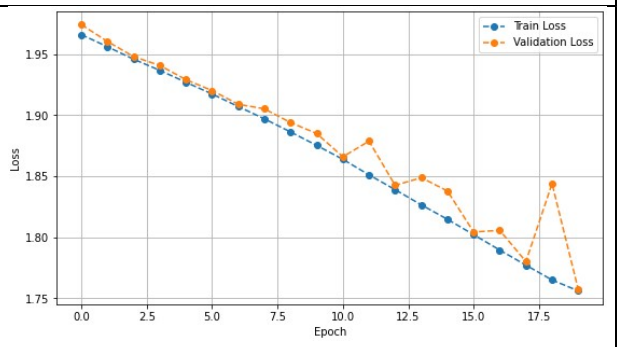
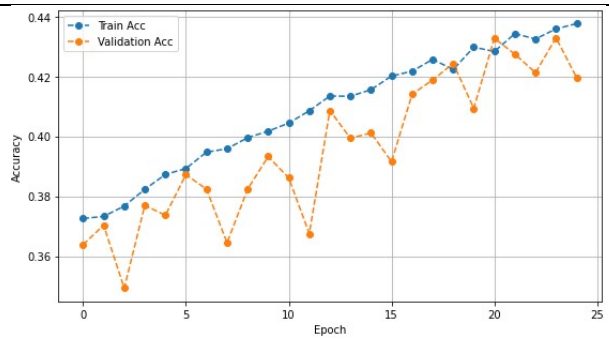
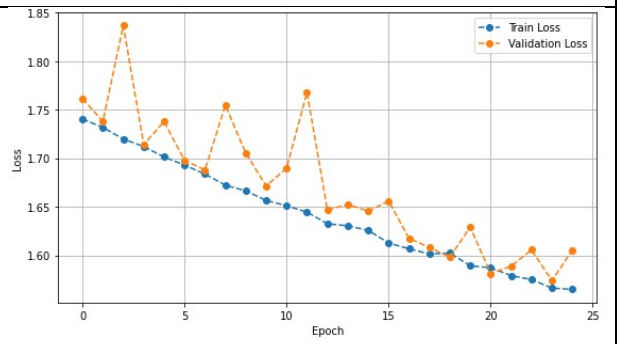
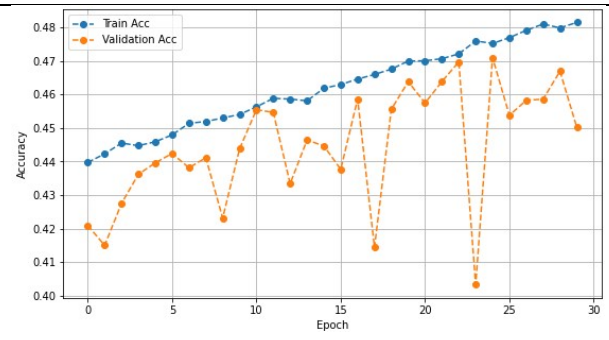
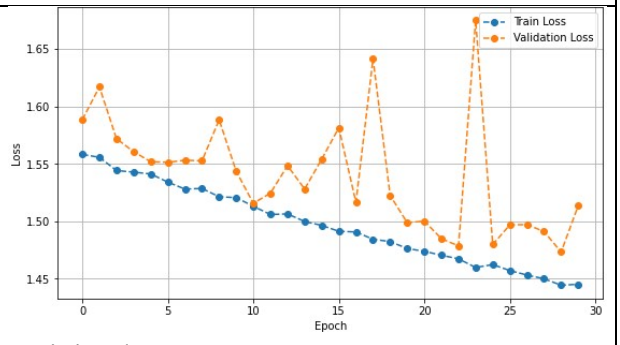
Epoch 10	 Training accuracy: 0.2296 Validation accuracy: 0.2268	 Training loss: 2.2644 Validation loss: 2.2605
Epoch 15	 Training accuracy: 0.2749 Validation accuracy: 0.2766	 Training loss: 1.9934 Validation loss: 1.9957
Epoch 20		

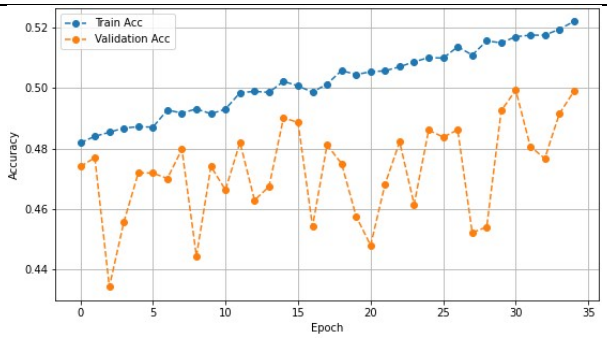
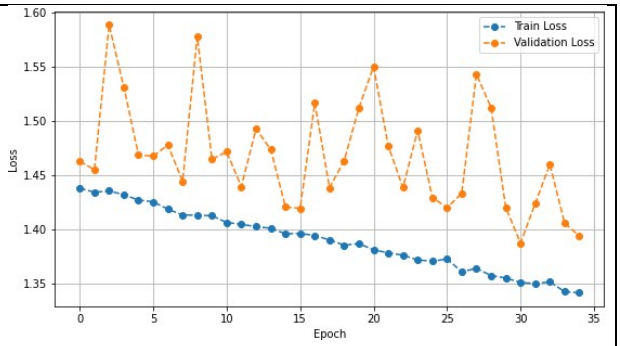
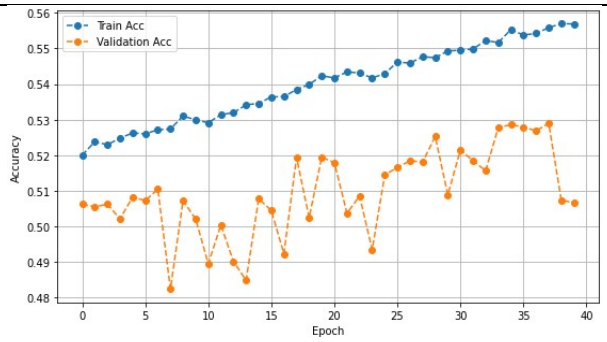
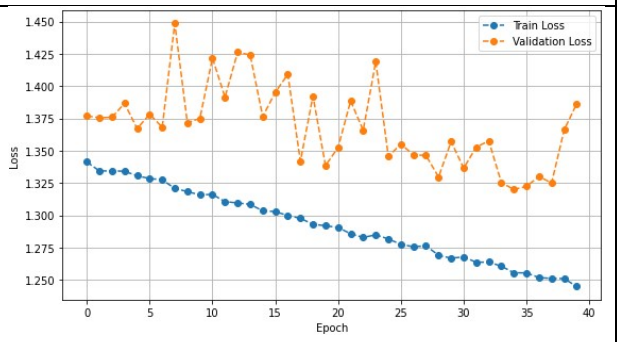
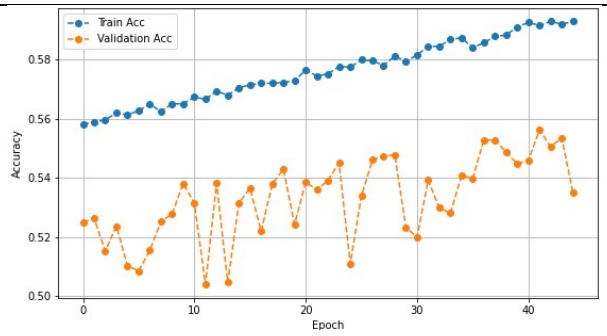
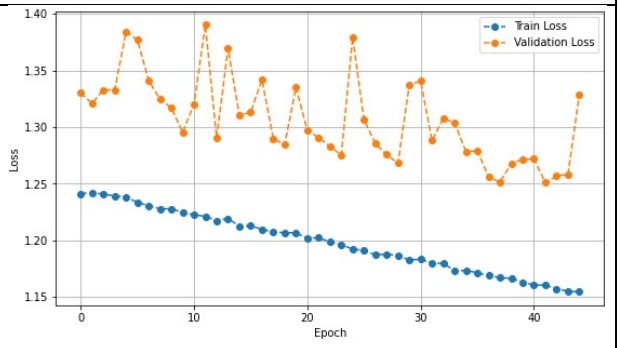
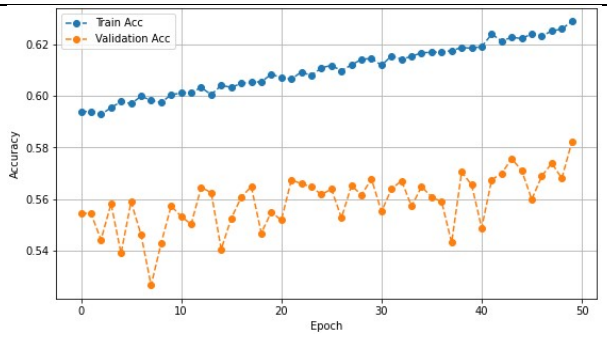
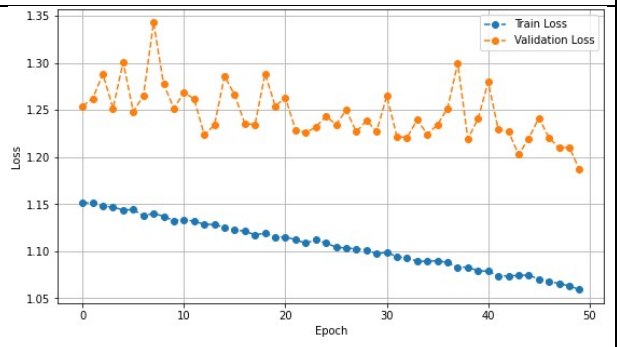
	Training accuracy: 0.3539 Validation accuracy: 0.3545	Training loss: 1.8016 Validation loss: 1.8041
Epoch 25	 Training accuracy: 0.4365 Validation accuracy: 0.4208	 Training loss: 1.5807 Validation loss: 1.6044
Epoch 30	 Training accuracy: 0.4842 Validation accuracy: 0.4643	 Training loss: 1.4435 Validation loss: 1.5003
Epoch 35	 Training accuracy: 0.5237 Validation accuracy: 0.5082	 Training loss: 1.3384 Validation loss: 1.3831
Epoch 40		

	Training accuracy: 0.5632 Validation accuracy: 0.5343	Training loss: 1.2352 Validation loss: 1.3154
Epoch 45	 Training accuracy: 0.6023 Validation accuracy: 0.5631	 Training loss: 1.1367 Validation loss: 1.2412
Epoch 50	 Training accuracy: 0.6354 Validation accuracy: 0.5845	 Training loss: 1.0402 Validation loss: 1.1860

