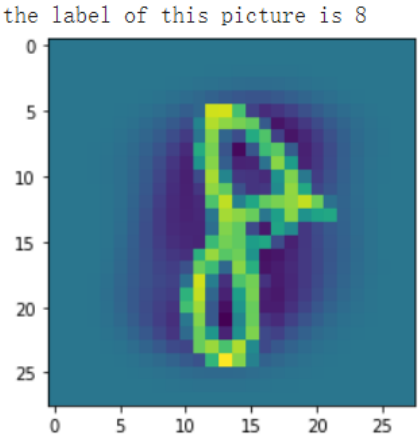
EE 228 HW4 report

Q1:

Data preprocessing:

- 10000 for training set
- 10000 for test set
- normalize



Set up nn.model:

- ReLU activation
- Cross_entropy
- He initialization
- Single hidden layer
- Learning rate: 0.003

```
SNN(
  (hidden): Linear(in_features=784, out_features=10, bias=True)
  (output): Linear(in_features=10, out_features=10, bias=True)
  (softmax): Softmax(dim=None)
  (relu): ReLU()
  (dropout): Dropout(p=0, inplace=False)
)
```

test model:

- K=10
- Q = 1 (no dropout)

```
Epoch [79/80], Train Loss: 1.60500, Test Loss: 1.60716, Train Accuracy: 0.88500, Test Accuracy: 0.87500
```

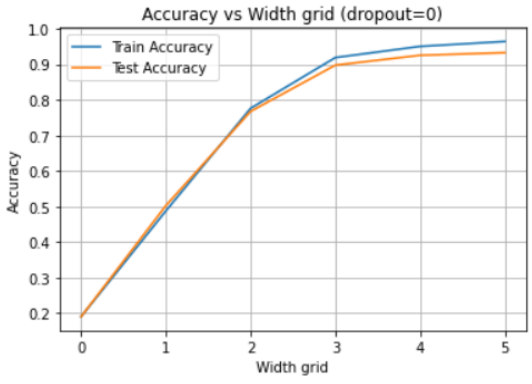
Q2:

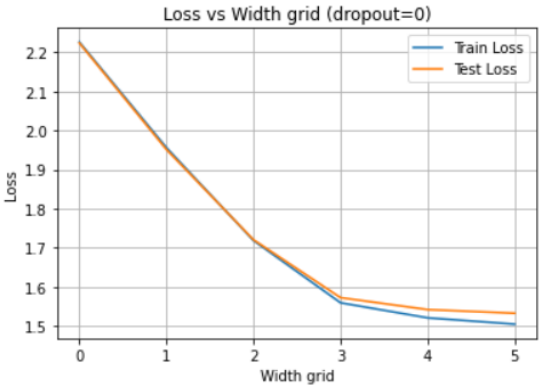
Width grid = [1,3,5,10,15,25,40]

Dropout grid = [0.1,0.5,1]

2.1

As k increases, does the performance improve? **Yes,the results improved significantly.**





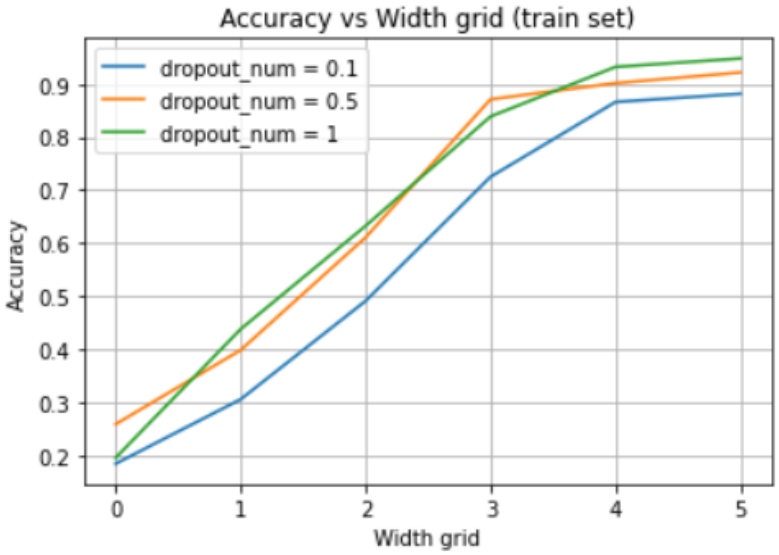
At what k, training accuracy becomes 100%?

