K-Nearest Neighbors (KNN) Algorithm

Dr. Muhammad Wasim

K-nearest neighbors (KNN) algorithm

- **Idea:** If we can measure the distance of the unseen example with the examples in the training data, we can predict label for the unseen example.

K-nearest neighbors (KNN) algorithm for Classification

Training Dataset

Sepal Length	Sepal Width	Petal Length	Petal Width	Class / Label
5.1	3.5	1.4	0.2	Iris setosa
4.9	3.0	1.4	0.2	Iris setosa
7.0	3.2	4.7	1.4	Iris versicolor
6.4	3.2	4.5	1.5	Iris versicolor
6.3	3.3	6.0	2.5	Iris virginica
5.8	3.3	6.0	2.5	Iris viginica

Unseen Example

Sepal Length	Sepal Width	Petal Length	Petal Width	Class / Label
5.2	2.9	6.0	2.3	?

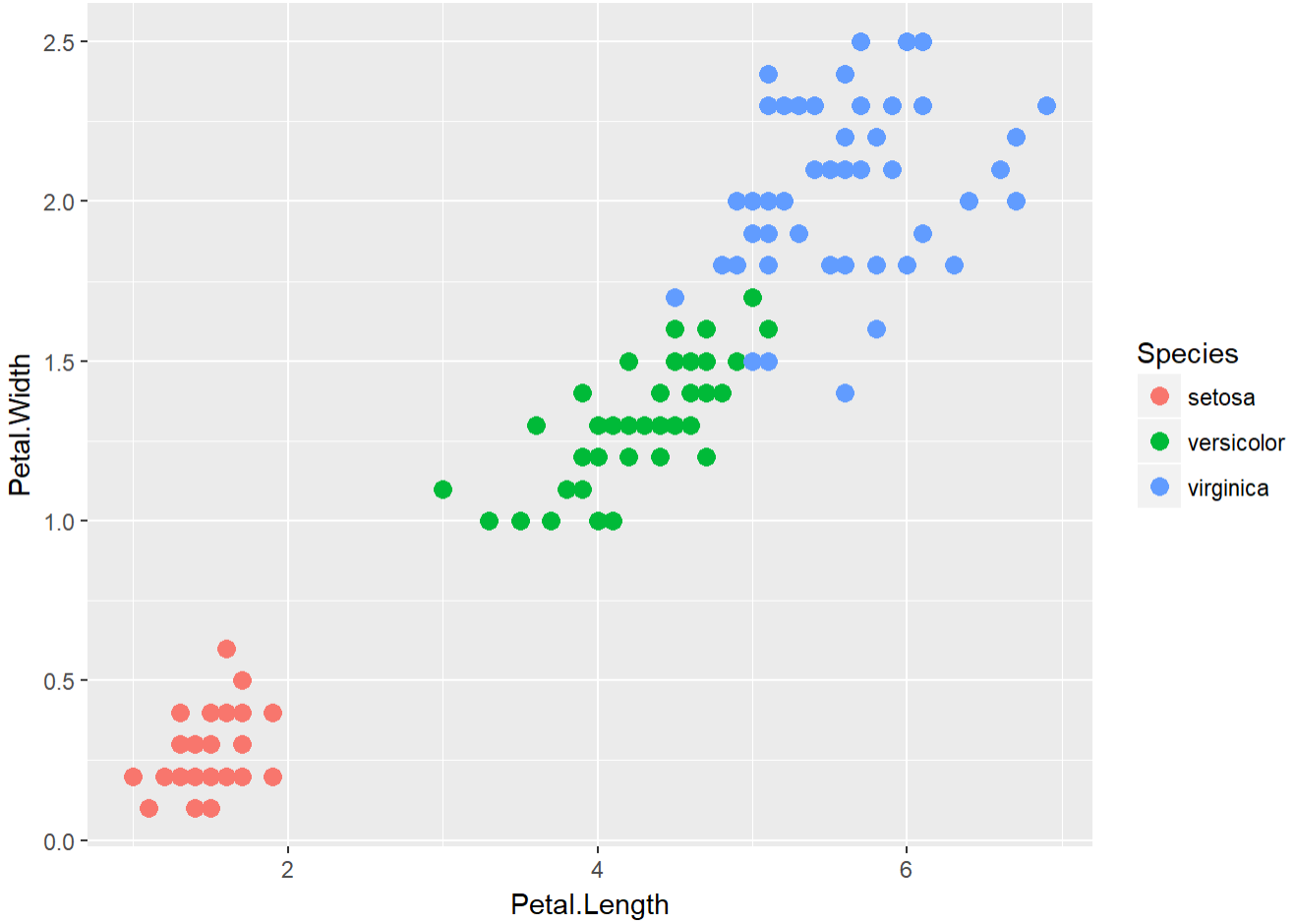
K-nearest neighbors (KNN) algorithm for Classification