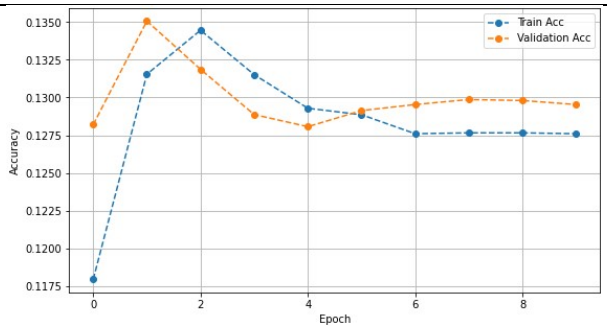
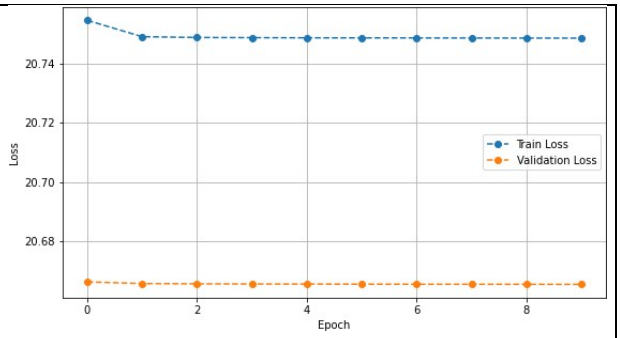
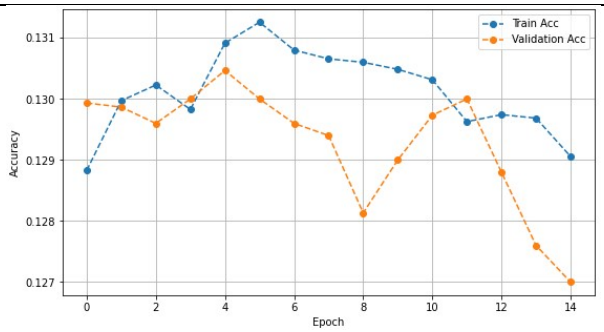
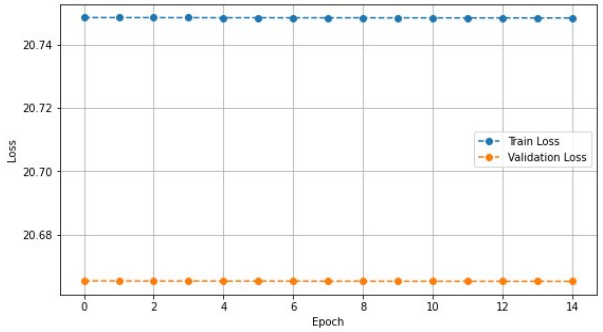
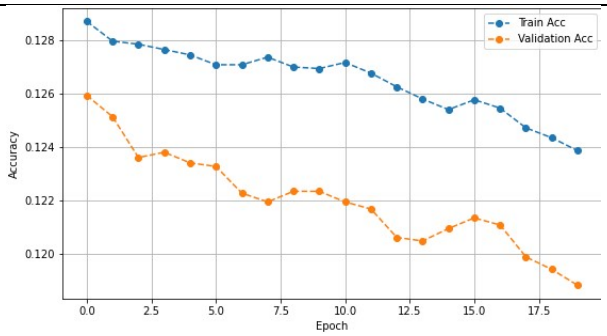
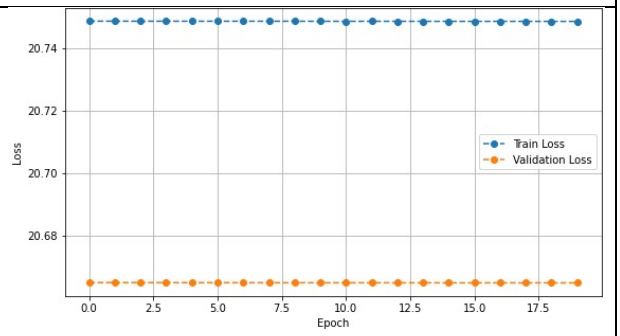
5. Loss function = Sparse categorical crossentropy

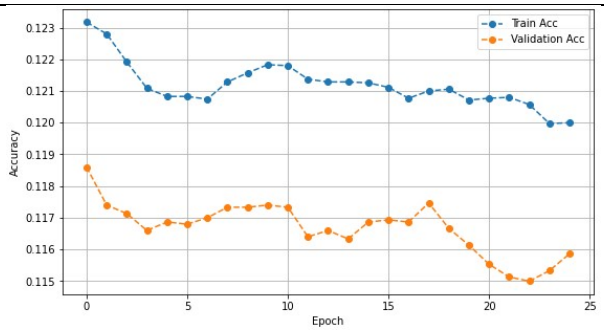
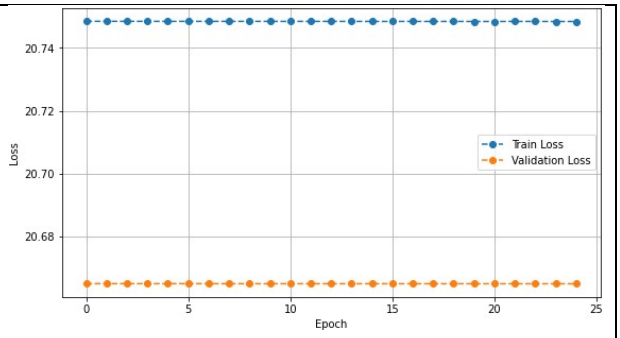
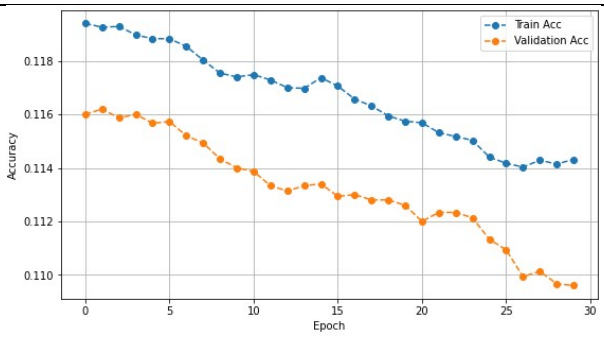
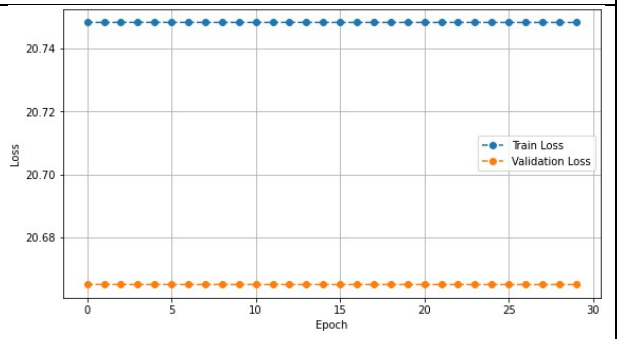
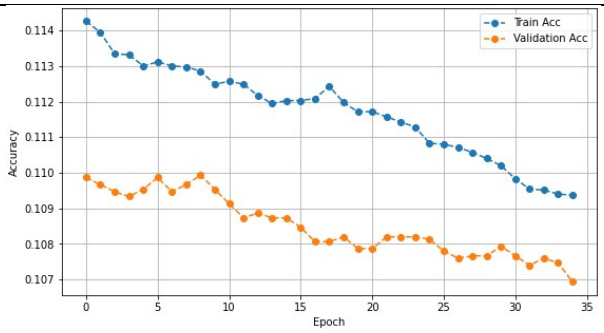
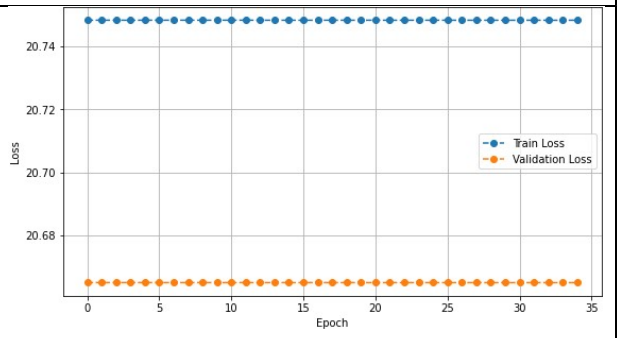
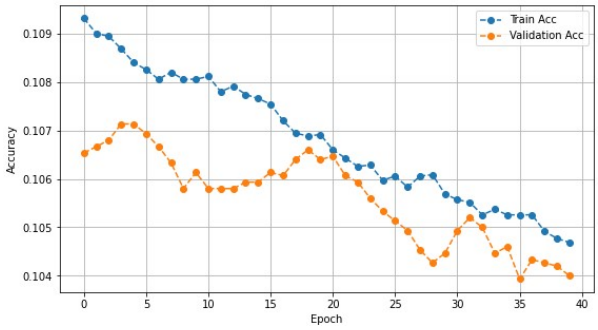
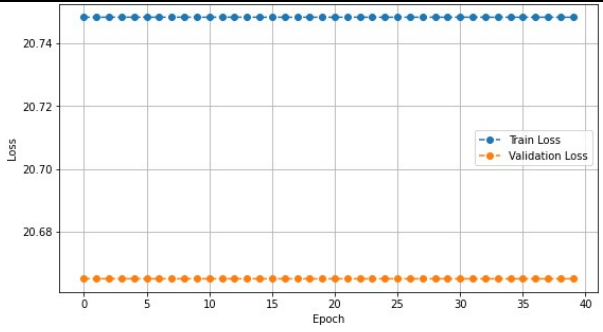
Epoch 10	 Training accuracy: 0.1908 Validation accuracy: 0.1895	 Training loss: 2.2681 Validation loss: 2.2645

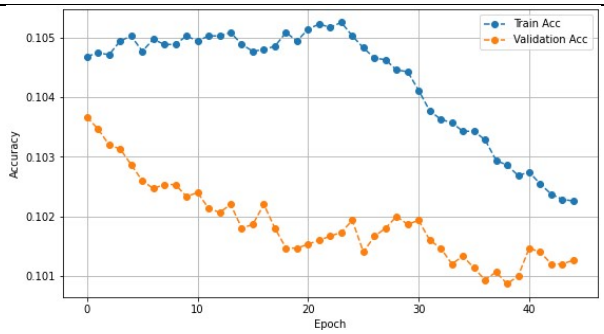
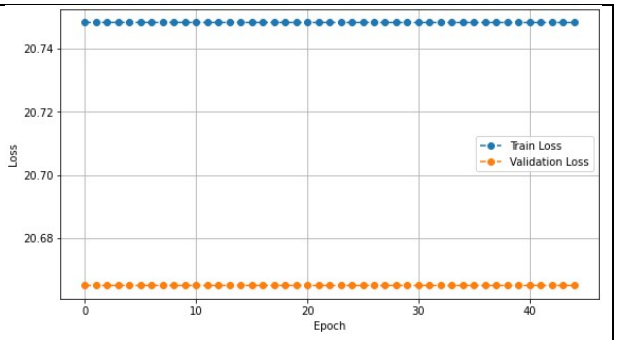
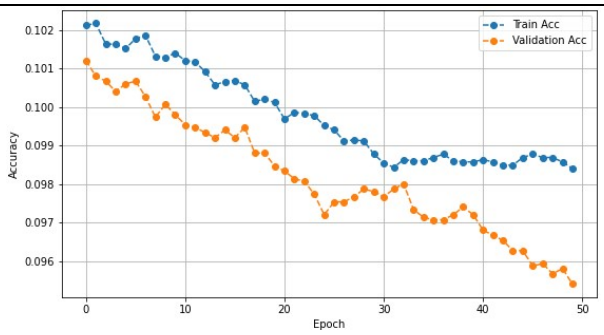
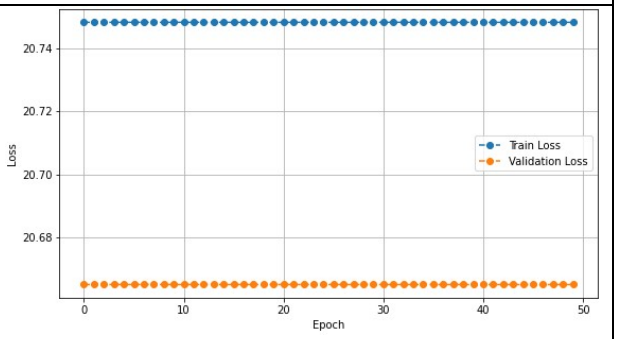
Epoch 15	 Training accuracy: 0.2803 Validation accuracy: 0.2805	 Training loss: 1.9768 Validation loss: 1.9784
Epoch 20	 Training accuracy: 0.3642 Validation accuracy: 0.3690	 Training loss: 1.7561 Validation loss: 1.7574
Epoch 25	 Training accuracy: 0.4376 Validation accuracy: 0.4196	 Training loss: 1.5649 Validation loss: 1.6053
Epoch 30	 Training accuracy: 0.4815 Validation accuracy: 0.4501	 Training loss: 1.4452 Validation loss: 1.5137

Epoch 35	 Training accuracy: 0.5220 Validation accuracy: 0.4989	 Training loss: 1.3414 Validation loss: 1.3939
Epoch 40	 Training accuracy: 0.5569 Validation accuracy: 0.5067	 Training loss: 1.2449 Validation loss: 1.3864
Epoch 45	 Training accuracy: 0.5929 Validation accuracy: 0.5350	 Training loss: 1.1546 Validation loss: 1.3282
Epoch 50	 Training accuracy: 0.6289 Validation accuracy: 0.5822	 Training loss: 1.0592 Validation loss: 1.1867

6. Loss function = Kullback leibler divergence

Epoch 10	 Training accuracy: 0.1276 Validation accuracy: 0.1295	 Training loss: 20.7487 Validation loss: 20.6653
Epoch 15	 Training accuracy: 0.1291 Validation accuracy: 0.1270	 Training loss: 20.7485 Validation loss: 20.6652
Epoch 20	 Training accuracy: 0.1239 Validation accuracy: 0.1188	 Training loss: 20.7484 Validation loss: 20.6651

