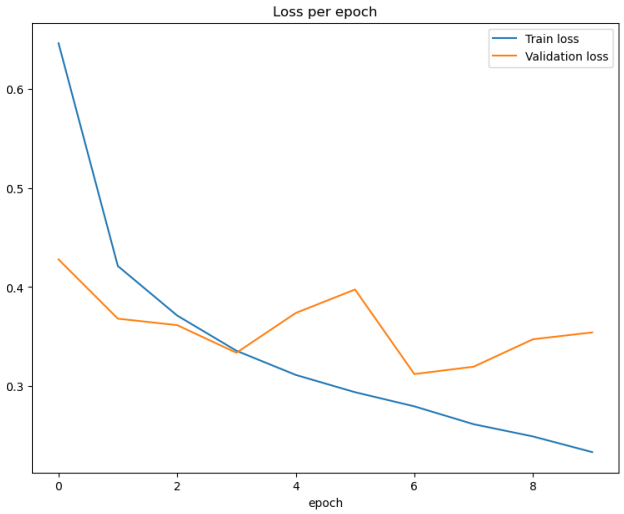
CSC 480/680 – Assignment 2

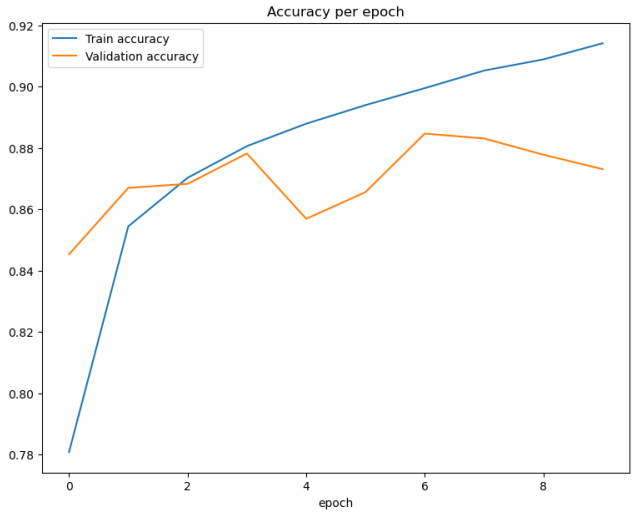
Name: Dhruv Jain

Answer 1.

Using the train and test datasets and then transforming that into a vector representation for MLP. Then using tensor showing test and train set for 2D- CNN model. Then training the CNN 2d trains set and MLP as 1D training set. Running epochs 10 times we get best accuracy at 91.42% for MLP model

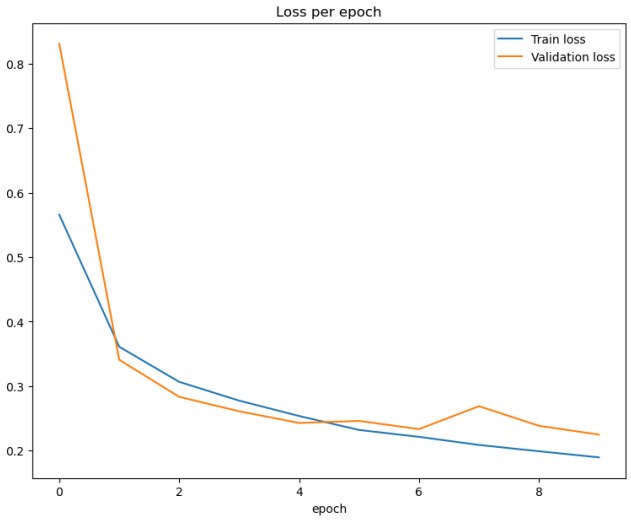
Showing the loss per epoch and accuracy per epoch for MLP in below figure.

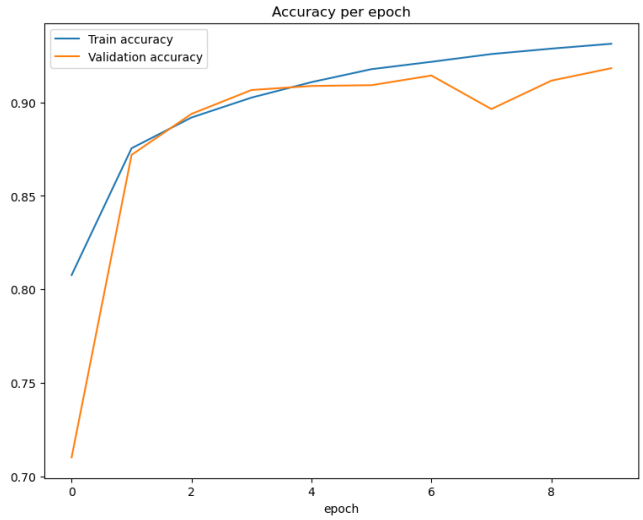




Now, running epochs 10 times we get best accuracy at 93.13% for CNN model

Showing the loss per epoch and accuracy per epoch for CNN in below figure.