Training Dataset

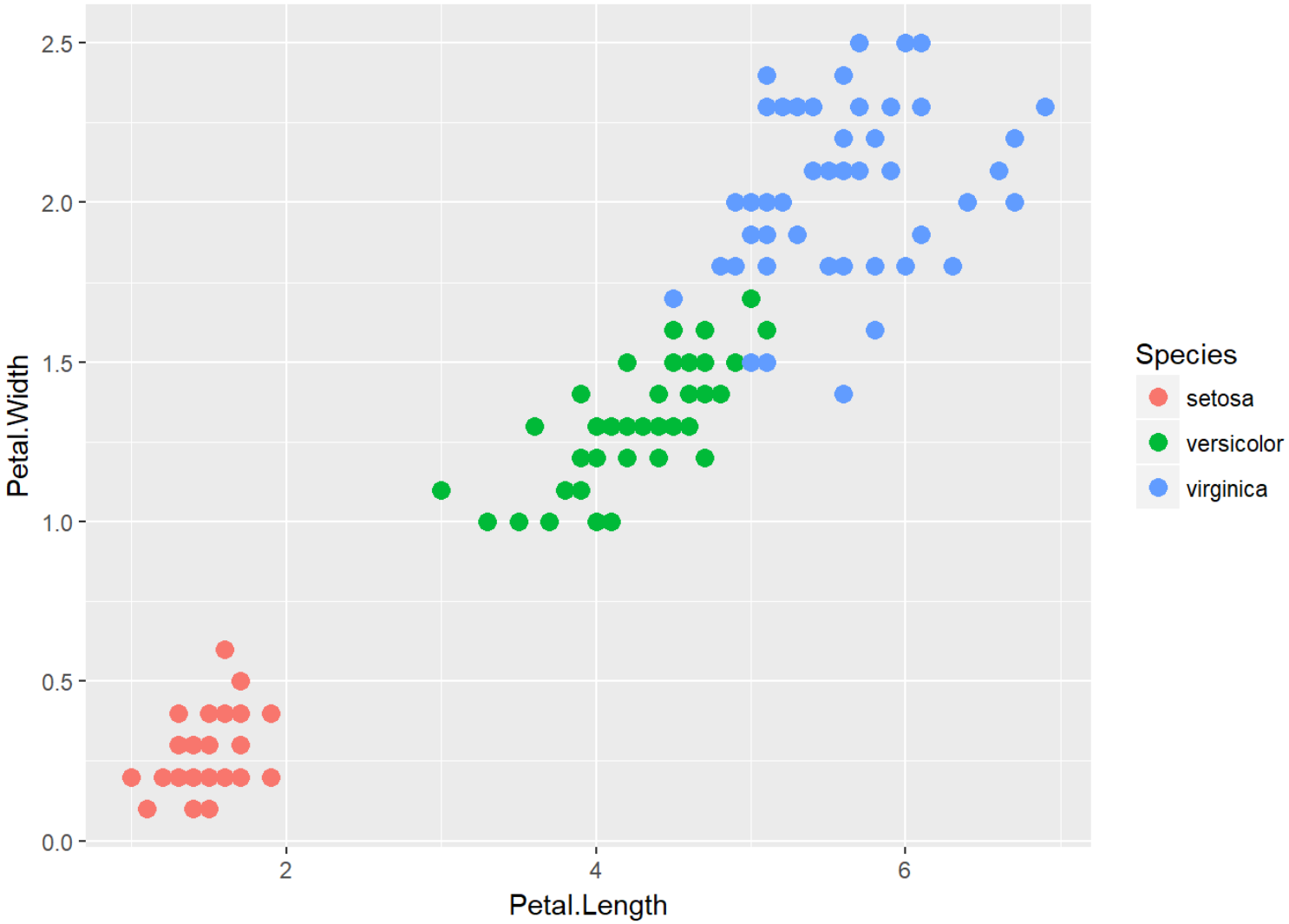
Petal Length	Petal Width	Class / Label
1.4	0.2	Iris setosa
1.4	0.2	Iris setosa
4.7	1.4	Iris versicolor
4.5	1.5	Iris versicolor
6.0	2.5	Iris virginica
6.0	2.5	Iris viginica

