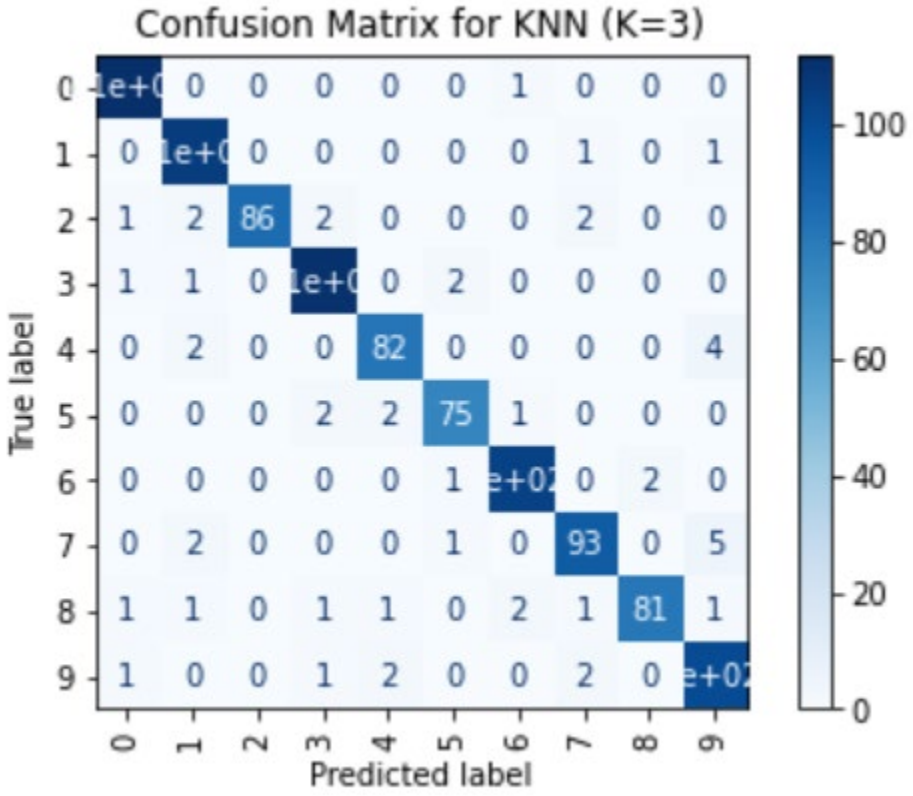
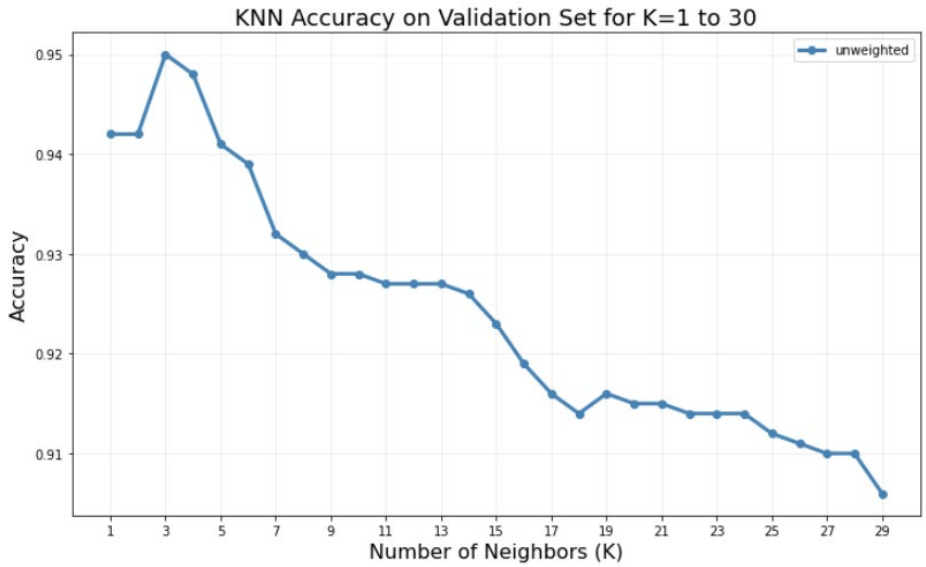
Week 4 kNN Peer Review Project



These observations suggest that digits 7 and 9, in particular, are the most prone to misclassification in this KNN model.



The highest accuracy is achieved at K = 3, where the accuracy peaks at approximately 0.95

```
|: # TODO : Complete the first subtask for max_depth

   # your code here
   tree1 = build_dt(X_train, y_train, max_depth = 2)
   y_hat = tree1.predict(X_test)
   prec1 = calculate_precision(y_test, y_hat, pos_label_value=1.0)
   rec1 = calculate_recall(y_test, y_hat, pos_label_value=1.0)

   print("Tree precision is: ", prec1)
   print("Tree recall is: ", rec1)
   print("Tree depth is: " , tree1.get_depth())
```

```
Tree precision is:  0.8413793103448276
Tree recall is:  0.6931818181818182
Tree depth is:  2
```

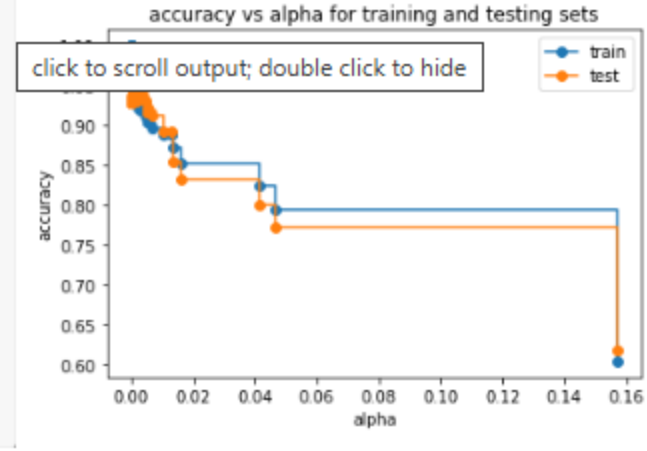
```
In [30]: # TODO : Complete the second subtask for max_depth

         # your code here
         tree1 = build_dt(X_train, y_train, max_depth = 2)
         y_hat = tree1.predict(X_test)
         prec1 = calculate_precision(y_test, y_hat, pos_label_value=1.0)
         rec1 = calculate_recall(y_test, y_hat, pos_label_value=1.0)

         print("Tree precision is: ", prec1)
         print("Tree recall is: ", rec1)
         print("Tree depth is: " , tree1.get_depth())
```

```
Tree precision is:  0.8413793103448276
Tree recall is:  0.6931818181818182
Tree depth is:  2
```

Max depth results in higher precision but lower recall because the tree cannot grow deep enough to capture complex patterns. Modifying the leaf nodes can yield higher recall but may reduce precision if the tree overfits the training data.



The parameter alpha in the plot appears to act as a regularization parameter, controlling the complexity of the model. The performance is modulated by alpha in such a way that lower values help the model capture data complexity, while higher values overly constrain the model, reducing its predictive power