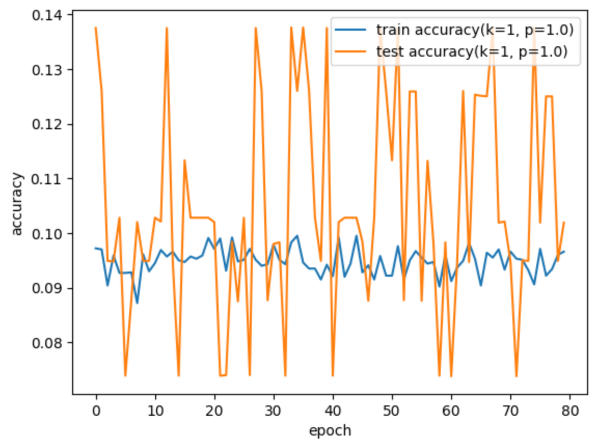
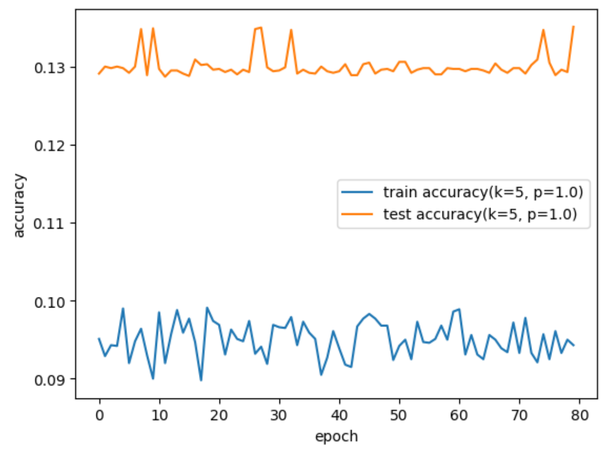
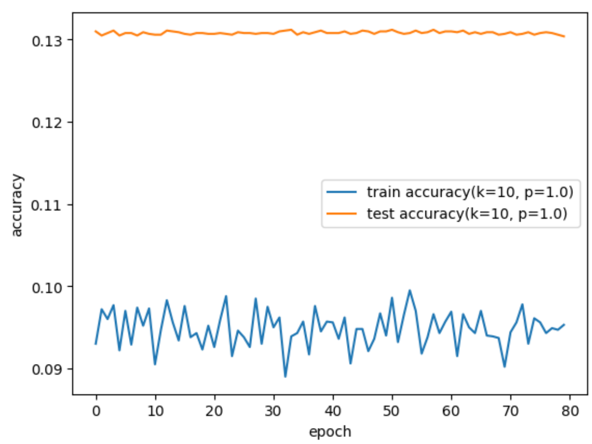
HW#3 - MNIST with Dropout and Noise

