

```
In [2]: import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
```

```
In [3]: from sklearn import datasets
```

```
In [4]: iris = datasets.load_iris()
```

```
In [5]: X = iris.data[:, :2]
```

```
In [6]: y = iris.target
```

```
In [7]: df_X=pd.DataFrame(X,columns=['SepalLengthCm','PetalLengthCm'])
```

```
In [8]: df_X
```

Out[8]:

	SepalLengthCm	PetalLengthCm
0	5.1	3.5
1	4.9	3.0
2	4.7	3.2
3	4.6	3.1
4	5.0	3.6
...
145	6.7	3.0
146	6.3	2.5
147	6.5	3.0
148	6.2	3.4

	SepalLengthCm	PetalLengthCm
149	5.9	3.0

150 rows × 2 columns

```
In [9]: df_y=pd.DataFrame(y,columns=['Species'])
```

```
In [10]: df_y
```

```
Out[10]:
```

	Species
0	0
1	0
2	0
3	0
4	0
...	...
145	2
146	2
147	2
148	2
149	2

150 rows × 1 columns

```
In [11]: from sklearn.preprocessing import StandardScaler
scaler = StandardScaler()
#scaler.fit(df_X)

x = scaler.fit_transform(df_X)
```

```
In [12]: df_y['Species'].value_counts()
```

```
Out[12]: 2    50
         1    50
         0    50
         Name: Species, dtype: int64
```

```
In [13]: from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(x, df_y, test_size=
0.20, random_state=3)
```

```
In [14]: from sklearn.neighbors import KNeighborsClassifier
classifier = KNeighborsClassifier(algorithm='auto', leaf_size=30, metri
c='minkowski',
    metric_params=None, n_jobs=1, n_neighbors=3, p=2, weights='uniform')
classifier.fit(X_train, y_train)
```

C:\Users\USER\anaconda3.x\lib\site-packages\ipykernel_launcher.py:4: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
after removing the cwd from sys.path.

```
Out[14]: KNeighborsClassifier(n_jobs=1, n_neighbors=3)
```

```
In [15]: k_pred=classifier.predict(X_test)
```

```
In [16]: k_pred
```

```
Out[16]: array([0, 0, 0, 0, 0, 2, 1, 0, 1, 2, 1, 0, 1, 1, 2, 0, 2, 2, 1, 0, 1,
1,
           2, 1, 0, 2, 2, 2, 1, 2])
```

```
In [17]: from sklearn.metrics import accuracy_score
from sklearn.metrics import confusion_matrix
```

```
In [18]: accuracy=accuracy_score(k_pred,y_test)
```

```
In [19]: accuracy
```

```
Out[19]: 0.7333333333333333
```

```
In [20]: matrix=confusion_matrix(k_pred,y_test)
```

```
In [21]: matrix
```

```
Out[21]: array([[10,  0,  0],
                [ 0,  6,  4],
                [ 0,  4,  6]], dtype=int64)
```

```
In [22]: final=pd.DataFrame(data=matrix,index=['setosa','versicolor','virginica'],columns=['setosa','versicolor','virginica'])
```

```
In [23]: final
```

```
Out[23]:
```

	setosa	versicolor	virginica
setosa	10	0	0
versicolor	0	6	4
virginica	0	4	6

```
In [24]: from sklearn.model_selection import cross_val_score
```

```
In [25]: print(cross_val_score(classifier,x,y,scoring='accuracy',cv=10).mean())
0.7333333333333334
```

```
In [26]: from sklearn.metrics import classification_report
```

```
In [27]: report=classification_report(k_pred,y_test)
```

```
In [28]: report
```

```
Out[28]: '
           precision    recall  f1-score   support\n\n
1.00         1.00         1.00         10\n
0.60         10\n          2         0.60         0.60         10\n\n
\n  accuracy                    0.73         30\n  macro avg
0.73         0.73         0.73         30\nweighted avg
0.73         30\n'
```

```
In [99]: error=[]
for i in range(1,30):
    knn=KNeighborsClassifier(n_neighbors=i)
    knn.fit(X_train,y_train)
    pred_i=knn.predict(X_test)
    error.append(np.mean(pred_i!=y_test))
```

C:\Users\USER\anaconda3.x\lib\site-packages\ipykernel_launcher.py:4: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().
after removing the cwd from sys.path.

```
-----
-----
ValueError                                Traceback (most recent call last)
<ipython-input-99-450aa01a55ce> in <module>
      4     knn.fit(X_train,y_train)
      5     pred_i=knn.predict(X_test)
----> 6     error.append(np.mean(pred_i!=y_test))
      7

~\anaconda3.x\lib\site-packages\pandas\core\ops\__init__.py in f(self,
other)
    767     def f(self, other):
    768
--> 769         other = _align_method_FRAME(self, other, axis=None)
    770
    771         if isinstance(other, ABCDataFrame):
```

```

~\anaconda3.x\lib\site-packages\pandas\core\ops\__init__.py in _align_m
ethod_FRAME(left, right, axis)
    642
    643         if right.ndim == 1:
--> 644             right = to_series(right)
    645
    646         elif right.ndim == 2:

~\anaconda3.x\lib\site-packages\pandas\core\ops\__init__.py in to_serie
s(right)
    634         if len(left.columns) != len(right):
    635             raise ValueError(
--> 636                 msg.format(req_len=len(left.columns), given
_len=len(right))
    637             )
    638             right = left._constructor_sliced(right, index=left.
columns)

```

ValueError: Unable to coerce to Series, length must be 1: given 30

```
In [86]: from sklearn.model_selection import RandomizedSearchCV
```

```
In [87]: k=np.random.randint(1,50,10)
```

```
In [88]: params={'n_neighbors':k}
```

```
In [89]: Random_search=RandomizedSearchCV(classifier,params,n_iter=5,cv=5,n_jobs
=-1,verbose=0)
Random_search.fit(X_train,y_train)
```

```

C:\Users\USER\anaconda3.x\lib\site-packages\sklearn\model_selection\_se
arch.py:765: DataConversionWarning: A column-vector y was passed when a
1d array was expected. Please change the shape of y to (n_samples, ), f
or example using ravel().
    self.best_estimator_.fit(X, y, **fit_params)

```

```
Out[89]: RandomizedSearchCV(cv=5,
                        estimator=KNeighborsClassifier(n_jobs=1, n_neighbors

```

```
-15\
```

```
-1),  
n_iter=5, n_jobs=-1,  
param_distributions={'n_neighbors': array([ 4, 26, 3  
9, 21, 42, 26, 12, 25, 12, 1])})
```

```
In [90]: Random_search.best_params_
```

```
Out[90]: {'n_neighbors': 21}
```

```
In [92]: classifier = KNeighborsClassifier(algorithm='auto', leaf_size=30, metri  
c='minkowski',  
metric_params=None, n_jobs=1, n_neighbors=21, p=2, weights='uniform')  
classifier.fit(X_train, y_train)
```

C:\Users\USER\anaconda3.x\lib\site-packages\ipykernel_launcher.py:3: DataConversionWarning: A column-vector y was passed when a 1d array was expected. Please change the shape of y to (n_samples,), for example using ravel().

This is separate from the ipykernel package so we can avoid doing imports until

```
Out[92]: KNeighborsClassifier(n_jobs=1, n_neighbors=21)
```

```
In [95]: k_pred=classifier.predict(X_test)  
accuracy=accuracy_score(k_pred,y_test)
```

```
In [97]: accuracy
```

```
Out[97]: 0.7666666666666667
```

```
In [ ]:
```

```
In [ ]:
```