```
import pandas as pd
import seaborn as sns
import numpy as np
from sklearn.datasets import load_iris

iris = load_iris()
```

```
df = pd.DataFrame(data=iris.data, columns=iris.feature_names)
df['target'] = iris.target
df
```

```
df.columns=['sl','sw','pl','pw','label']
df
```

```
from sklearn.model_selection import train_test_split
train,test=train_test_split(df,test_size=0.2)
```

```
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import cross_val_score
```

+ Code + Text

```
max_k_range=train.shape[0]//2
k_list=[]
for i in range(3,max_k_range,2):
  k_list.append(i)
```

```
cross_validation_scores=[]
x_train=train[['pl','pw']]
y_train=train[['label']]
```

```
for k in k_list:
  knn = KNeighborsClassifier(n_neighbors=k)
  scores=cross_val_score(knn,x_train,y_train.values.ravel(), cv=10,scoring='accuracy')
  cross_validation_scores.append(scores.mean())
```

```
cross_validation_scores
```

```
import matplotlib.pyplot as plt
```

```
plt.plot(k_list,cross_validation_scores)
plt.xlabel('k value')
```

```
plt.ylabel('accuracy')
plt.show()
new_k=k_list[cross_validation_scores.index(max(cross_validation_scores))]
```

```
knn = KNeighborsClassifier(n_neighbors=new_k)
```

```
knn.fit(x_train,y_train.values.ravel())
```

```
x_test=test[['pl','pw']]
y_test=test[['label']]
```

```
predictions = knn.predict(x_test)
predictions
```

```
from sklearn.metrics import accuracy_score
```

```
print("accuracy is " +str( accuracy_score(y_test.values.ravel(),predictions)))
```

```
 comparison = pd.DataFrame(
    {'pred':predictions, 'truth':y_test.values.ravel()})
comparison
```