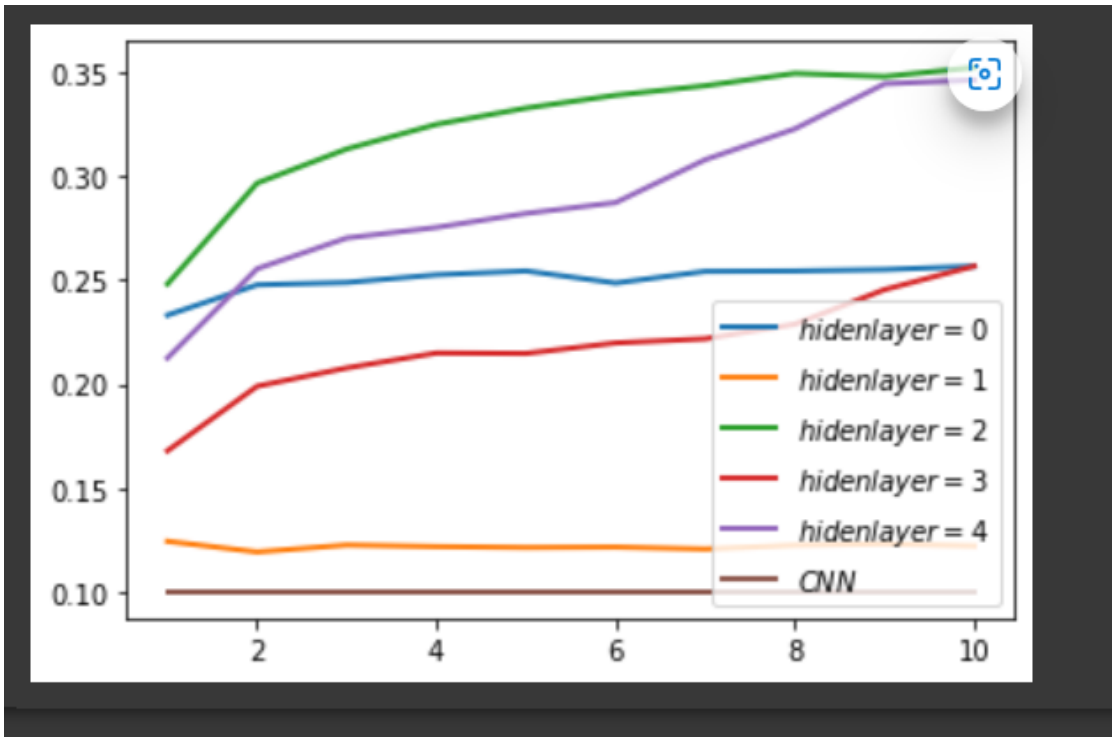
6289

Wang hang
Qs 1:



Qs1:

The CNN model has the best result since it could extract spatial info on animage, and build complex feature maps that help the classification. In short, better feature
For the fully connected network, as the number of hidden layer increase, the performance decrease. This decrement in performance is result from the overfitting.

Qs2:

The biggest advantage of ReLu is indeed non-saturation of its gradient, which makes the model with ReLU performs better than sigmoid. Moreover, the model with ReLU is less expensive in computation resource.

Qs3:

validation accuracy for the model with dropout is better than the model without dropout. This is because the dropout technique prevents the model to overfit. It turns off a percentage of neurons in the layer so the network has to adjust itself to cover up the loss of neurons by changing the values of other neurons so it doesn't get the chance to overfit or memorize any particular values. Although the training accuracy will be higher on the model without dropout, and that is only because of overfitting. So we should always consider the validation accuracy for measuring the performance of the model.