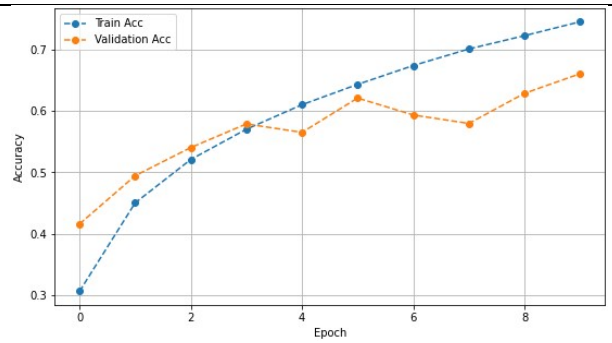
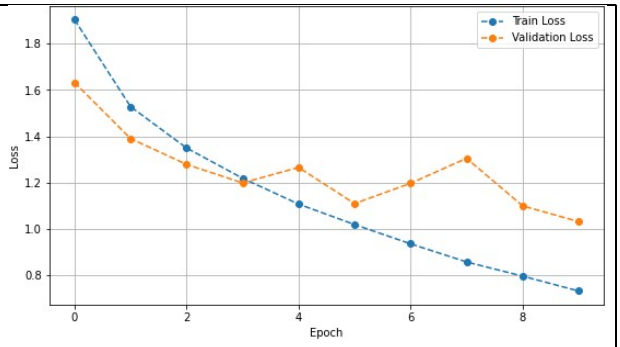
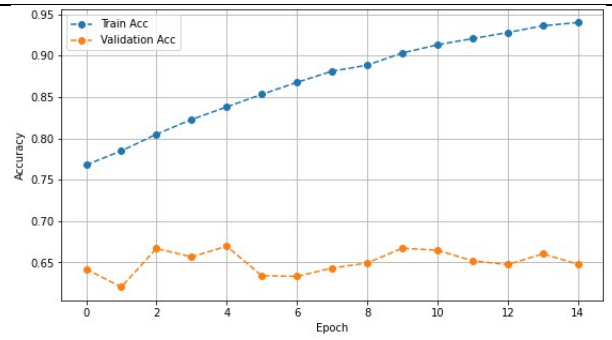
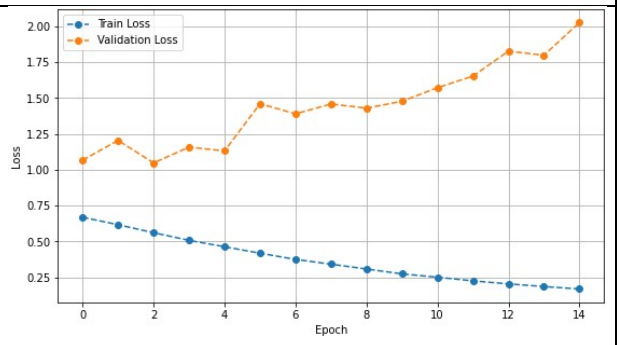
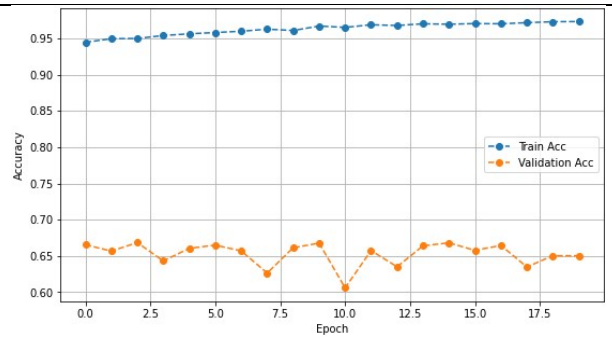
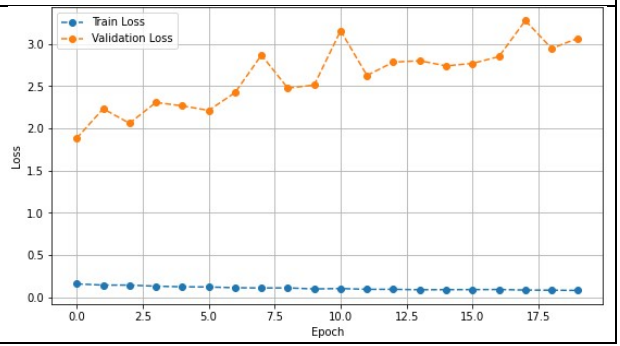
Epoch 25	 Training accuracy: 0.1200 Validation accuracy: 0.1159	 Training loss: 20.7484 Validation loss: 20.6650
Epoch 30	 Training accuracy: 0.1143 Validation accuracy: 0.1096	 Training loss: 20.7483 Validation loss: 20.6650
Epoch 35	 Training accuracy: 0.1094 Validation accuracy: 0.1069	 Training loss: 20.7483 Validation loss: 20.6650
Epoch 40	 Training accuracy: 0.1047 Validation accuracy: 0.1040	 Training loss: 20.7483 Validation loss: 20.6650

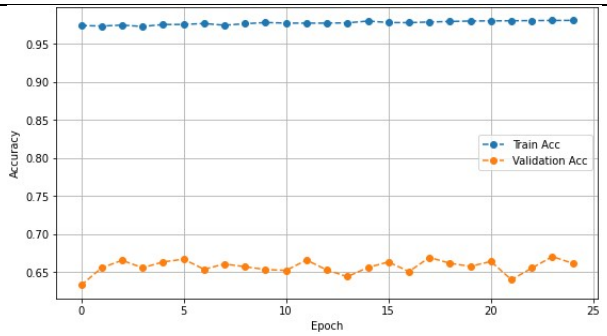
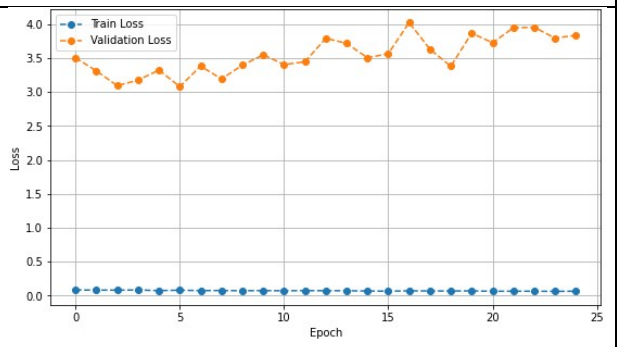
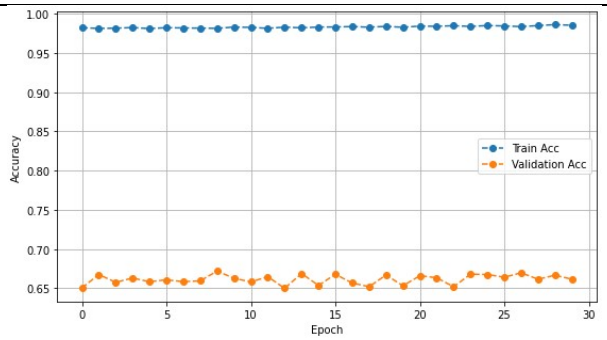
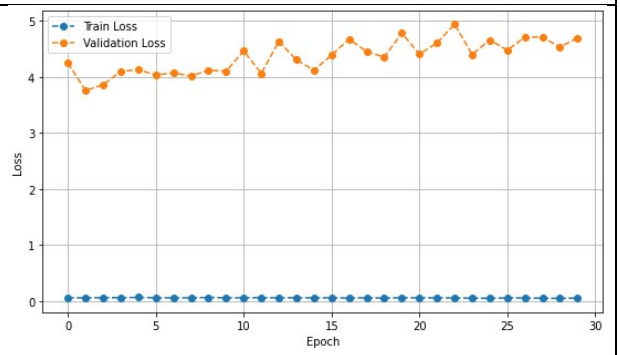
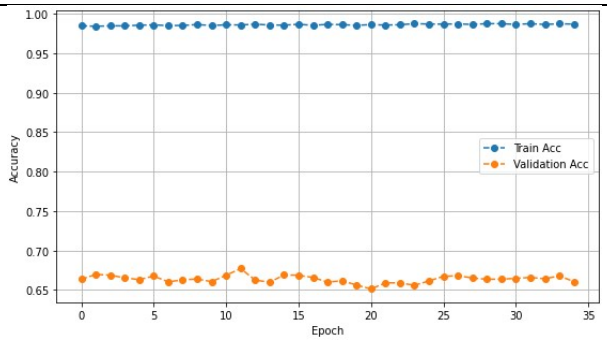
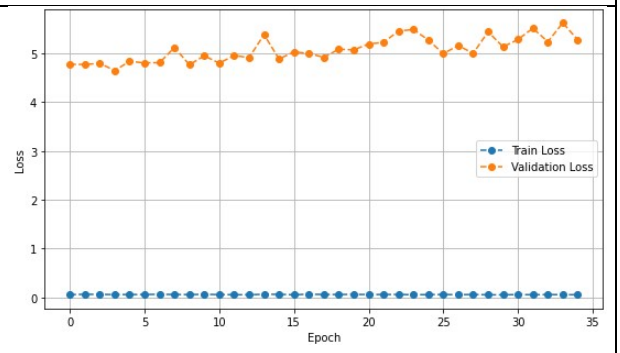
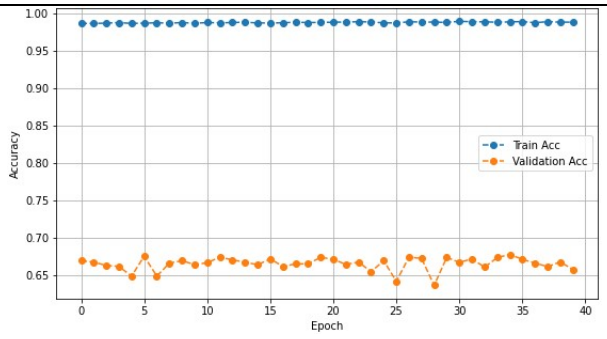
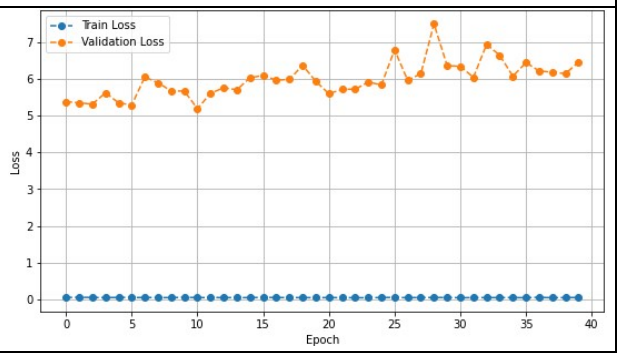
Epoch 45	 Training accuracy: 0.1023 Validation accuracy: 0.1013	 Training loss: 20.7483 Validation loss: 20.6650
Epoch 50	 Training accuracy: 0.0984 Validation accuracy: 0.0954	 Training loss: 20.7483 Validation loss: 20.6650

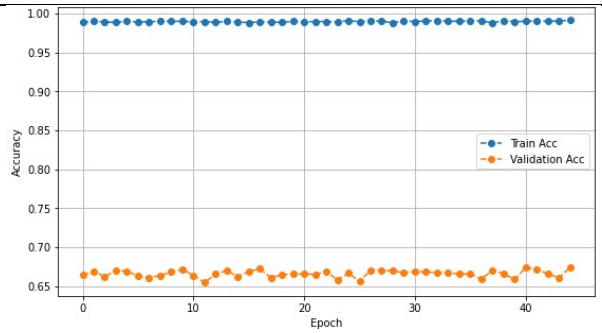
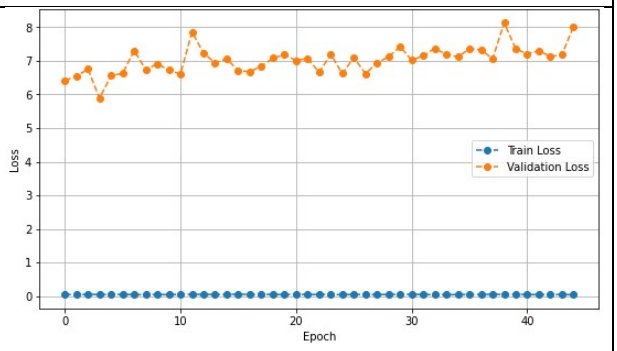
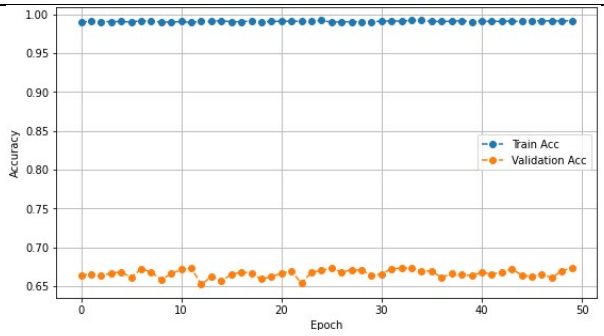
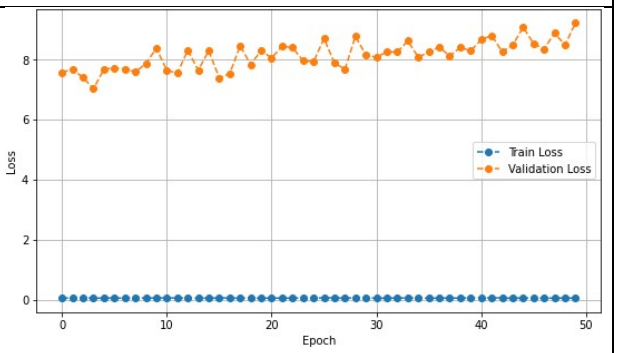
Experiment 7-9: RMSprop