Unseen Example

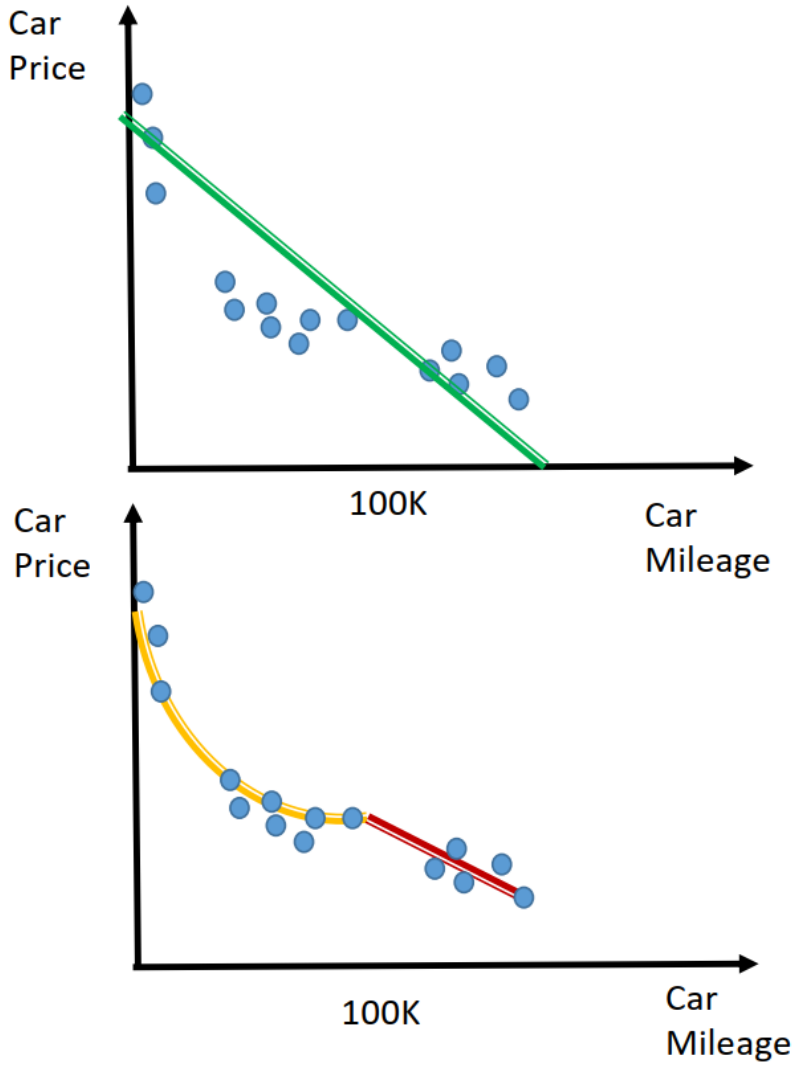
Petal Length	Petal Width	Class / Label
6.0	2.3	?



