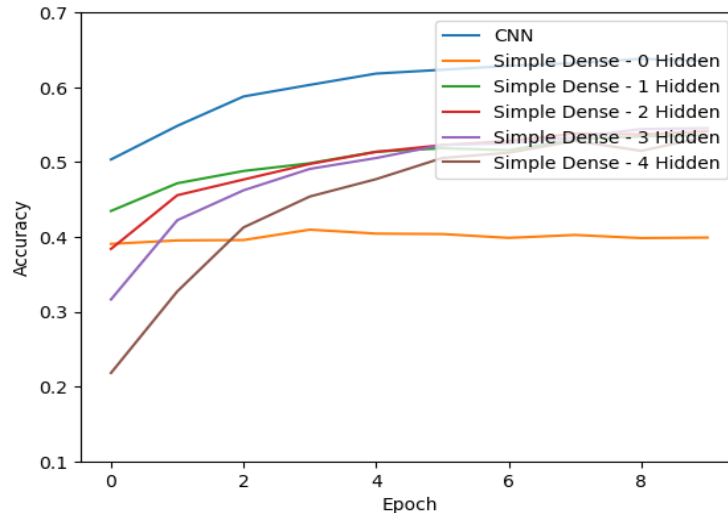


HW2 – Baldur Hua

[GitHub link](#)

(a)

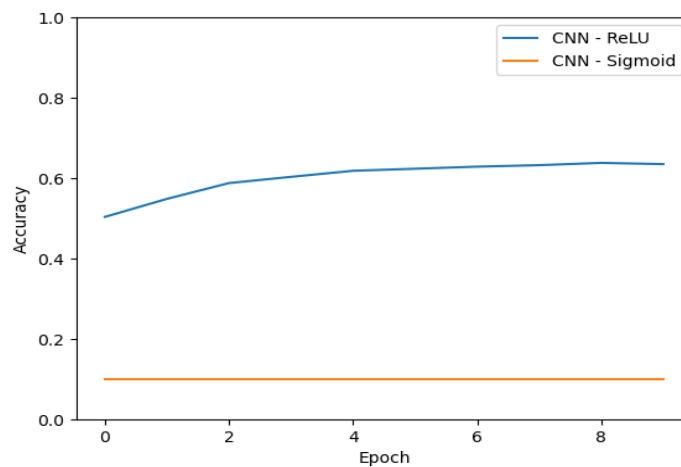


The CNN predicted the data with the highest accuracy among all other simple dense neural networks, as it lowers the intensity of the weights. Therefore, it increases the simplicity of achieving optimization, reducing the amount of data for assimilation

The increasing hidden layers of the simple dense neural networks results lower accuracy, as the model seems to be unable to achieve the best optimization.

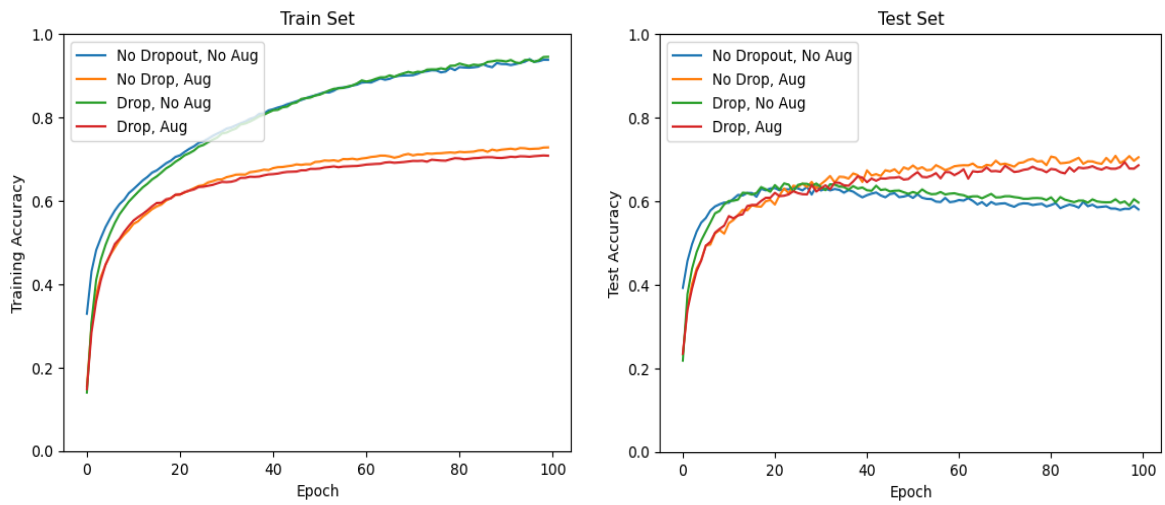
For the dense neural network with 2 or more hidden layers should be equivalent to the one with one hidden layer. However, the accuracies vary widely.

(b)



The ReLU units provide higher accuracy than sigmoid units. Sigmoid units hardly change the gradient, as the gradient is close to zero when start off from the tail side.

(c)



The model of No-dropout-No-data-augmentation and Drop-No-data-augmentation have the best assimilation rate. However, as the number of epochs increases, it overfits the model. While data augmentation performs a little worse with no dropout, however, significantly reducing the overfitting. Also, with both dropout and data augmentation, the model cannot perform well. Yet, the model does not seem to be overfitting. The results are subject to change due to the choice of augmentation and dropout rate.