Train acc	Test acc	-
0.914	0.873	MLP model
0.931	0.918	CNN model

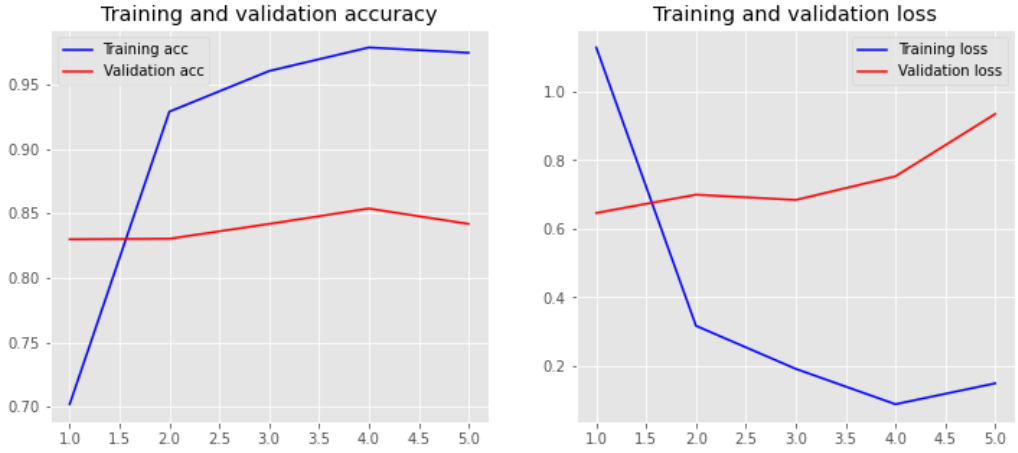
Multilayer Perceptron use one perceptron for each input for a pixel in an image and the number of weights rapidly becomes unmanageable for large images. Moreover, sometimes they react differently to an input and its might shift the origin data set quality.

Convolution neural network can work better with spatial method. I think that CNNs has a ability to develop an internal representation of a 2-D image. Due to this CNN works better than MLP models.

Answer 2.

Baseline model with MLP:

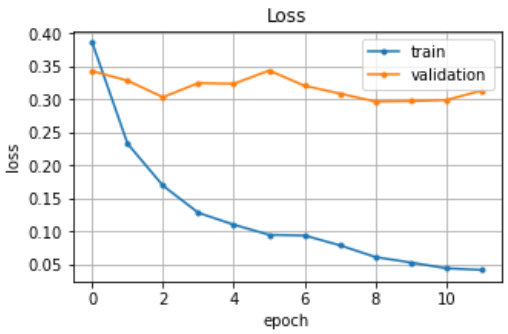
```
Training Accuracy: 0.9956, Validation Accuracy:  0.8418
```

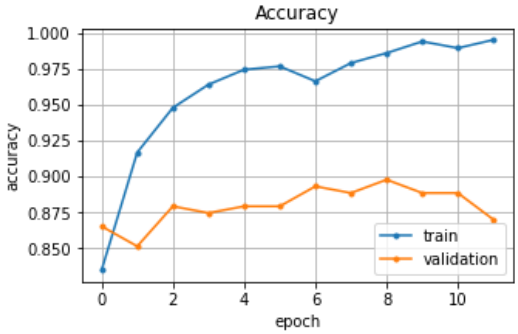



```
Accuracy: 0.978359
Precision: 0.976743
Recall: 0.978359
F1 score: 0.977015
```

Final model MLP with hyper parameters:

```
Training Accuracy: 0.9981 , Validation Accuracy:  0.8572
```






```
Accuracy: 0.975969
Precision: 0.977256
Recall: 0.975969
F1 score: 0.976533
```

We can observe that in baseline model we are getting 97.8% accuracy while in final model using hyper parameters, we are getting 97.6% accuracy. Generally, this using hyper parameters should increase the accuracy but in this we have achieved the same accuracy for both baseline and final model. Hyperparameter tuning consists of finding a set of optimal hyperparameter values for a learning algorithm while applying this optimized algorithm to this data set. That combination of hyperparameters maximizes the model's performance, minimizing a predefined loss function to produce better results with fewer errors