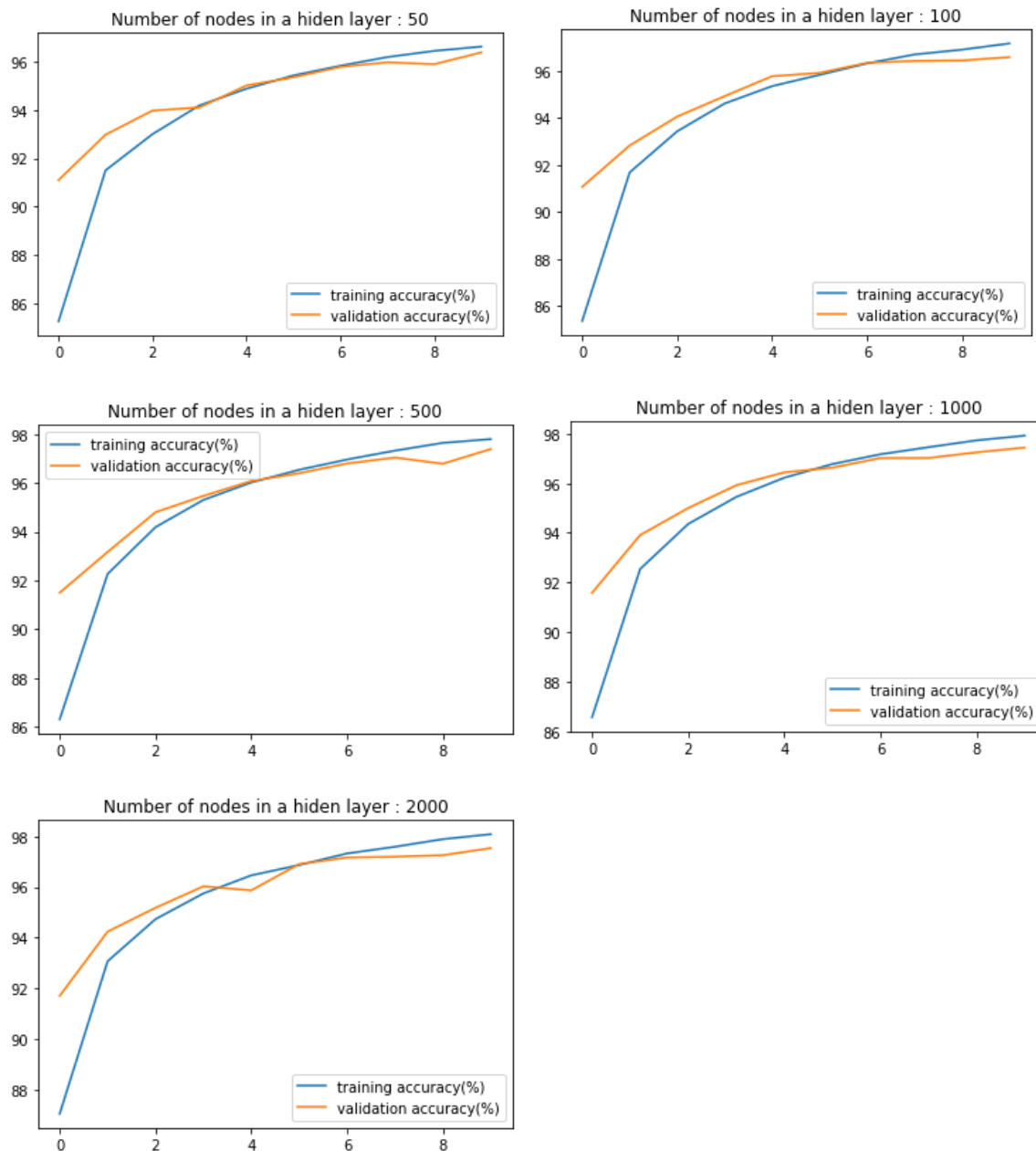


## Assignment 2 – Neural Networks

1970030 park guen a

### 1. Performance over different layer width

A. Show your 5 different plots of training & validation accuracy with different node # [50, 100, 500, 1000, 2000] in the single hidden layer.

