

Weighted distance in sklearn KNN

Weighting the points means that instead of each neighbor class points having an equal vote, their vote is weighted. This means that if for some sample, in 3 nearest-neighbors, 2 of them are of class Type_A but have low weights, and 1 of them is of class Type_B and has a high weight, KNN might still pick class Type_B even though more neighbors are of class Type_A.

Let us consider following Iris flowers classification using KNN classifiers, import iris dataset and use KNeighborsClassifier for classification of iris species.

```
from sklearn.neighbors import
KNeighborsClassifier, kneighbors_graph
from sklearn.model_selection import train_test_split
from sklearn import datasets
knn_classifier = KNeighborsClassifier(n_neighbors=3,
weights='distance')
import pandas as pd
iris = datasets.load_iris()
X = iris.data[:, :2]
y = iris.target
X_train, X_test, y_train, y_test = train_test_split(X, y,
test_size = 0.2, random_state = 0)
knn_classifier = KNeighborsClassifier(n_neighbors=3,
weights='distance')
knn_classifier.fit(X_train, y_train)
y_pred=knn_classifier.predict(X_test)
```

Output:

Hence species are classified into three species like(0,1,2 resp.)

```
array([1, 2, 0, 2, 0, 1, 0, 1, 2, 1, 1, 2, 1, 2, 2, 0, 2, 1, 0, 0, 1, 1,
       0, 0, 2, 0, 0, 2, 1, 0])
```

Parameters:

n_neighbors=3 (default=5)

Number of neighbors to use by default for kneighbors queries.

‘distance’ : weight points by the inverse of their distance. In this case, closer neighbors of a query point will have a greater influence than neighbors which are further away.

```
A = kneighbors_graph(X=X_test, n_neighbors=3,
mode='connectivity', include_self=True)
print("Computes the (weighted) graph of k-Neighbors for points
in X_test, which predicted by knn_classifier")
A.toarray()
```

Computes the weights of k-Neighbors for points for testing dataset, which predicted by KNN Classifier as follows,

[illegible]