2.2

What is the role of p on training accuracy? **Prevent model overfitting**

When p is smaller, is it easier to optimize or more difficult? **More difficult**

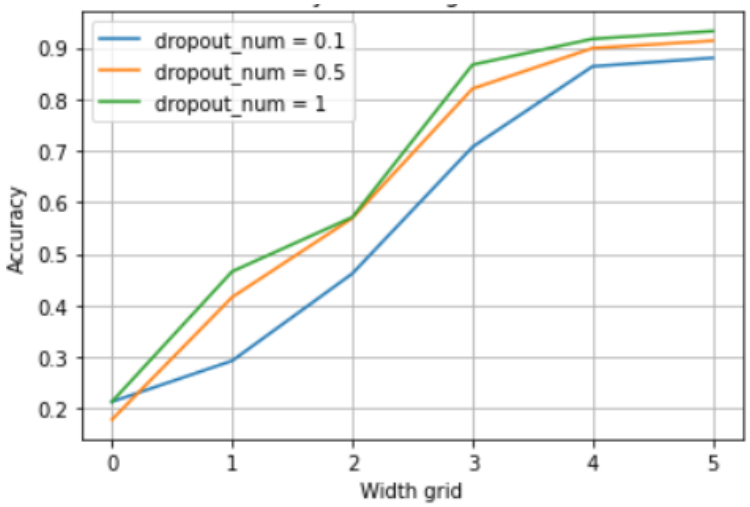
For each choice of p, determine at what choice of k, training accuracy becomes 100%.



2.3

Does dropout help with the test accuracy? **Not significant I think.**

For which (k, p) configuration: Do you achieve the best test accuracy? **I think it should be when k = 40 and p =1.**

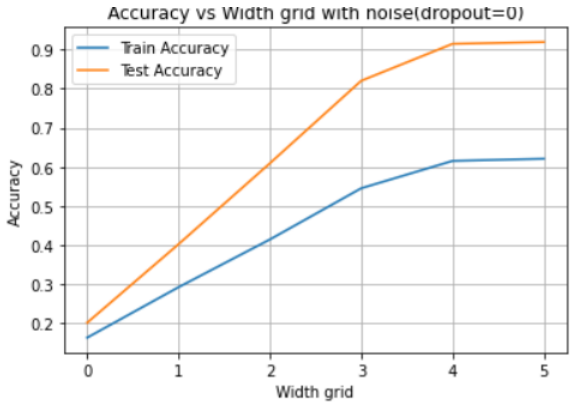


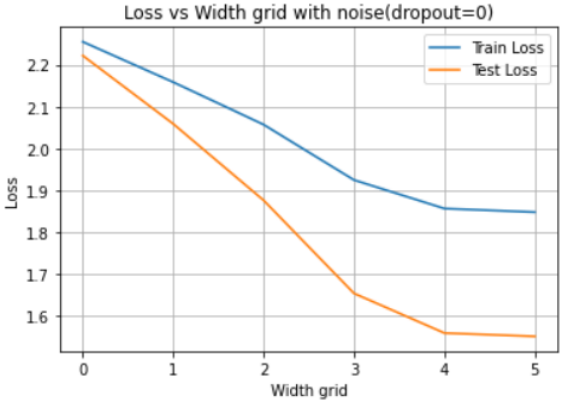
Q3. adding noise

- **Pick 40% of the training examples at random and assign their label to a number from 1 to 9 at random.**

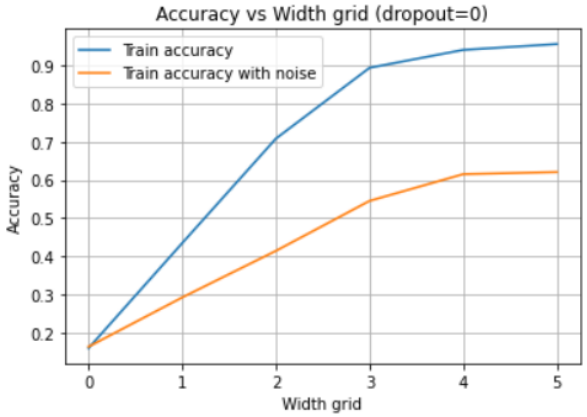
3.1

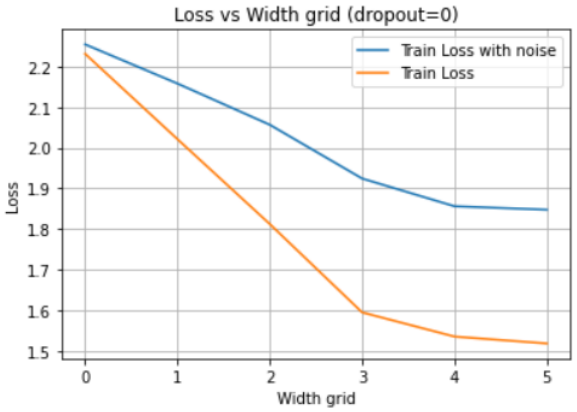
From the result, we can see that the training accuracy decreased significantly, but the test accuracy is stable.