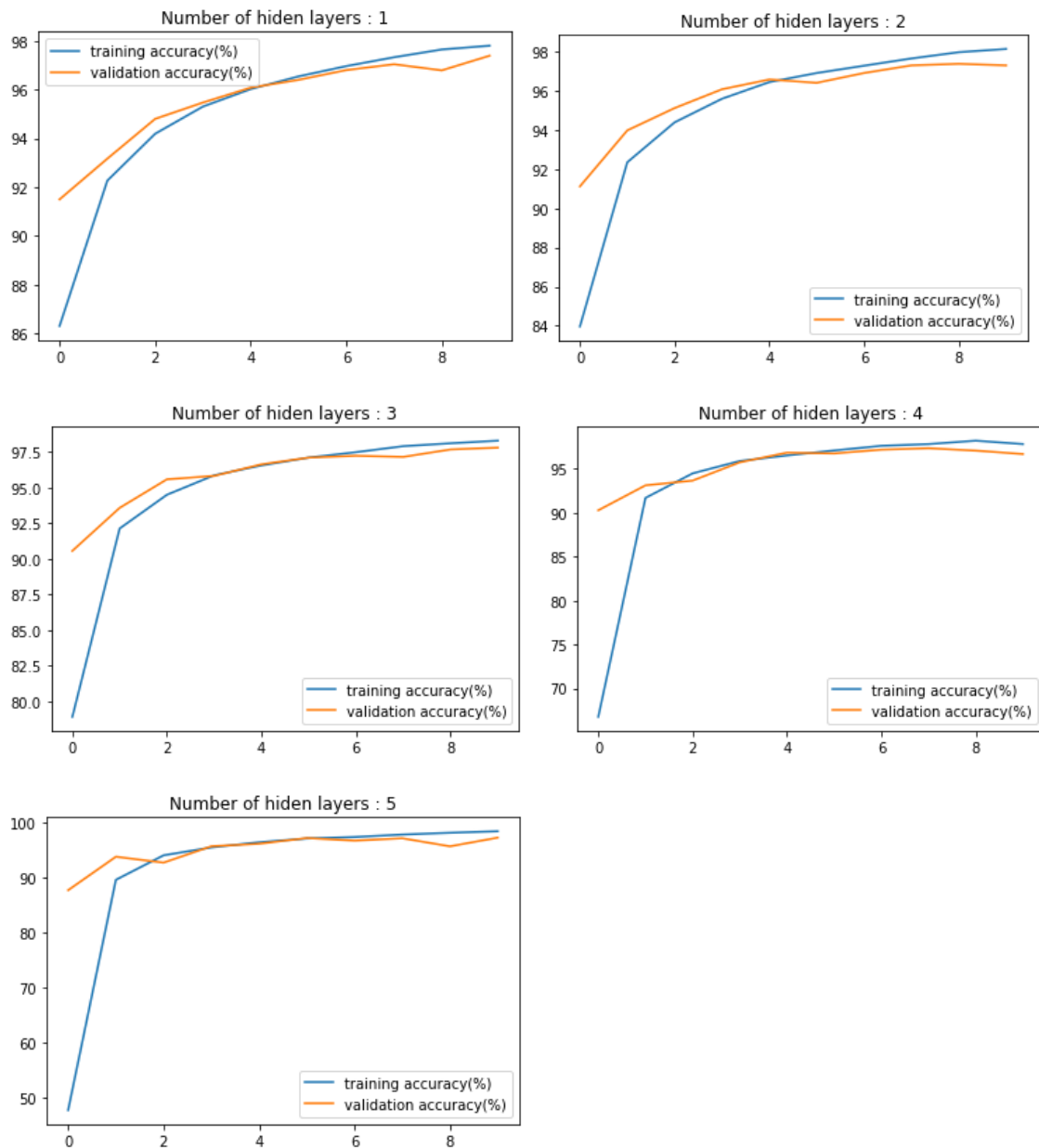


B. Based on the derived results, briefly analyze performance over different layer width verbally.

우선 hidden layer의 node 수를 늘릴수록 training accuracy는 높아졌다. 이때 validation accuracy 역시 높아졌으나 training loss는 계속 감소하는 반면 validation loss는 node를 500에서 1000개로 늘릴 때 증가하였기 때문에 overfitting이 발생함을 알 수 있다.

## 2. Performance over different layer depth

A. Show your 5 different plots of training & validation accuracy with different # of hidden layers [1, 2, 3, 4, 5].



B. Based on the derived results, briefly analyze performance over different layer depth verbally.

우선 hidden layer의 수를 늘릴수록 training accuracy는 점점 높아지다가 hidden layer가 4개일 때 accuracy가 떨어진다. 그리고 training loss와 validation loss 역시 layer를 늘릴수록 줄어들다가 layer가 4개일 때는 2개의 loss 모두, 5개일때는 validation loss가 늘어나는 것을 확인할 수 있다. 그래서 이를 통해 overfitting이 발생함을 알 수 있다.

### 3. Optimizing the neural networks beyond 97.5%

A. Optimize the network so that the validation accuracy exceeds 97.5%.

B. What efforts have you made to improve accuracy of your final model?

앞선 1,2번에서 성능이 가장 좋았던 hidden layer 3개를 이용하려고 했다. 그리고 node 수를 다양하게 바꿔가며 train해보았는데 (784,128,256,512,10)으로 node 수를 점점 늘려가는 방법과 (784,512,256,128,10)로 점점 줄여가는 방법 중 줄여가는 방법이 더 성능이 좋아 선택했다. 그리고 learning rate를 바꿔가며 test해보았을 때 0.05일 때가 가장 성능이 좋았다. 또한 epoch이 40이 넘어가자 training accuracy가 100이 되어서 마무리해주었다. 마지막으로 초기 W값을 잘 정해주기 위해서 Xavier initialization을 사용해줬더니 성능이 좋아졌다. CNN을 쓰면 성능이 더 좋아질 것이다. 최종결과는 validation accuracy가 98.2000이 나왔다.