Optimizer	RMSprop
Batch size	128
Learning rate	0.001
Validation split	0.3
Epoch	10-50

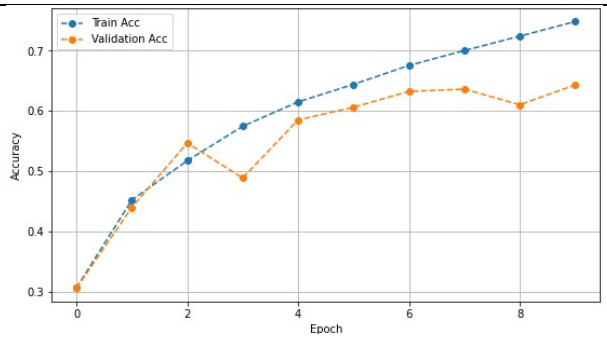
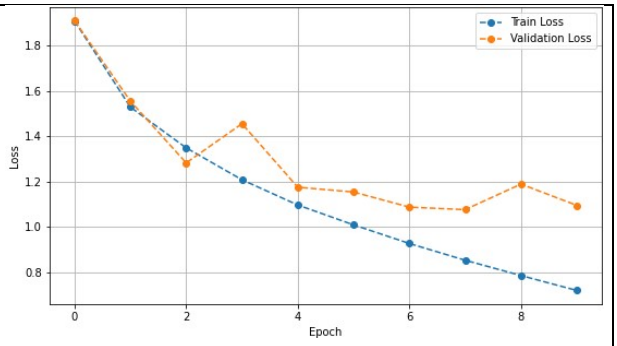
7. Loss function = Categorical crossentropy

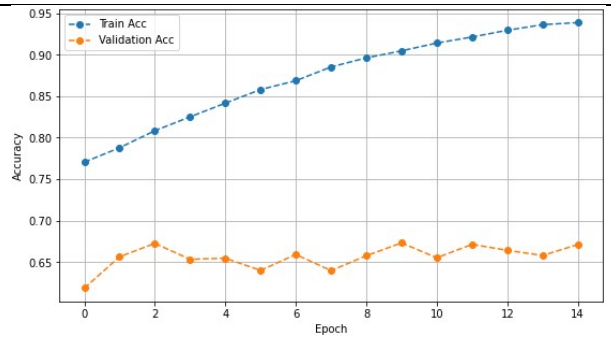
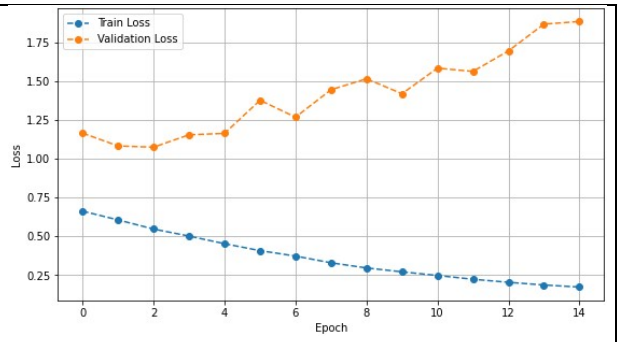
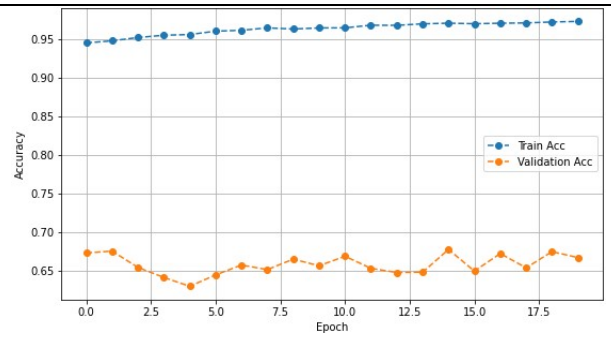
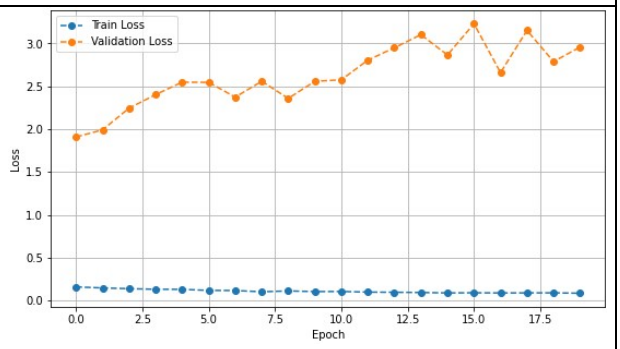
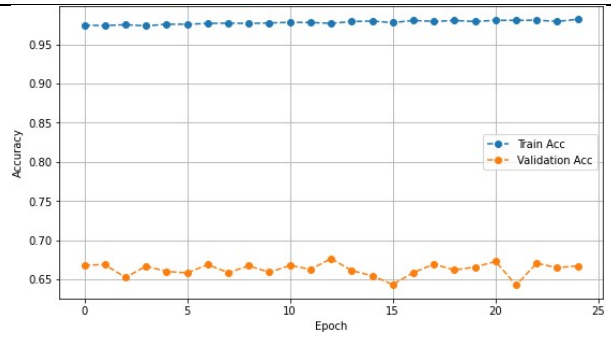
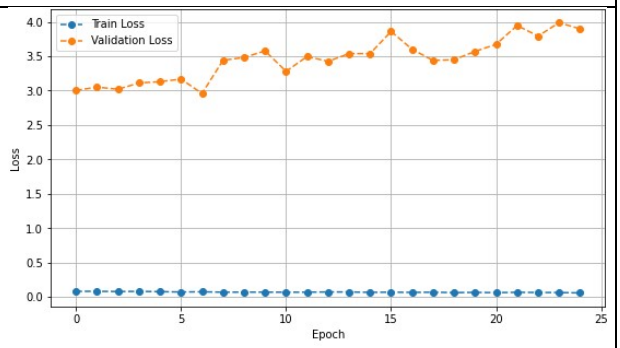
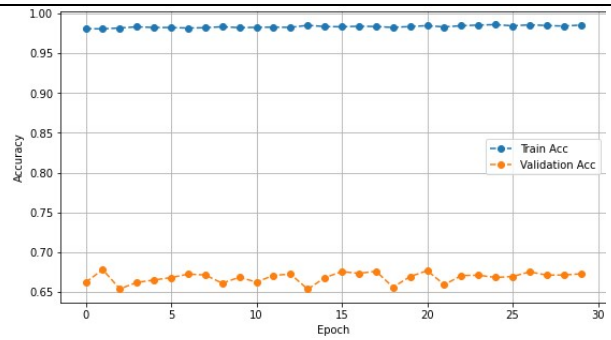
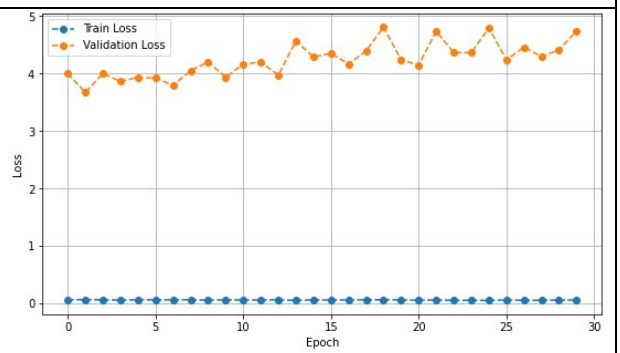
Epoch 10	 Training accuracy: 0.7455 Validation accuracy: 0.6608	 Training loss: 0.7317 Validation loss: 1.0312
Epoch 15	 Training accuracy: 0.9401 Validation accuracy: 0.6477	 Training loss: 0.1690 Validation loss: 2.0269
Epoch 20		

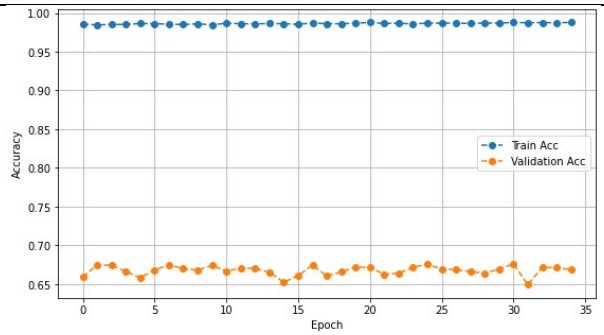
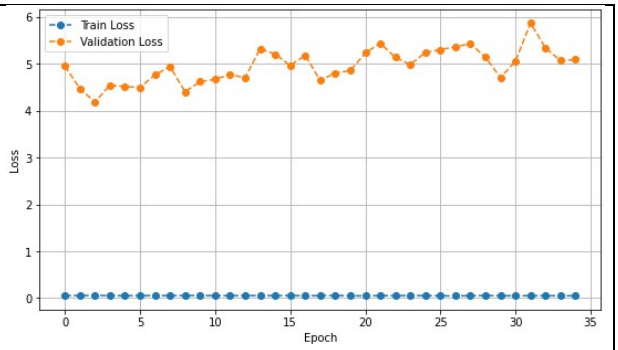
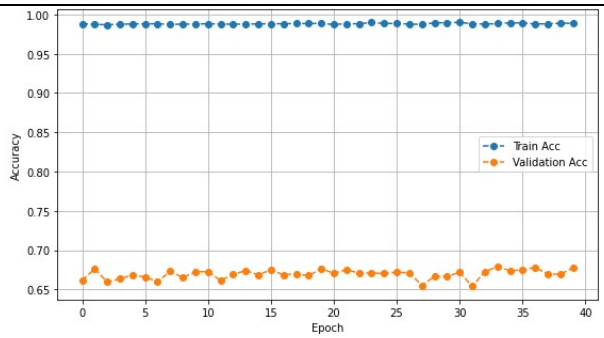
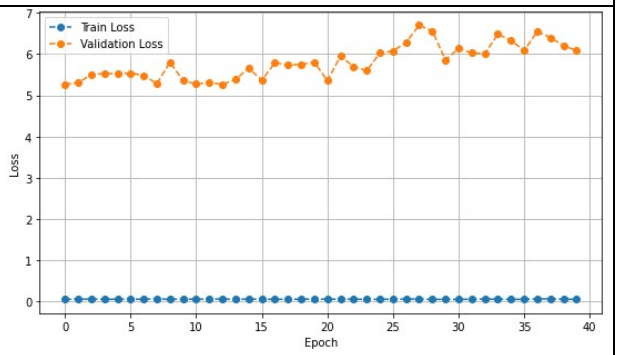
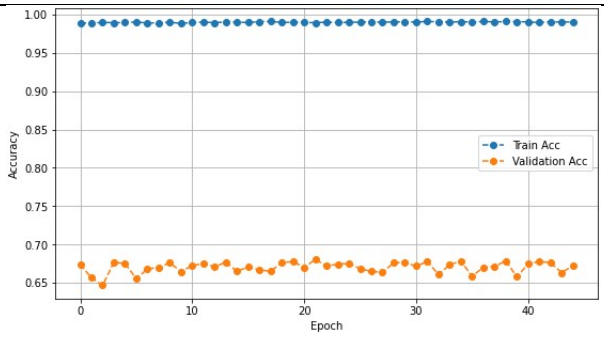
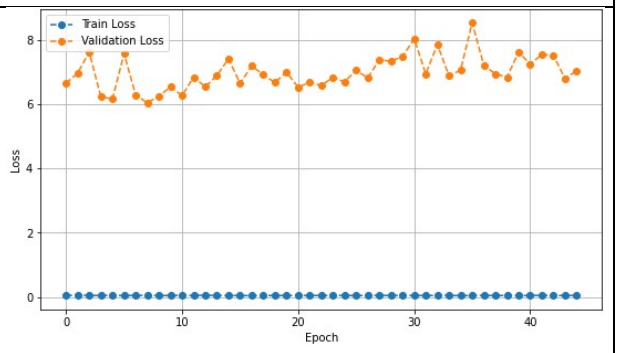
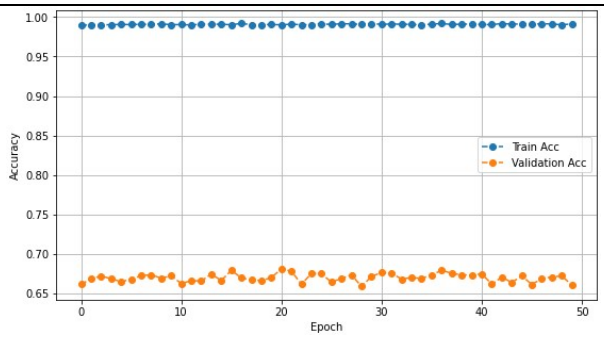
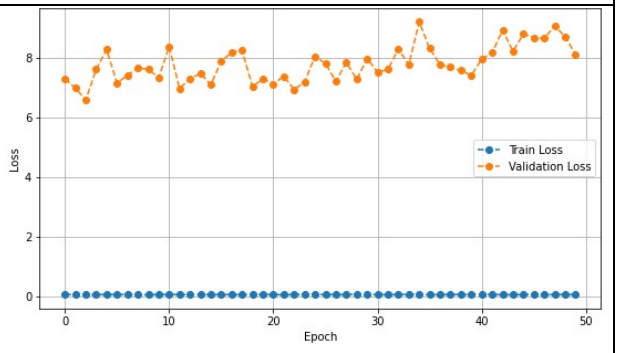
	Training accuracy: 0.9732 Validation accuracy: 0.6503	Training loss: 0.0825 Validation loss: 3.0659
Epoch 25	 Training accuracy: 0.9809 Validation accuracy: 0.6617	 Training loss: 0.0664 Validation loss: 3.8357
Epoch 30	 Training accuracy: 0.9853 Validation accuracy: 0.6615	 Training loss: 0.0592 Validation loss: 4.6994
Epoch 35	 Training accuracy: 0.9869 Validation accuracy: 0.6605	 Training loss: 0.0521 Validation loss: 5.2773
Epoch 40		