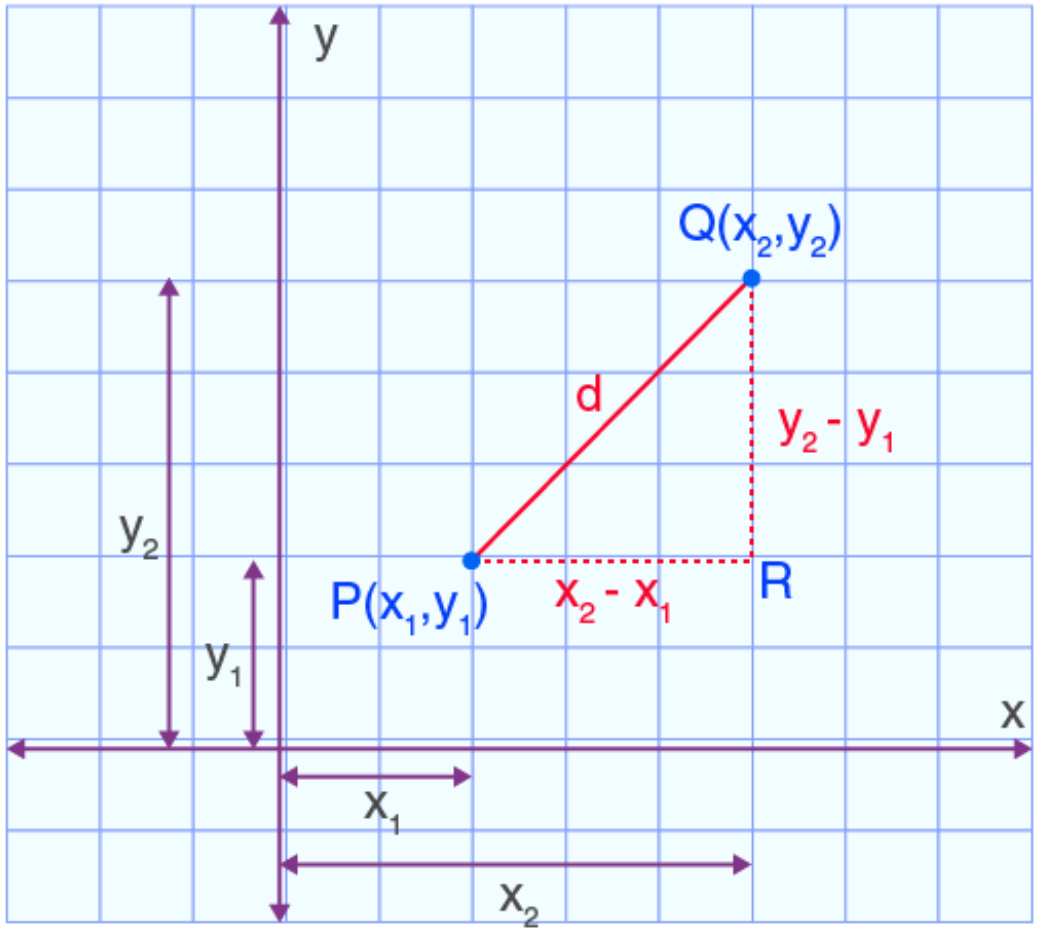
K-nearest neighbors (KNN) algorithm for Classification



K-nearest neighbors (KNN) algorithm for Regression



Distance Measure: Euclidean Distance



$$d(P,Q) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

Similarity Measures

- Minkowski Distance

$$\left(\sum_{i=1}^{n} \left| x_i - y_i \right|^p\right)^{1/p}$$

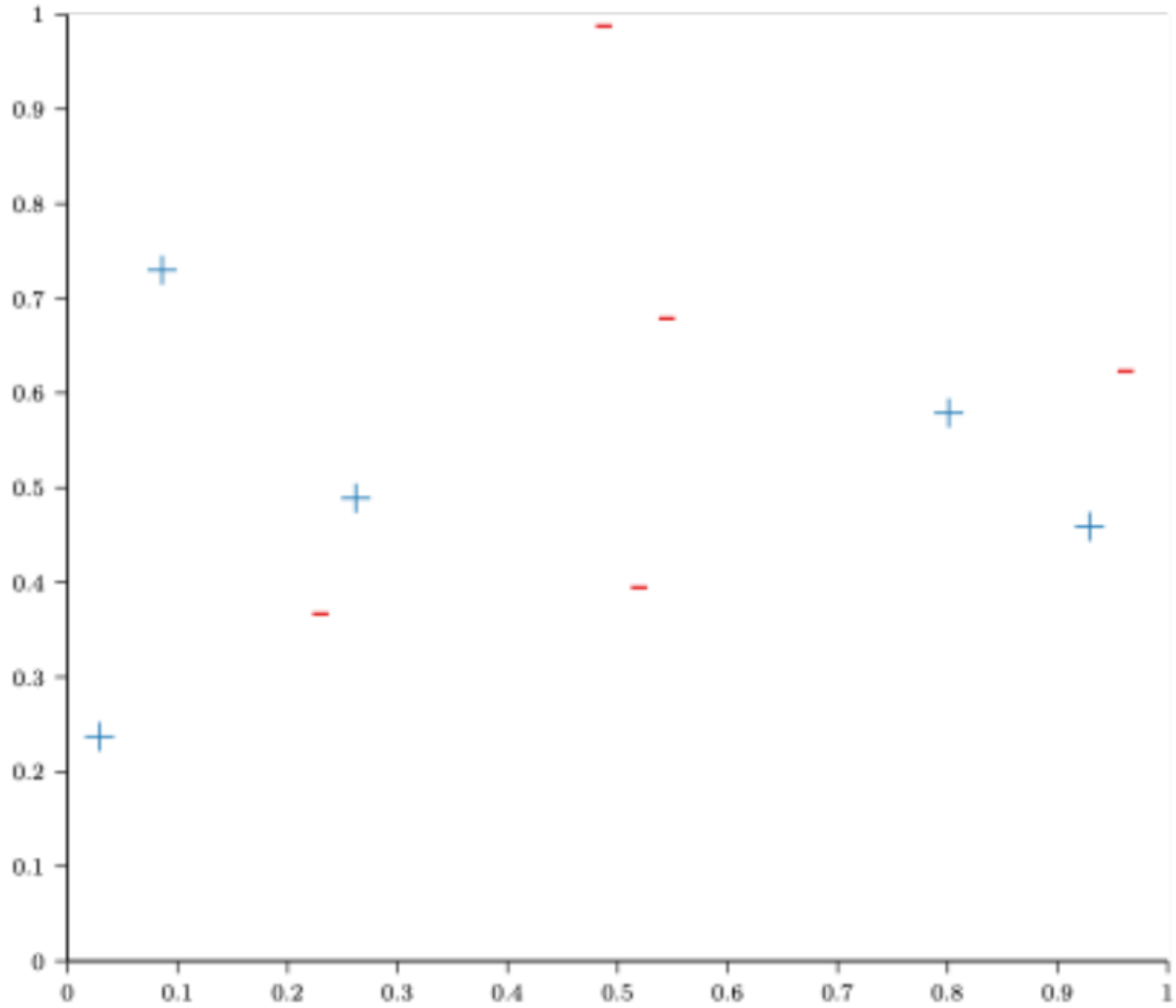
- Manhattan Distance (p=1)