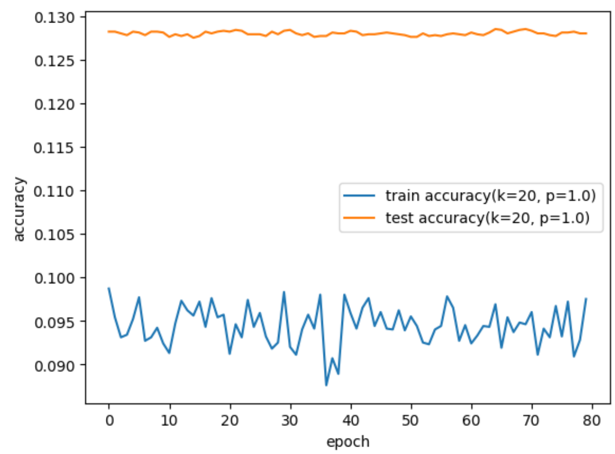
Yin-Yu Chang

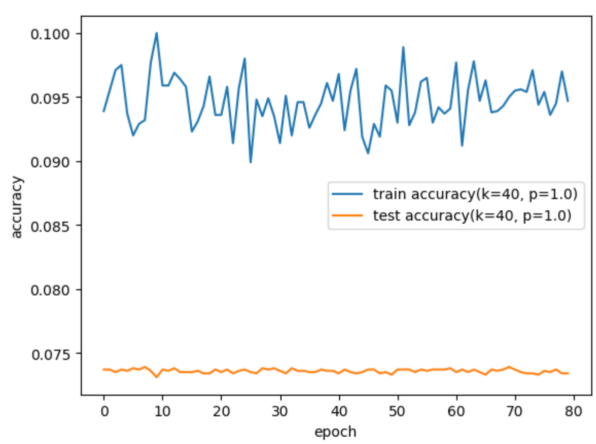
2-1. Fix p=1.0





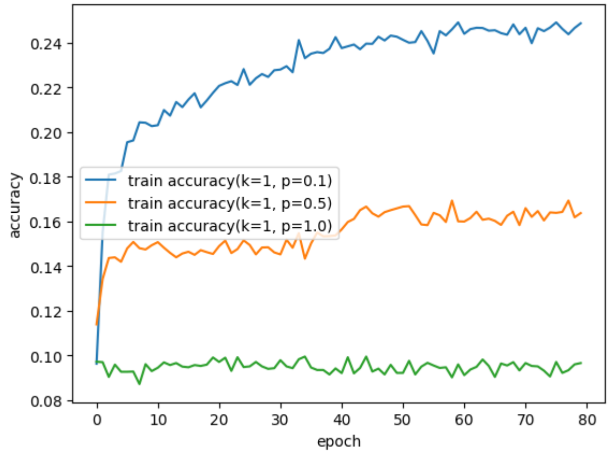


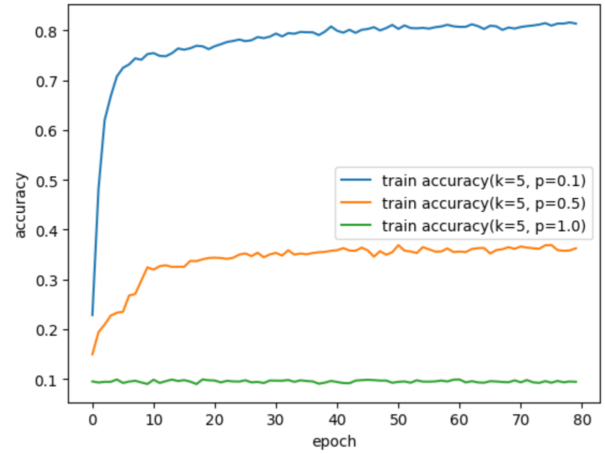


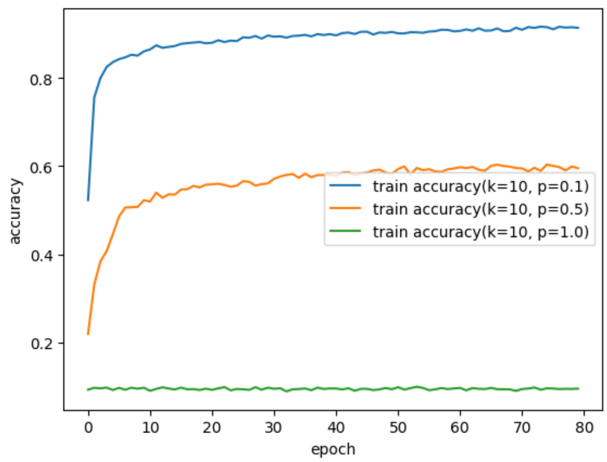


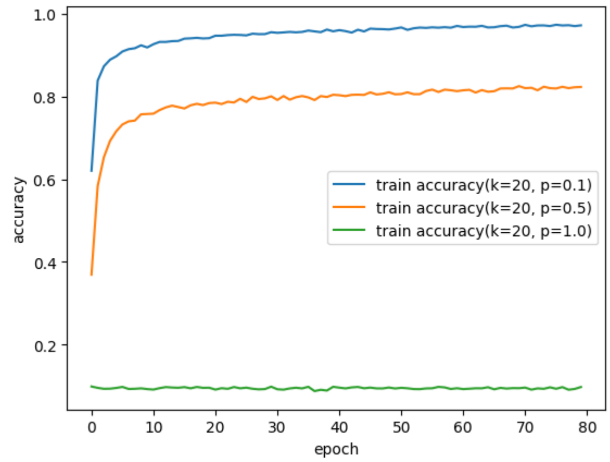
As k increases, the test accuracy decreases. So the performance doesn't improve. Training accuracy doesn't become 100 in any of k. It remains between 0.09 and 0.1 in any of k.

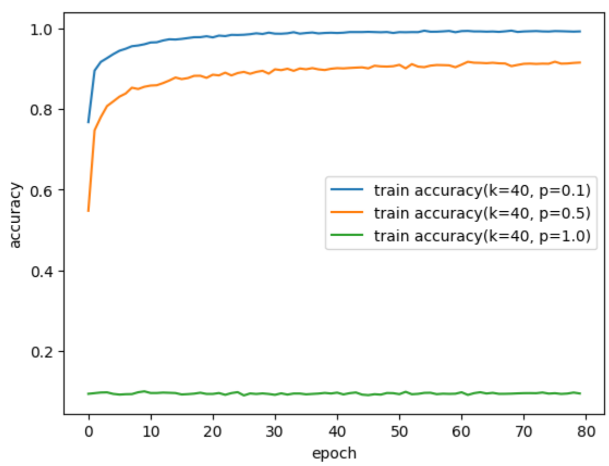
2-2. training accuracy as a function of k and for different p