	Training accuracy: 0.9887 Validation accuracy: 0.6575	Training loss: 0.0539 Validation loss: 6.4458
Epoch 45	 Training accuracy: 0.9914 Validation accuracy: 0.6739	 Training loss: 0.0419 Validation loss: 8.0274
Epoch 50	 Training accuracy: 0.9917 Validation accuracy: 0.6738	 Training loss: 0.0473 Validation loss: 9.2313

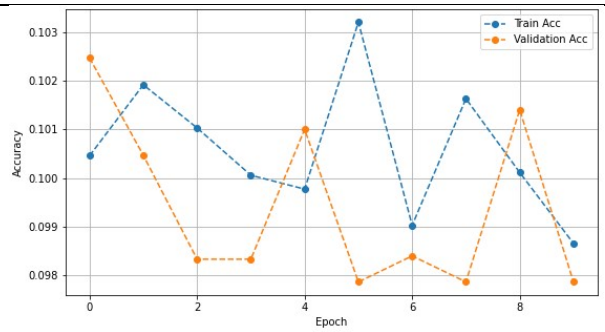
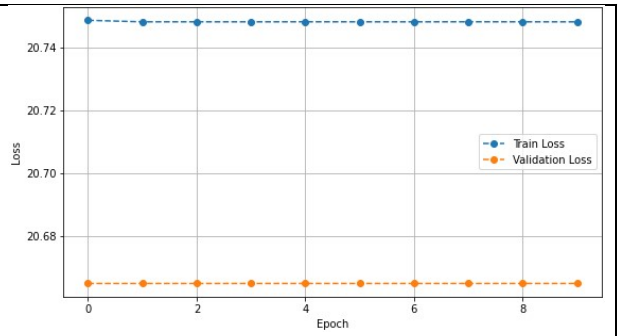
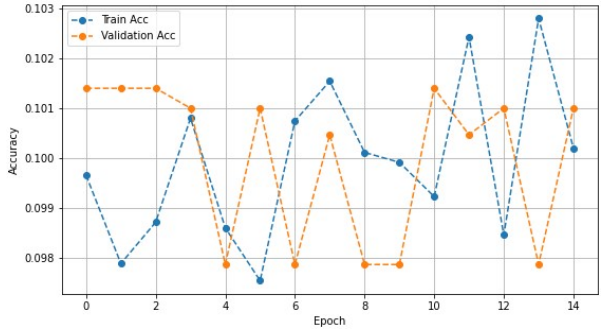
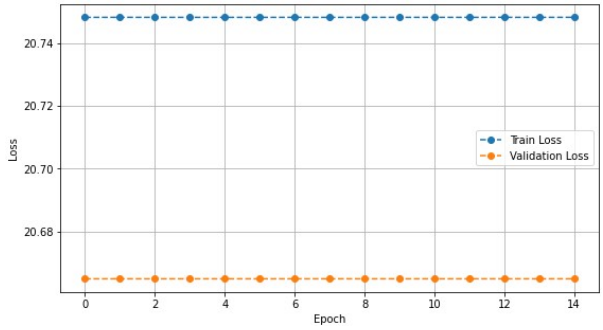
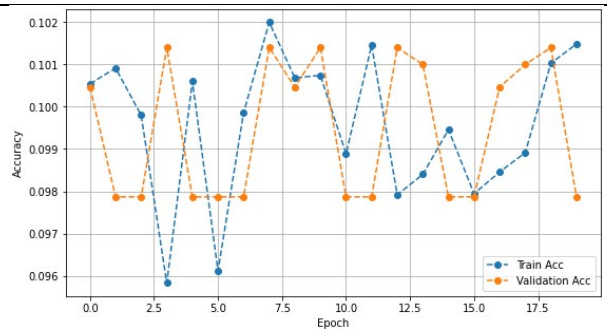
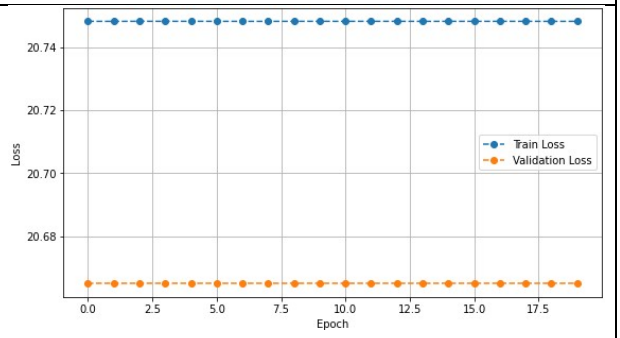
8. Loss function = Sparse categorical crossentropy

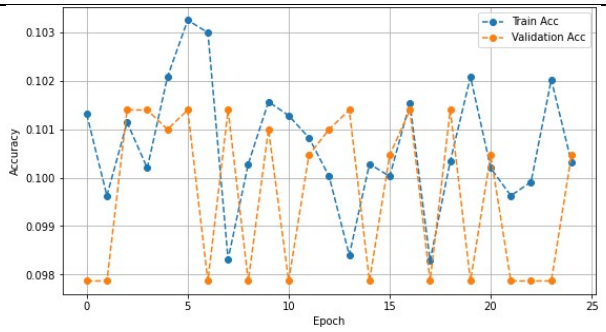
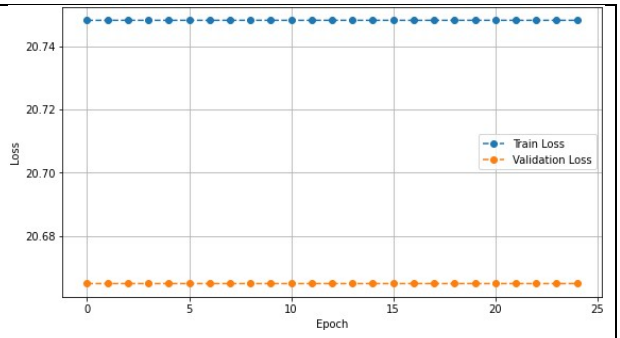
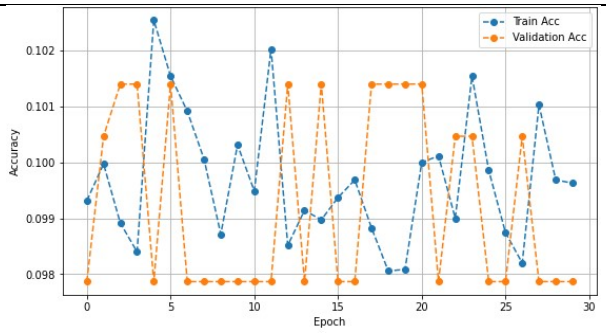
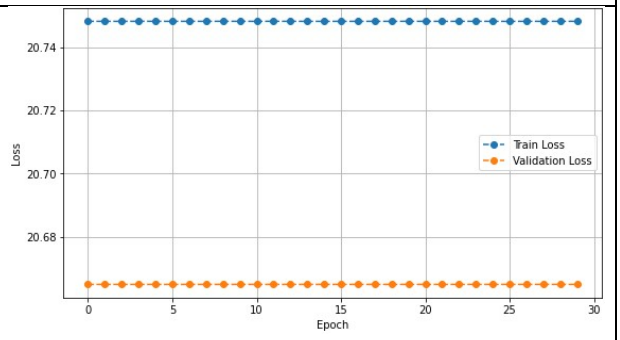
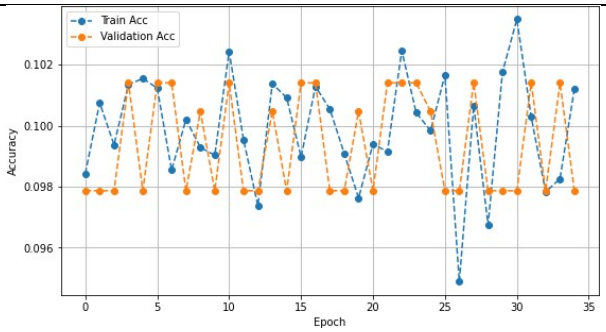
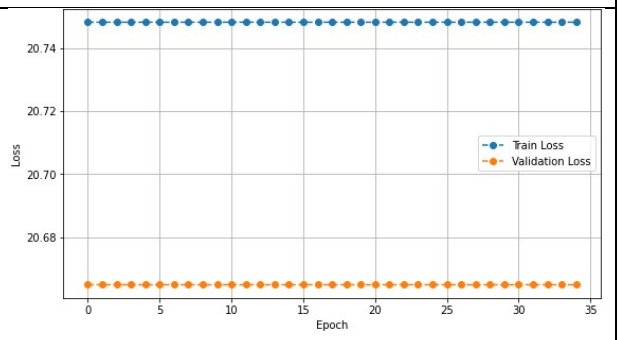
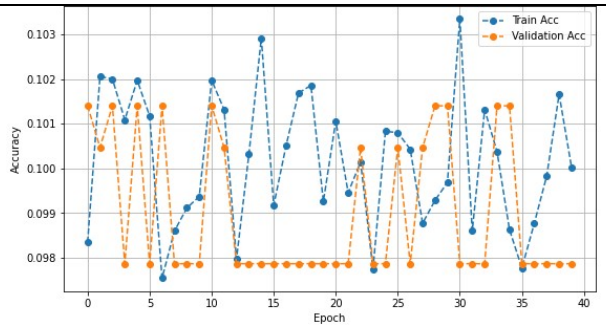
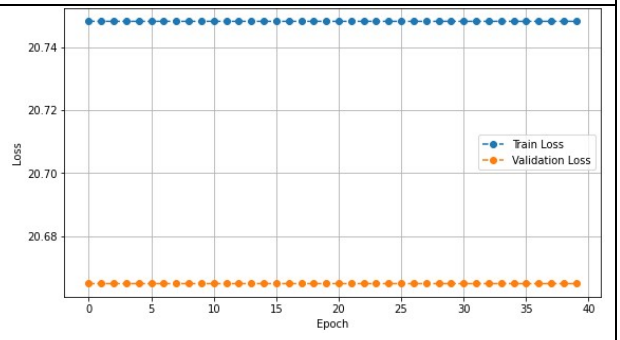
Epoch 10	 Training accuracy: 0.7485 Validation accuracy: 0.6433	 Training loss: 0.7195 Validation loss: 1.0948