$$\left(\sum_{i=1}^{n} \left| x_i - y_i \right|\right)$$

- Euclidean Distance (p=2)

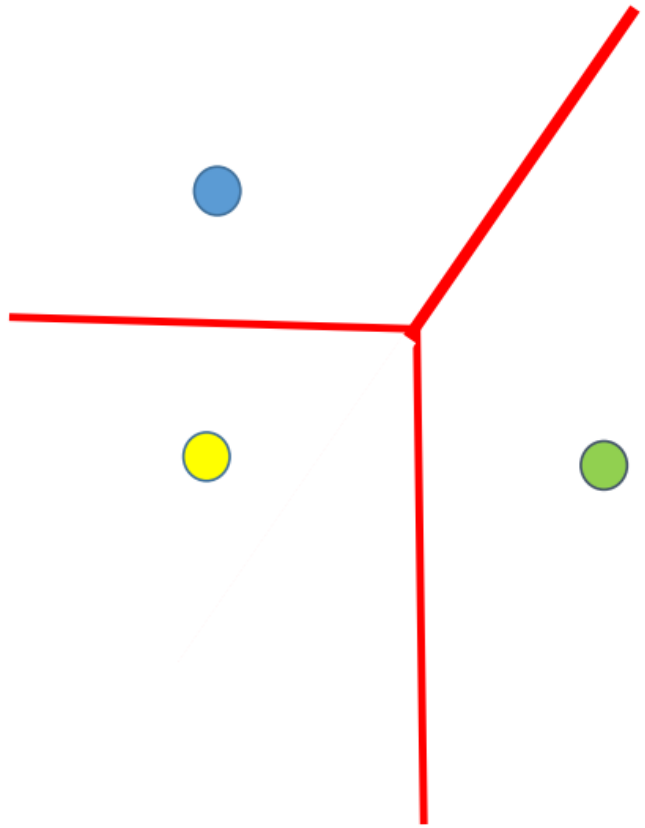
$$\left(\sum_{i=1}^{n} (x_i - y_i)^2\right)^{1/2}$$