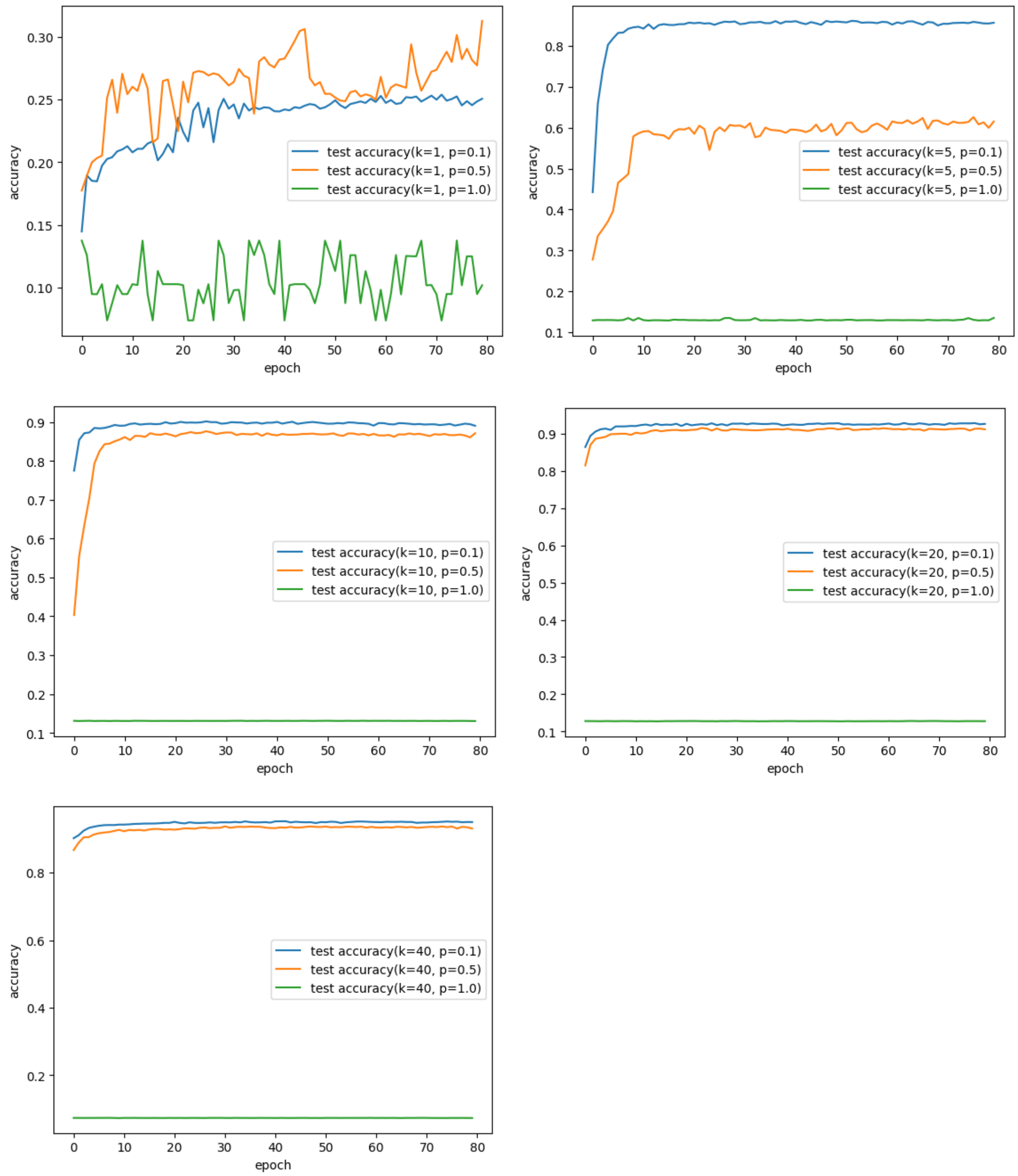






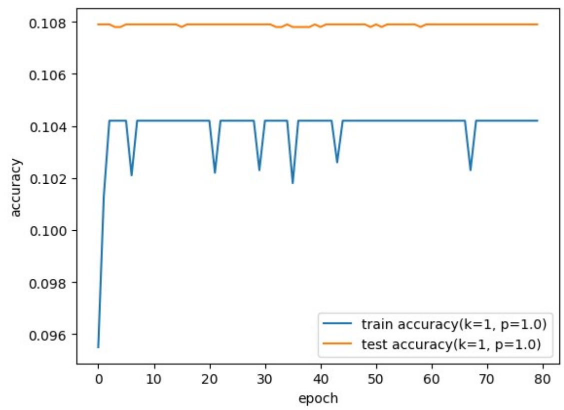
We can see when p is smaller, the accuracy would increase. So, when p is smaller, it is easier to optimize. When we choose p = 0.1, the accuracy when k equals 20 or 40 will be near 100%; when we select p = 0.5 or 1.0, no accuracy will be near 100% for any of k.

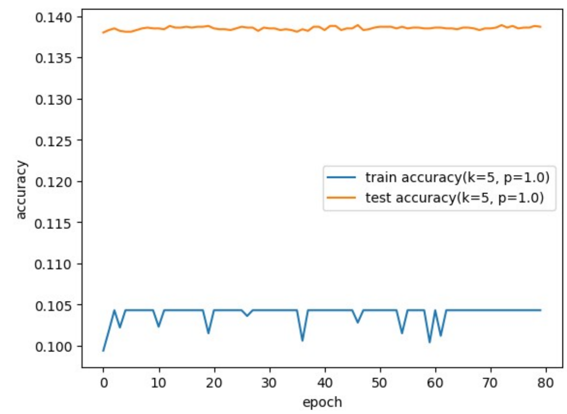
2-3. test accuracy as a function of k and for different p