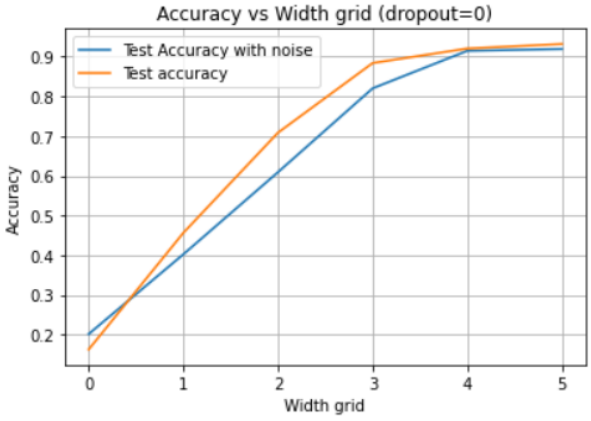


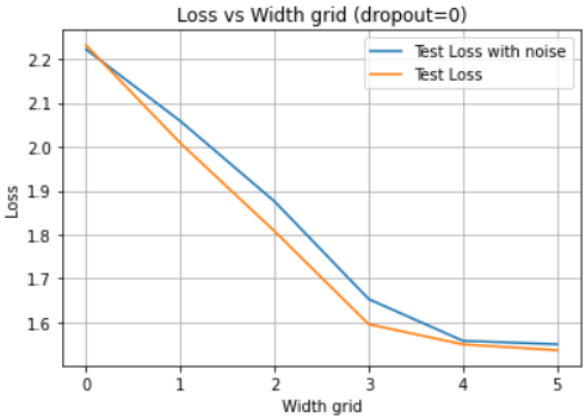
Compare the training accuracy and loss with the previous question(without noise). The noise has a great influence on the training set.





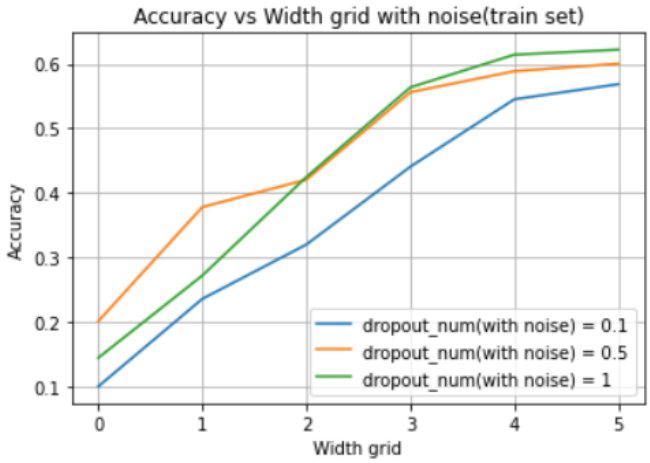
But not significant on the test set.





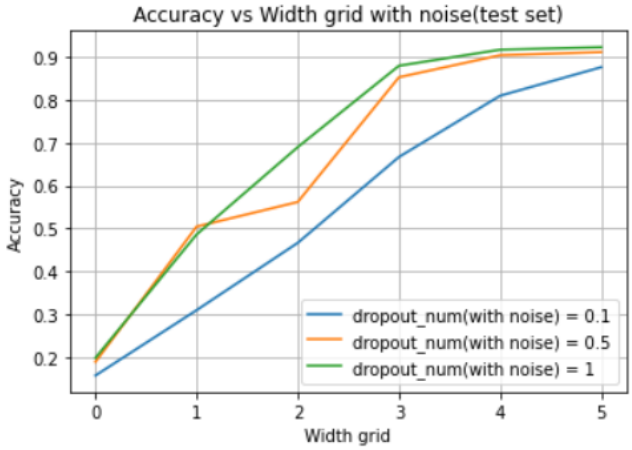
3.2

Compared to the result in 3.1, we can see that the dropout grid can improve the training accuracy when adding noise in the training set.



3.3

Applied the dropout method on the test set, the improvement is not significant.



So we can conclude that the dropout method can improve the training accuracy when the data suffers from noise.

Q4:

Comment on the differences between Step 2 and Step 3. How does noise change things?

After adding noise on the training set, the training accuracy decreased significantly. So noise is not good for model training.

But because no noise was added to the test set, the results were stable..

For which setup dropout is more useful?

For question 3, the dropout is more useful. Especially on the training set which suffered from noise. So, in the fully connected model, dropout is useful for againsting noise.