Nearest Neighbor: Example



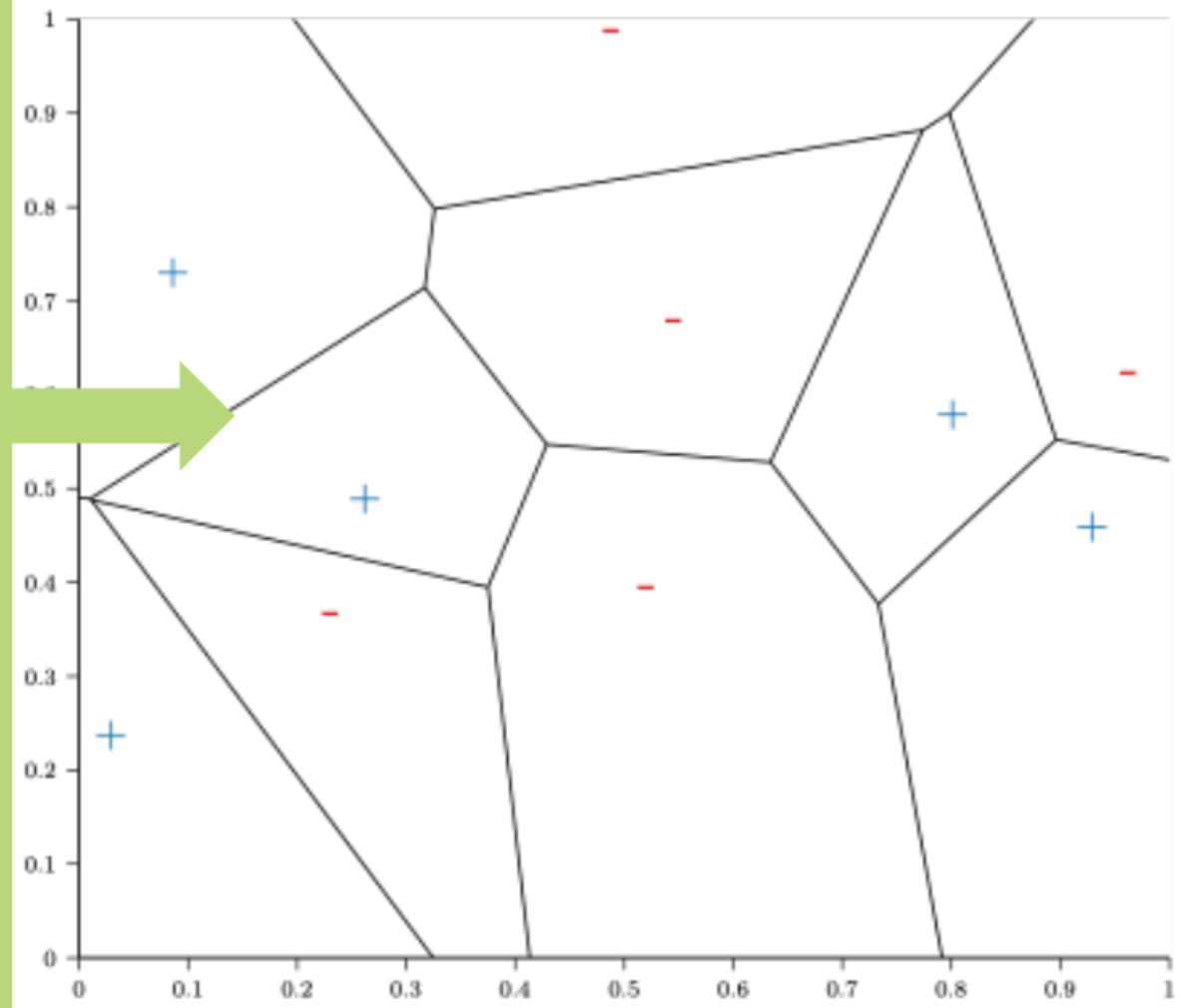
Voronoi Diagram

- Voronoi Cell of x: All points in a Voronoi cell are closer to x than to any other instance