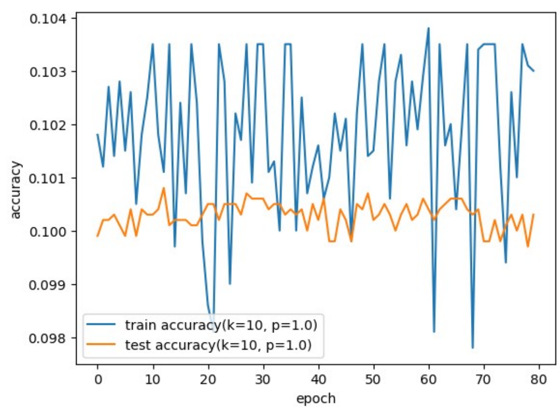


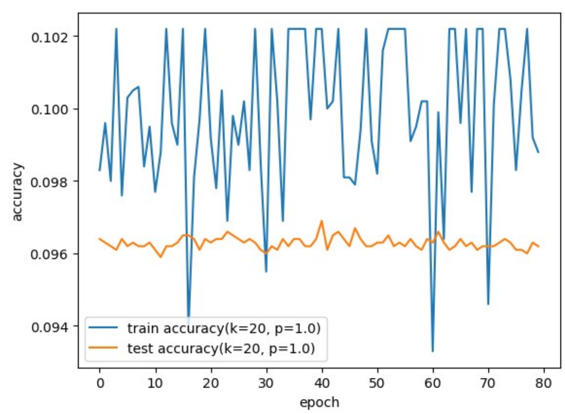
In most cases, when p = 0.1, we have the best accuracy. So, dropout helps with the accuracy. I achieve the best test accuracy in the (k,p) = (40, 0.1) configuration.