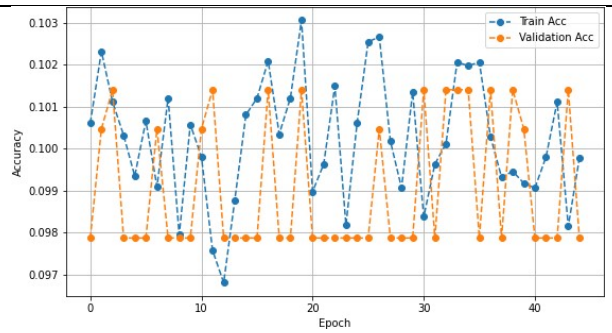
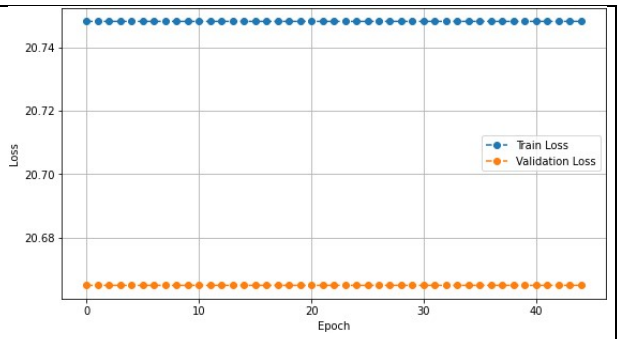
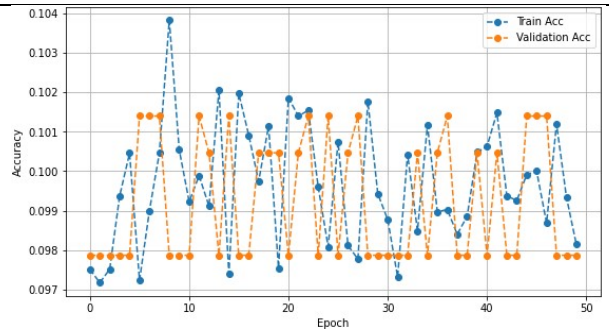
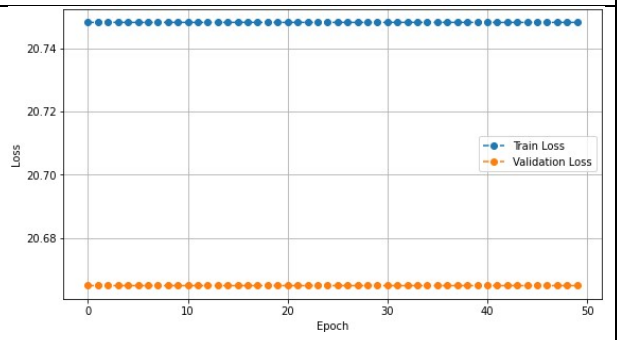
Epoch 15	 Training accuracy: 0.9388 Validation accuracy: 0.6710	 Training loss: 0.1712 Validation loss: 1.8850
Epoch 20	 Training accuracy: 0.9730 Validation accuracy: 0.6669	 Training loss: 0.0852 Validation loss: 2.9575
Epoch 25	 Training accuracy: 0.9819 Validation accuracy: 0.6673	 Training loss: 0.0602 Validation loss: 3.8972
Epoch 30	 Training accuracy: 0.9854 Validation accuracy: 0.6729	 Training loss: 0.0592 Validation loss: 4.7452

Epoch 35	 Training accuracy: 0.9879 Validation accuracy: 0.6687	 Training loss: 0.0503 Validation loss: 5.0892
Epoch 40	 Training accuracy: 0.9889 Validation accuracy: 0.6775	 Training loss: 0.0489 Validation loss: 6.0877
Epoch 45	 Training accuracy: 0.9902 Validation accuracy: 0.6725	 Training loss: 0.0501 Validation loss: 7.0440
Epoch 50	 Training accuracy: 0.9913 Validation accuracy: 0.6606	 Training loss: 0.0501 Validation loss: 8.1230

9. Loss function = Kullback leibler divergence

Epoch 10	 Training accuracy: 0.0987 Validation accuracy: 0.0979	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 15	 Training accuracy: 0.1002 Validation accuracy: 0.1010	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 20	 Training accuracy: 0.1015 Validation accuracy: 0.0979	 Training loss: 20.7483 Validation loss: 20.6649

Epoch 25	 Training accuracy: 0.1003 Validation accuracy: 0.1005	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 30	 Training accuracy: 0.0996 Validation accuracy: 0.0979	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 35	 Training accuracy: 0.1012 Validation accuracy: 0.0979	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 40	 Training accuracy: 0.1000 Validation accuracy: 0.0979	 Training loss: 20.7483 Validation loss: 20.6649

Epoch 45	 Training accuracy: 0.0998 Validation accuracy: 0.0979	 Training loss: 20.7483 Validation loss: 20.6649
Epoch 50	 Training accuracy: 0.0981 Validation accuracy: 0.0979	 Training loss: 20.7483 Validation loss: 20.6649

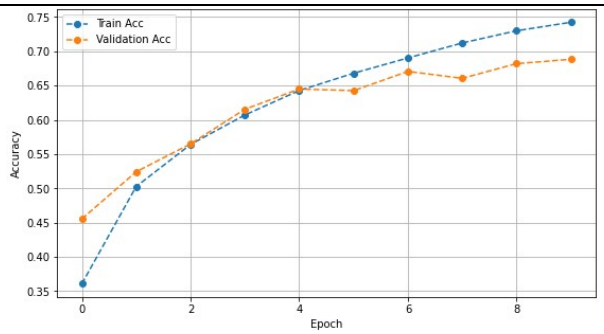
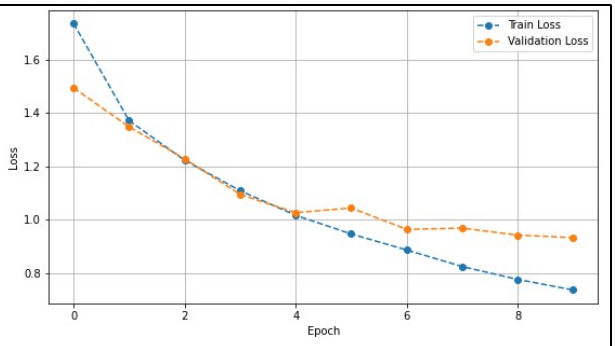
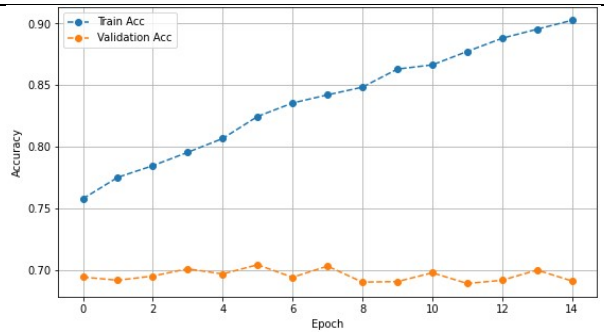
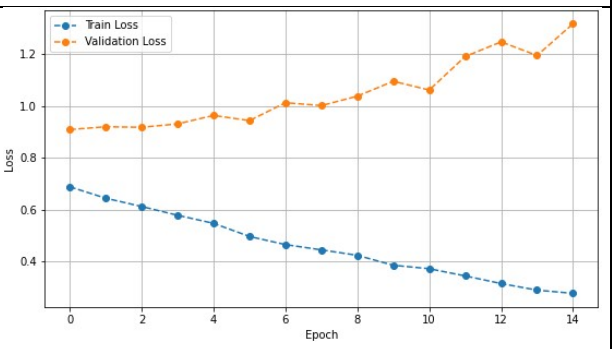
Observation: Highest accuracy 68.25% (0.6825) was achieved when,

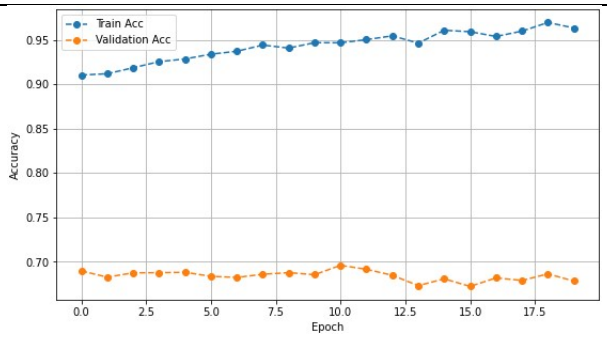
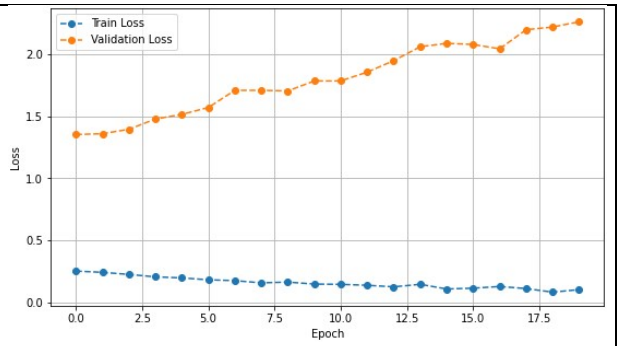
Optimizer	Adam
Loss function	Sparse categorical crossentropy
Batch size	128
Learning rate	0.001
Validation split	0.3
Epoch	10

Experiment 10: changing validation split

Keeping adam optimizer and sparse categorical crossentropy loss function constant, changing validation split from 0.3 to 0.2

Optimizer	Adam
Loss function	Sparse categorical crossentropy
Batch size	128
Learning rate	0.001
Validation split	0.2

Epoch 10	 Training accuracy: 0.7423 Validation accuracy: 0.6883	 Training loss: 0.7355 Validation loss: 0.9321
Epoch 15	 Training accuracy: 0.9026 Validation accuracy: 0.6911	 Training loss: 0.2769 Validation loss: 1.3174

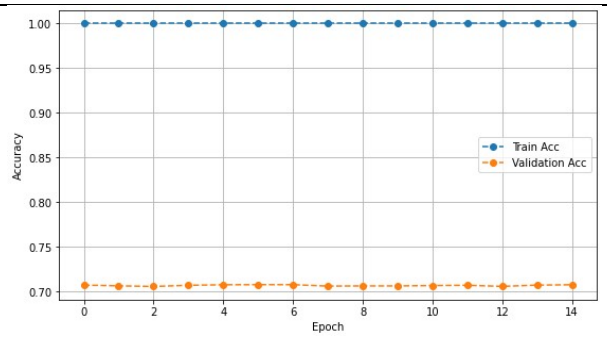
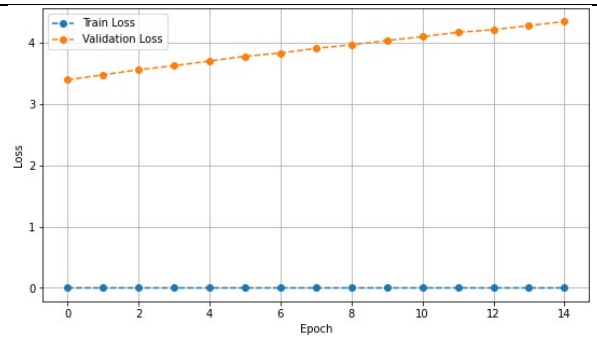
Epoch 20	 Training accuracy: 0.9633 Validation accuracy: 0.6783	 Training loss: 0.1043 Validation loss: 2.2585

Observation: Accuracy improved to 69.11% (0.6911) at epoch 15 when validation split was 0.2

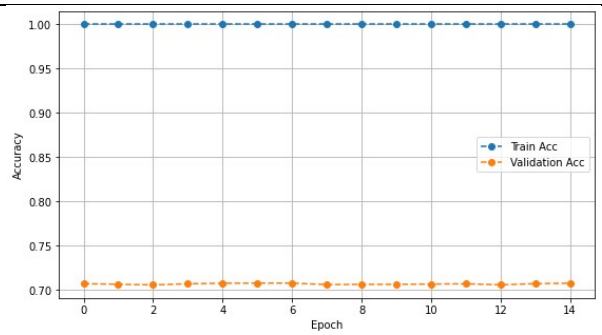
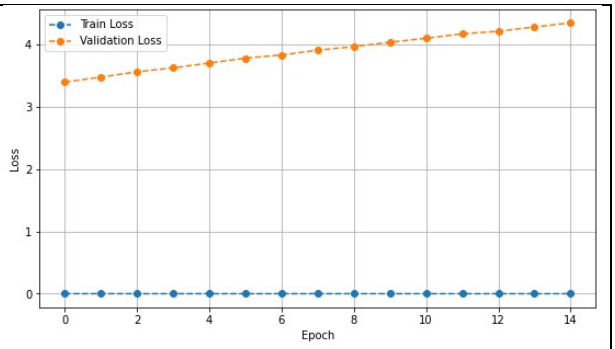
Experiment 11: changing learning rate