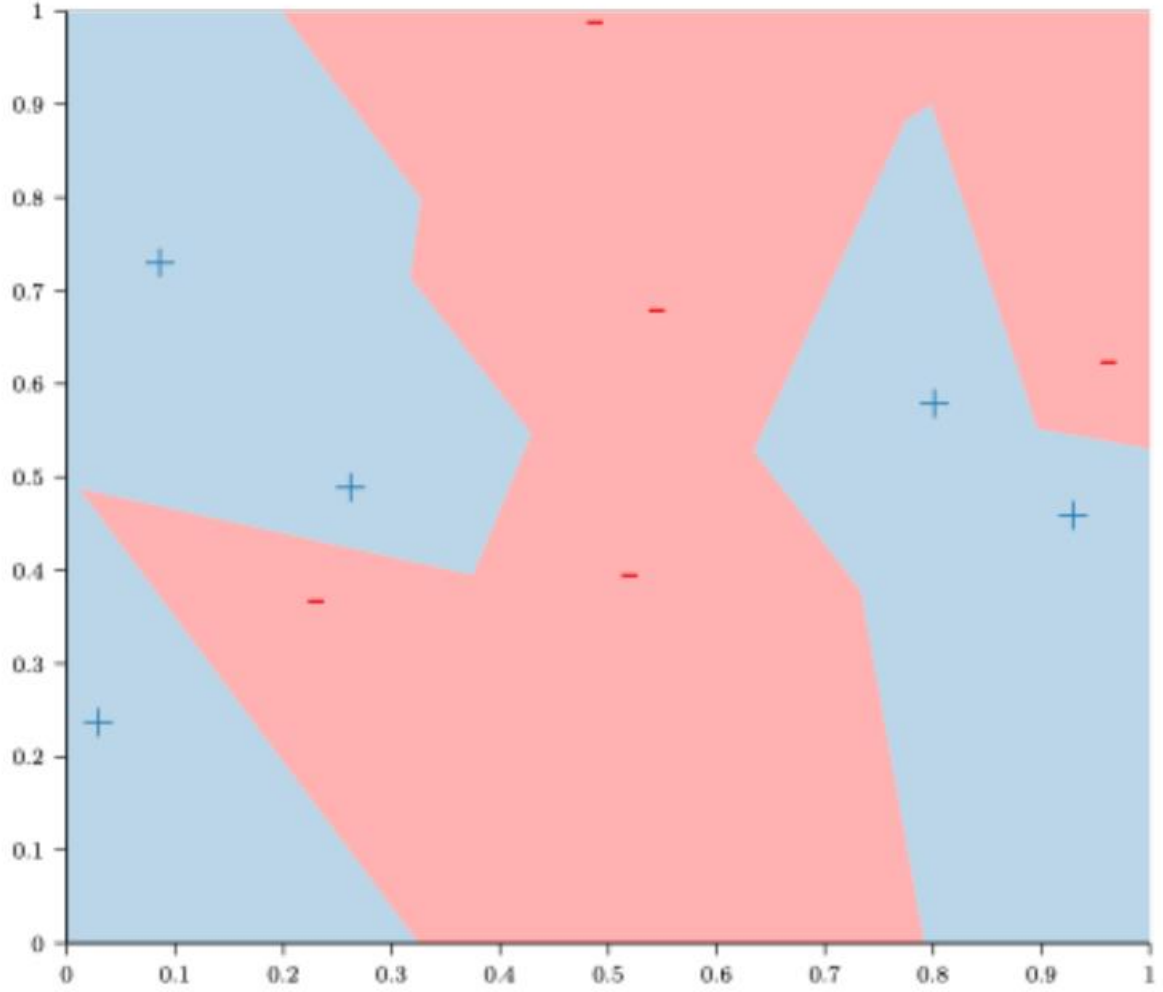


Nearest Neighbor: Example

- This is a **Voronoi diagram**
- Each **cell** contain one of our training examples
- **All points within a cell** are **closer** to that training example, than to any other training example
- **Points on the Voronoi line segments** are **equidistant** to one or more training examples