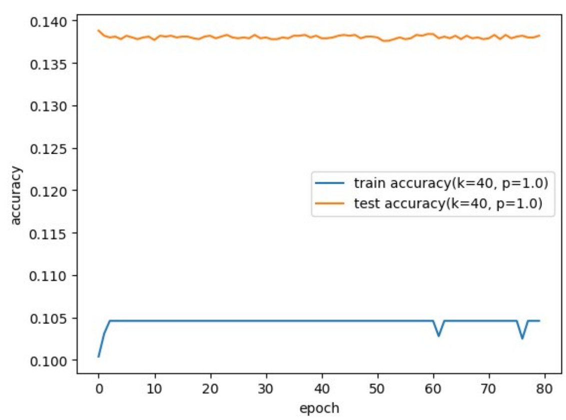
3-1. Fix p=1.0





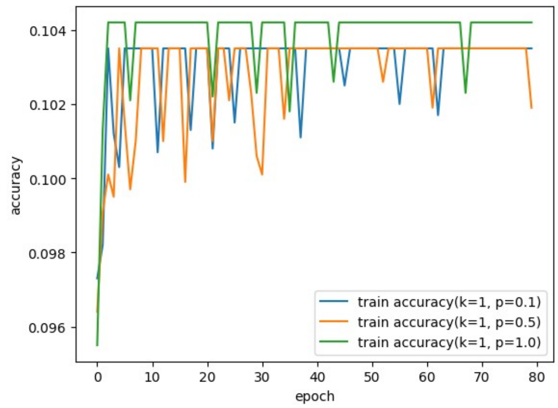


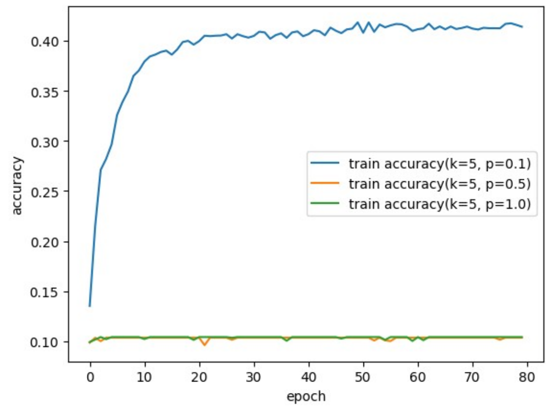


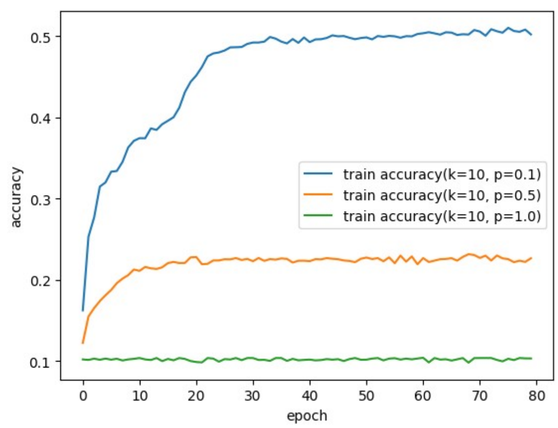


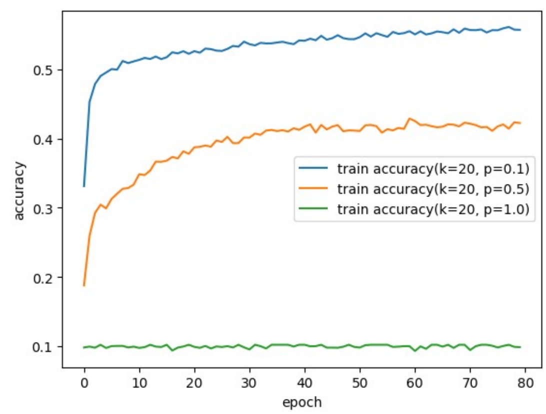
As k increases, the test accuracy decreases. So the performance doesn't improve. Training accuracy doesn't become 100 in any of k.

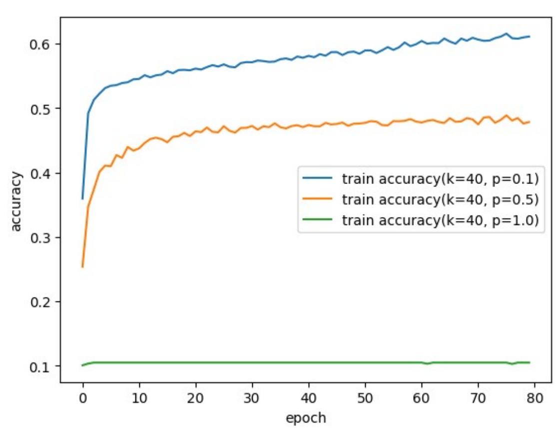
3-2. training accuracy as a function of k and for different p





