

K Nearest Neighbors

Imports

```
In [1]: import numpy as np
import matplotlib.pyplot as plt
from matplotlib.colors import ListedColormap
from sklearn import neighbors, datasets
from sklearn.neighbors import KNeighborsClassifier
import numpy as np
from sklearn import datasets
from KNN_iris import KNN_plot
```

Loading the dataset

```
In [2]: # Loading the dataset
iris = datasets.load_iris()
iris_X = iris.data
iris_y = iris.target

# Overlook on the dataset
print "Number of features: ", iris_X.shape[1]
print "Labels: ", np.unique(iris_y)
```

```
Number of features:  4
Labels:  [0 1 2]
```

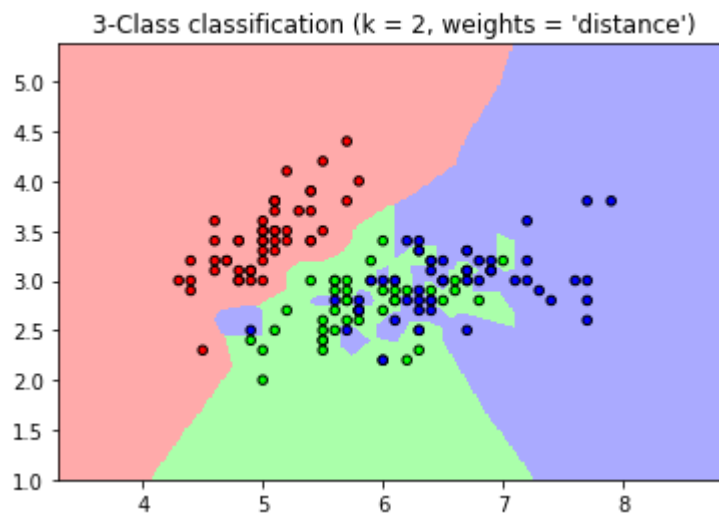
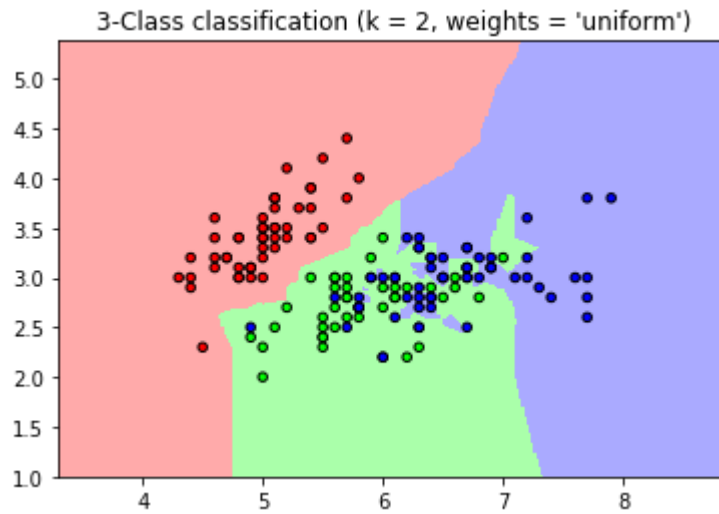
Changing the metric

Metric: The metric is the definition of distance used by the algorithm.

In KNN, when metric = 'distance', neighbors are weighted 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.

```
In [18]: metrics = ['uniform', 'distance']  
n_neighbors = 2  
  
for m in metrics:  
    KNN_plot(n_neighbors, metric = m)
```



Interpretation: When distance is used as a metric, singular points have a stronger influence.

Changing the projection space

In [19]: `X = [0, 1, 2, 3]`

```
n_neighbors = 10  
metric = 'uniform'
```

```
for x in X:  
    for y in X:  
        if y>x:  
            KNN_plot(n_neighbors, x1 = x, x2 = y, metric = m)
```