Optimizer	Adam
Loss function	Sparse categorical crossentropy
Batch size	128
Learning rate	0.005 - 0.0001
Validation split	0.2
Epoch	15

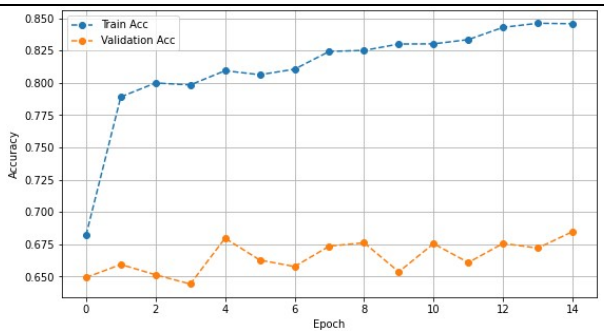
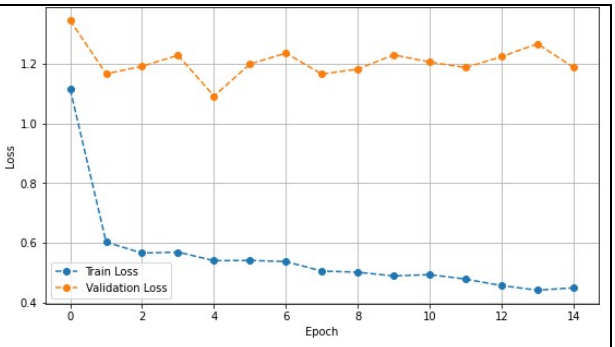
Learning rate = 0.0005

Epoch 15	 Training accuracy: 1.0000 Validation accuracy: 0.7064	 Training loss: 3.1700e-−04 Validation loss: 3.29040

Learning rate = 0.0001

Epoch 15	 Training accuracy: 1.0000 Validation accuracy: 0.7078	 Training loss: 1.5843e-05 Validation loss: 4.3468

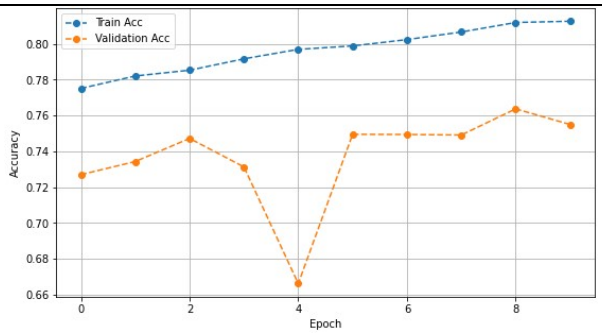
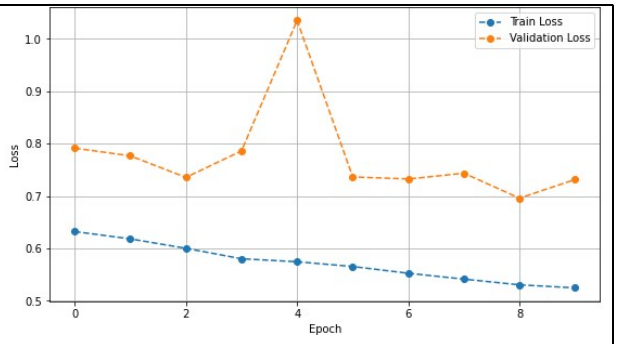
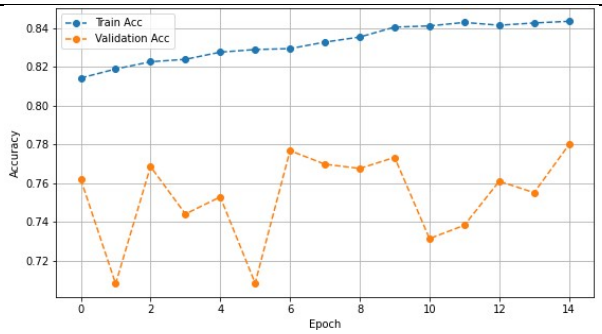
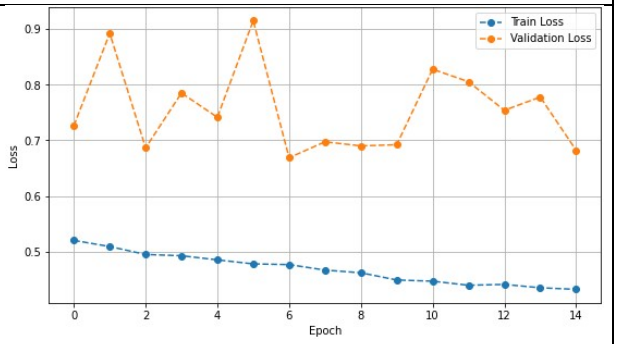
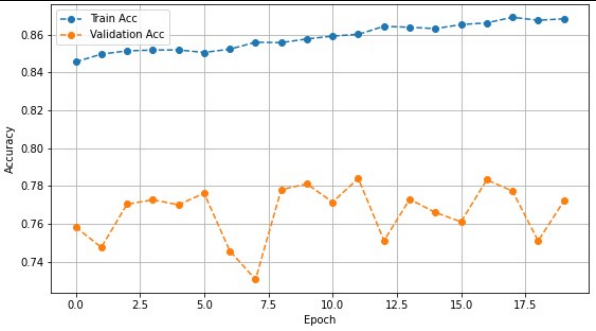
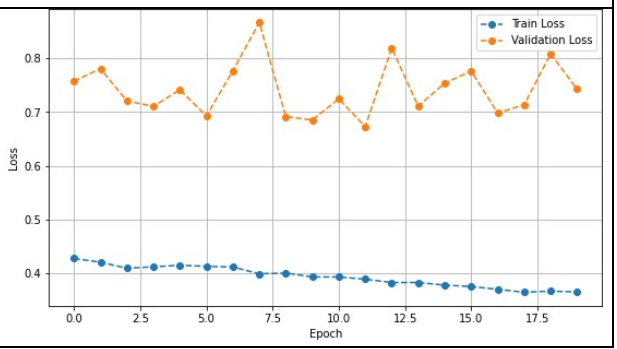
Learning rate: 0.005

Epoch 15	 Training accuracy: 0.8457 Validation accuracy: 0.6848	 Training loss: 0.4507 Validation loss: 1.1871

<u>Observation</u>: Accuracy is highest (70.78%) and (70.64%) when learning rate is 0.0001 and 0.0005, but the loss is too high (4.3468 and 3.29040). So, in the next experiment dropout will be introduced to see if we can reduce the loss and learning rate will be constant to 0.001.

Experiment 12: Adding dropout and batch normalization after each convolutional layers

Optimizer	Adam
Loss function	Sparse categorical crossentropy
Batch size	128
Learning rate	0.001
Validation split	0.2
Dropout	0.2

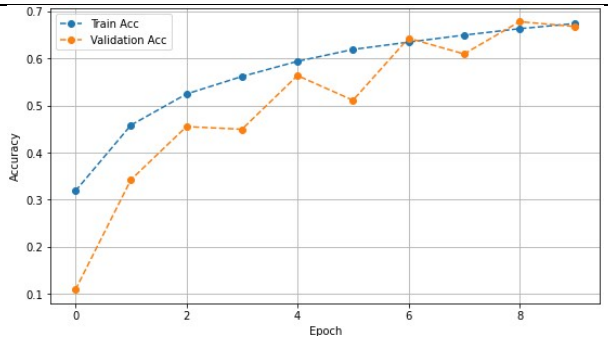
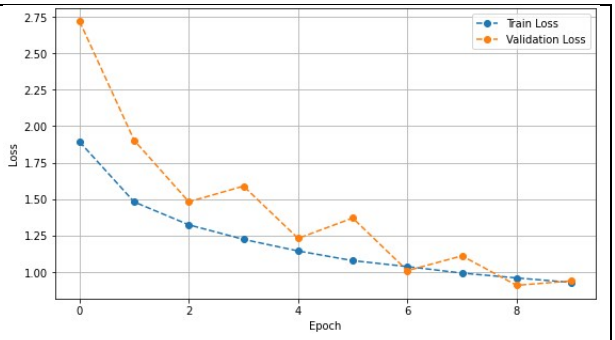
Epoch 10	 Training accuracy: 0.8126 Validation accuracy: 0.7550	 Training loss: 0.5246 Validation loss: 0.7315
Epoch 15	 Training accuracy: 0.8435 Validation accuracy: 0.7798	 Training loss: 0.4323 Validation loss: 0.6823
Epoch 20		

	Training accuracy: 0.8683 Validation accuracy: 0.7720	Training loss: 0.3656 Validation loss: 0.7427

Observation: Accuracy has been improved to 78% (0.7798) and loss has also reduced (0.6823).

Next, dropout was added to dense layers too to see if we can further improve our model.

Experiment 13: Adding dropout = 0.2 and batch normalization after each dense layer

Epoch 10	 Training accuracy: 0.6737 Validation accuracy: 0.6673	 Training loss: 0.9292 Validation loss: 0.9396

Observation: As we can see, accuracy was not improved when adding dropout to dense layers.

As of now the highest accuracy is 78% (0.7798) when,

Optimizer	Adam
Loss function	Sparse categorical crossentropy
Batch size	128
Learning rate	0.001
Validation split	0.2
Dropout	0.2 (after each convolutional layer)
Epoch	15

Next, data augmentation will be applied keeping these parameters constant.

Experiment14: Data augmentation

Augmentation Parameters: featurewise_center=False

samplewise_center = False

featurewise_std_normalization = False

samplewise_std_normalization = False

zca_whitening = False,

rotation_range = 15,

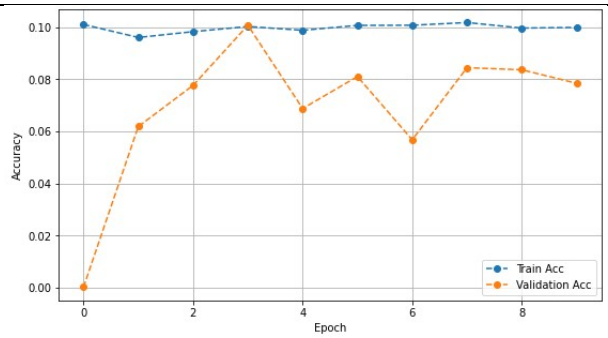
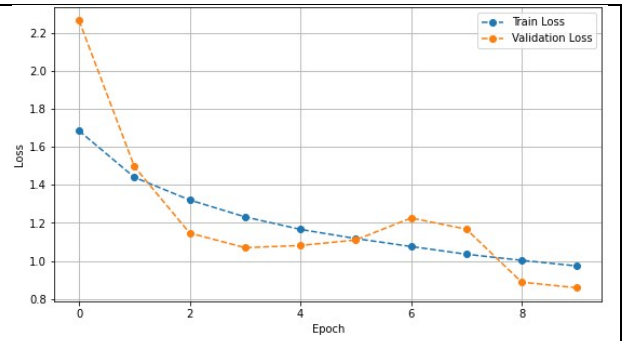
width_shift_range = 0.1,

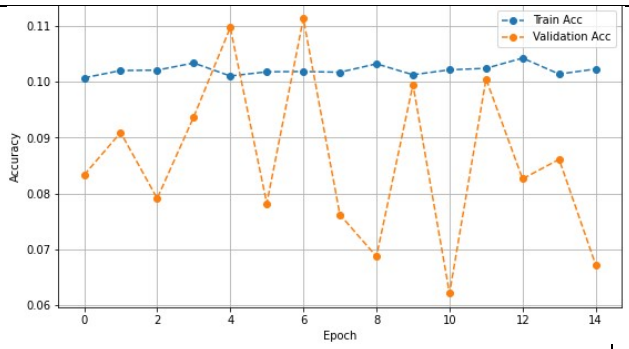
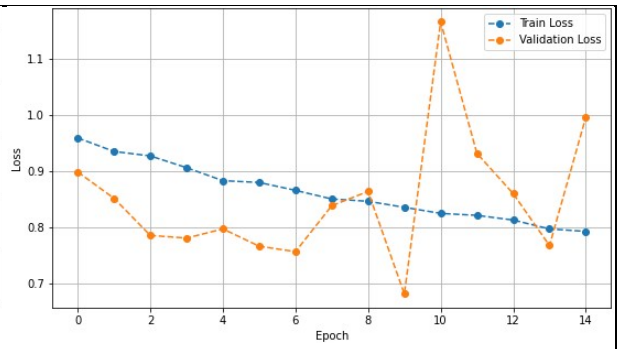
height_shift_range = 0.1

horizontal_flip=True

vertical_flip=False

Optimizer	Adam
Loss function	Sparse categorical crossentropy
Batch size	128
Learning rate	0.001
Validation split	0.2
Dropout	0.2 (after each convolutional layer)