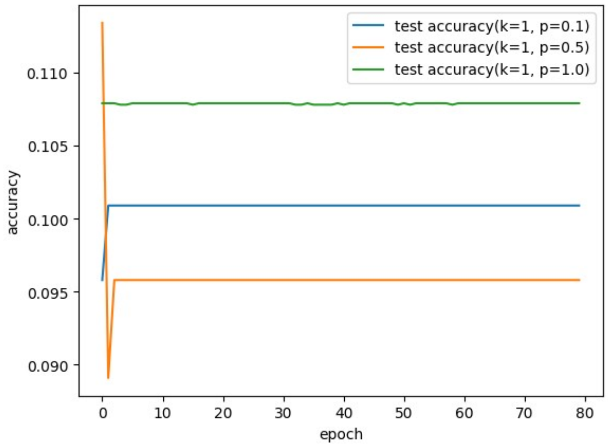


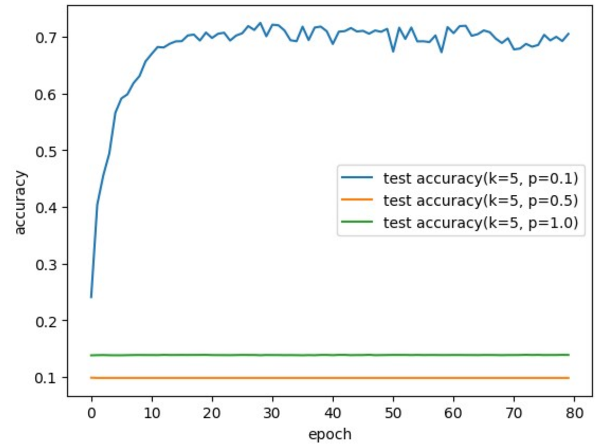


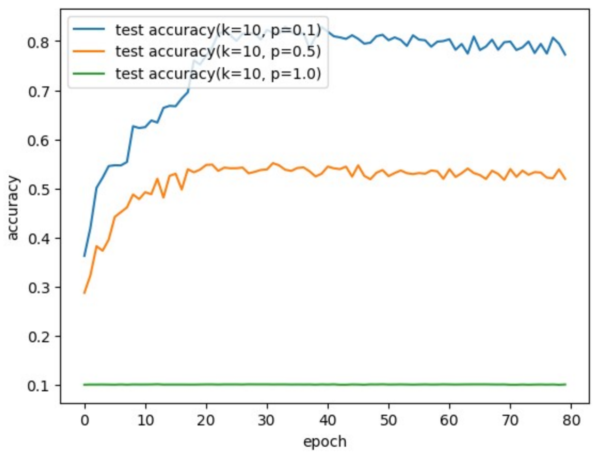


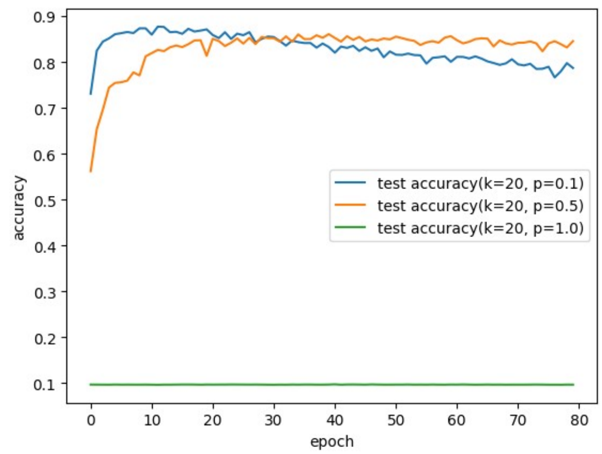
We can see when p is smaller, the accuracy would increase. So, when p is smaller, it is easier to optimize. No matter whether we choose p = 0.1 0.5 or 1.0, no accuracy will be near 100% for any of k.

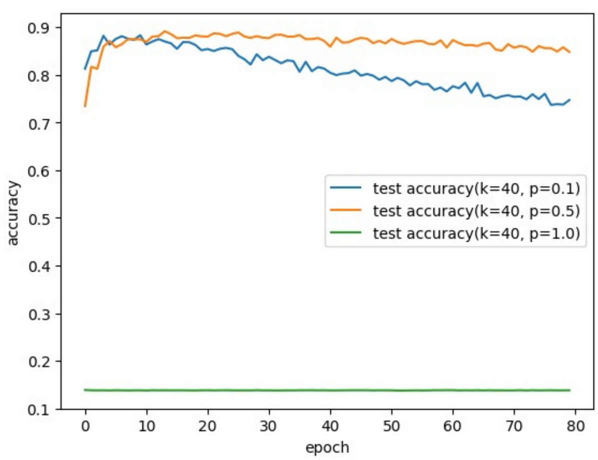
3-3. test accuracy as a function of k and for different p











When k = 5 or 10, when p = 0.1, it has the best accuracy. When k = 20 or 40, when p = 0.5, it has the highest accuracy. So, dropout helps with the accuracy when k is not large. I achieve the best test accuracy in the (k,p) = (40, 0.5) configuration.

4. Comment on the differences between Step 2 and Step 3. How does noise change things? For which setup dropout is more useful?
When we add noise, the training accuracy decrease. But the test accuracy can still be good when k is karge and p = 0.5. I think dropout is more useful when there is no noise. Or when thertr is noise and p = 0.5 and k is large.