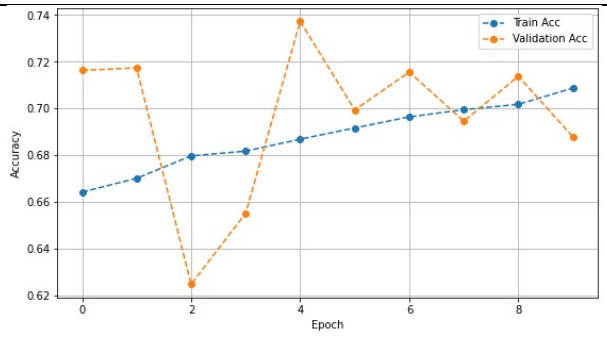
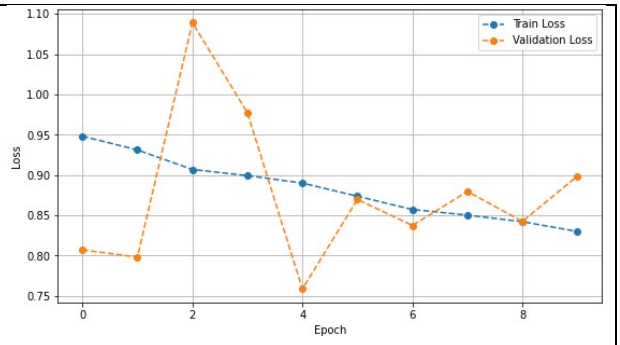
Epoch 10	 Training accuracy: 0.1000 Validation accuracy: 0.0785	 Training loss: 0.9732 Validation loss: 0.8583

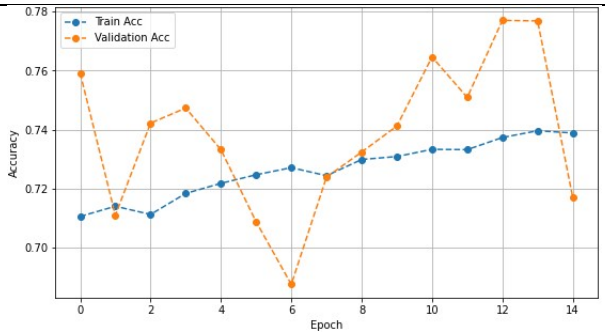
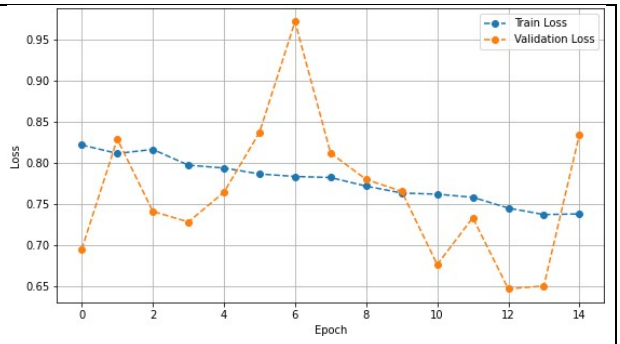
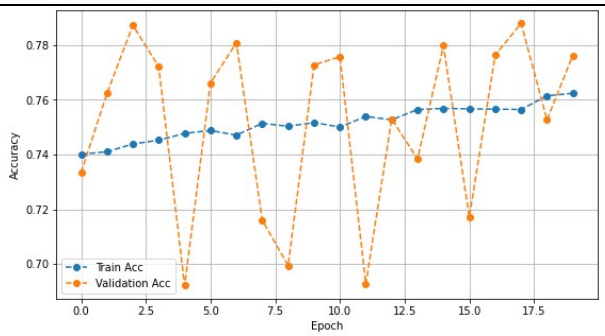
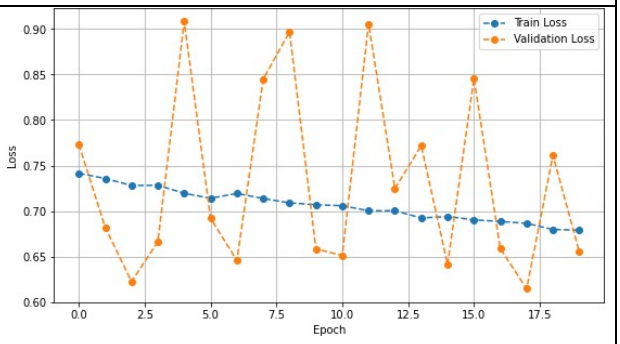
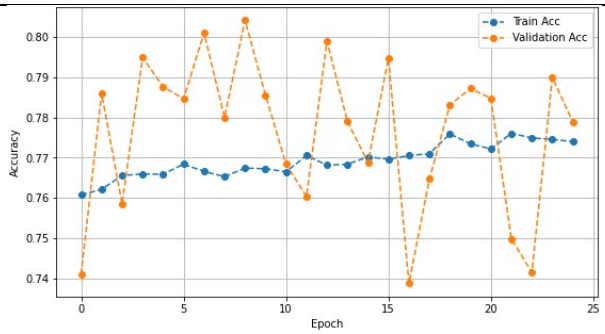
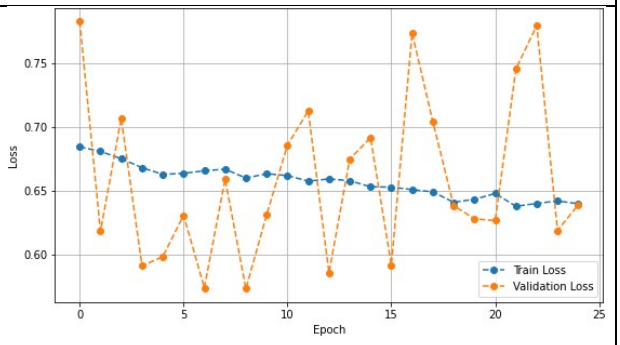
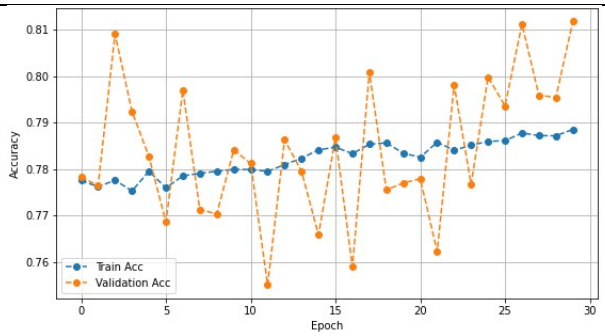
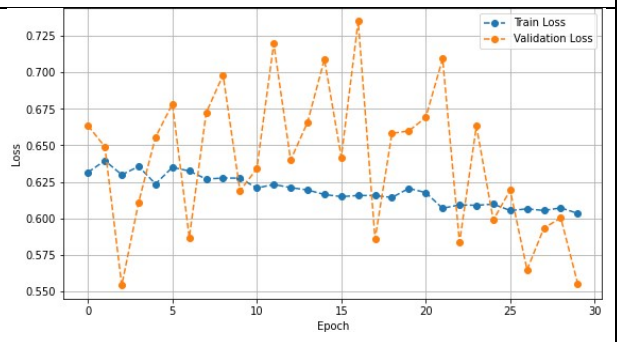
Epoch 15	 Training accuracy: 0.1023 Validation accuracy: 0.0672	 Training loss: 0.7928 Validation loss: 0.9959

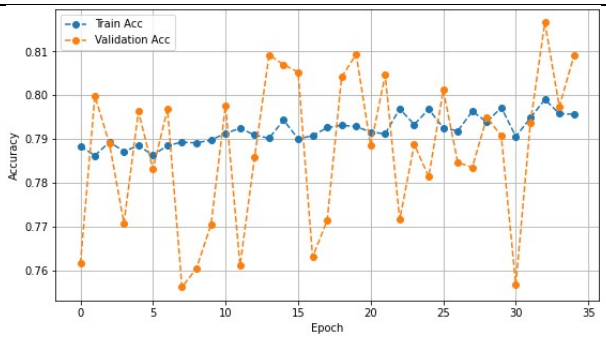
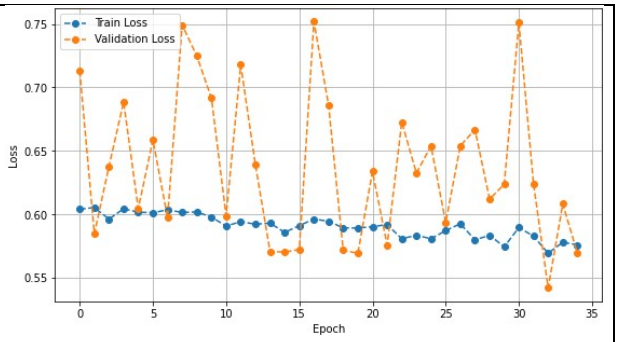
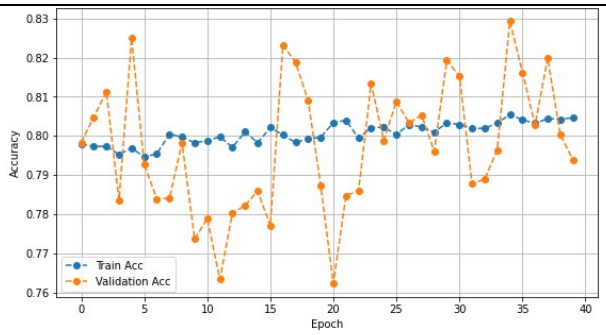
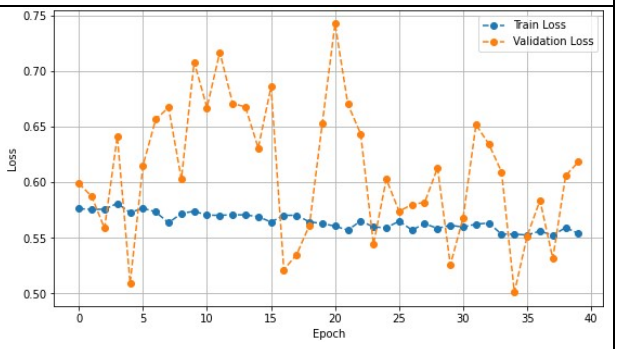
Observation: Here, the accuracy is not very good. So, in the next experiment the function will be changed to categorical crossentropy to improve.

Experiment 15: Data augmentation

Optimizer	Adam
Loss function	Categorical crossentropy
Batch size	128
Learning rate	0.001
Validation split	0.2
Dropout	0.2 (after each convolutional layer)

Epoch 10	 Training accuracy: 0.7087 Validation accuracy: 0.6879	 Training loss: 0.8299 Validation loss: 0.8987

Epoch 15	 Training accuracy: 0.7389 Validation accuracy: 0.7170	 Training loss: 0.7381 Validation loss: 0.8349
Epoch 20	 Training accuracy: 0.7624 Validation accuracy: 0.7763	 Training loss: 0.6789 Validation loss: 0.6553
Epoch 25	 Training accuracy: 0.7740 Validation accuracy: 0.7789	 Training loss: 0.6399 Validation loss: 0.6390
Epoch 30	 Training accuracy: 0.7878 Validation accuracy: 0.8112	 Training loss: 0.6065 Validation loss: 0.5647

Epoch 35	 Training accuracy: 0.7957 Validation accuracy: 0.8090	 Training loss: 0.5761 Validation loss: 0.5695
Epoch 40	 Training accuracy: 0.8046 Validation accuracy: 0.7939	 Training loss: 0.5540 Validation loss: 0.6184

<u>Observation</u>: Accuracy is highest 81.12% (0.8112) at epoch 30 and loss is also the lowest (0.5647).

But, we can see that the validation accuracy is higher than testing accuracy and validation loss is lower than training loss. This means the model has generalized or the dropout is slightly higher. As there is no dropout is testing, only in training, this may happen.

However, validation accuracy is 79.4% (0.7939) at epoch 40 which is better than previous final accuracy (78%). Moreover at epoch 40 the validation accuracy is slightly lower than testing accuracy and validation loss is higher than training loss which should be the case always.