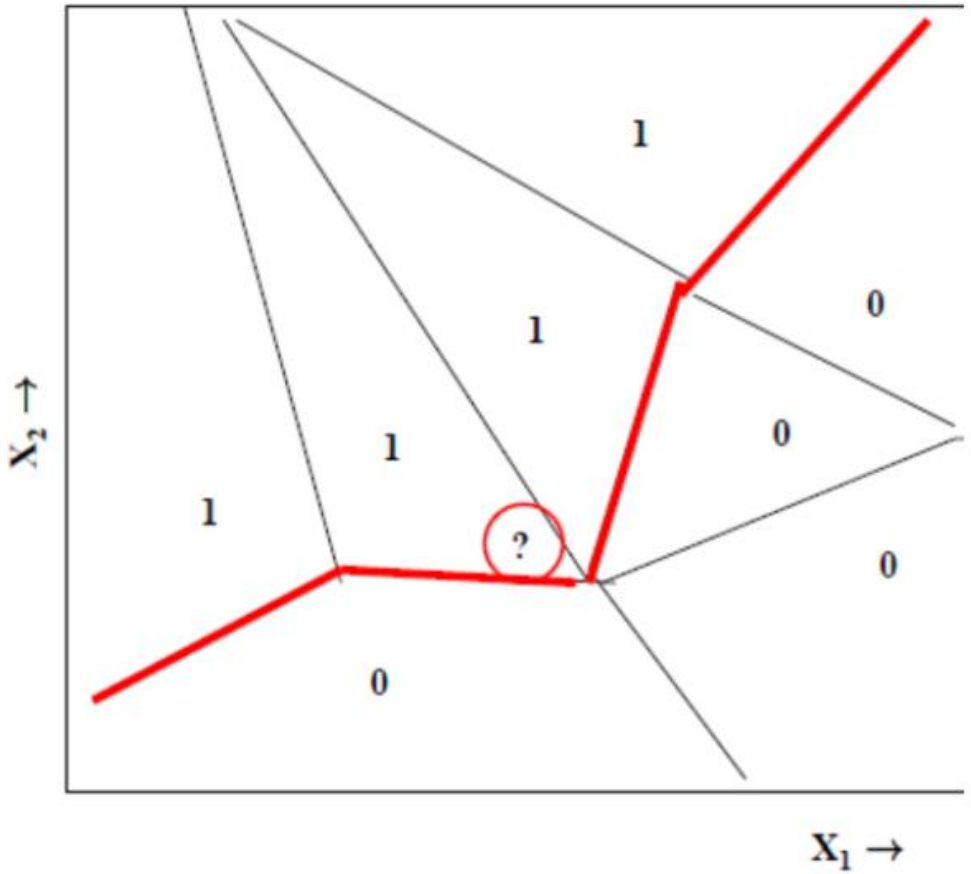


Nearest Neighbor: Example

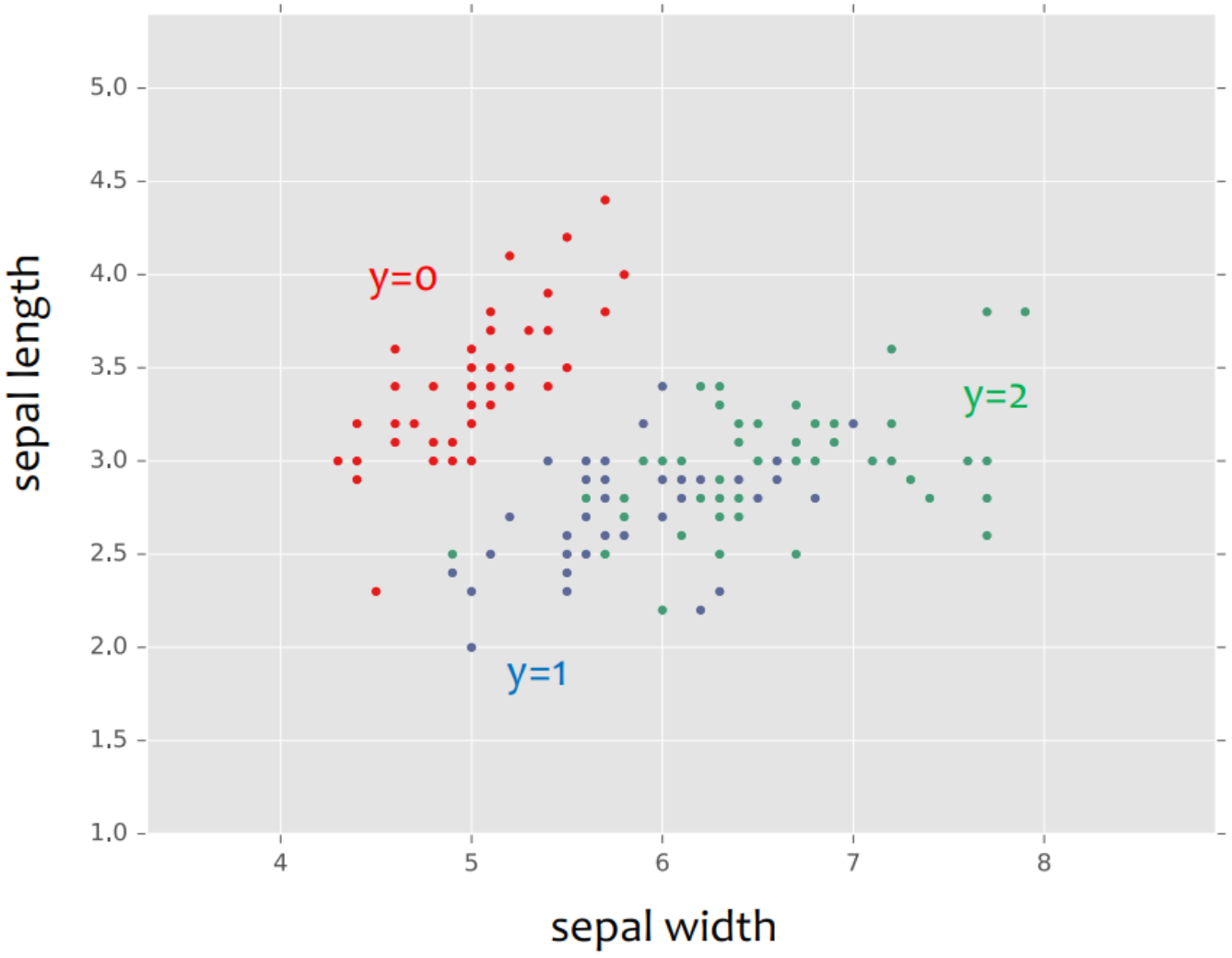


Decision Boundary

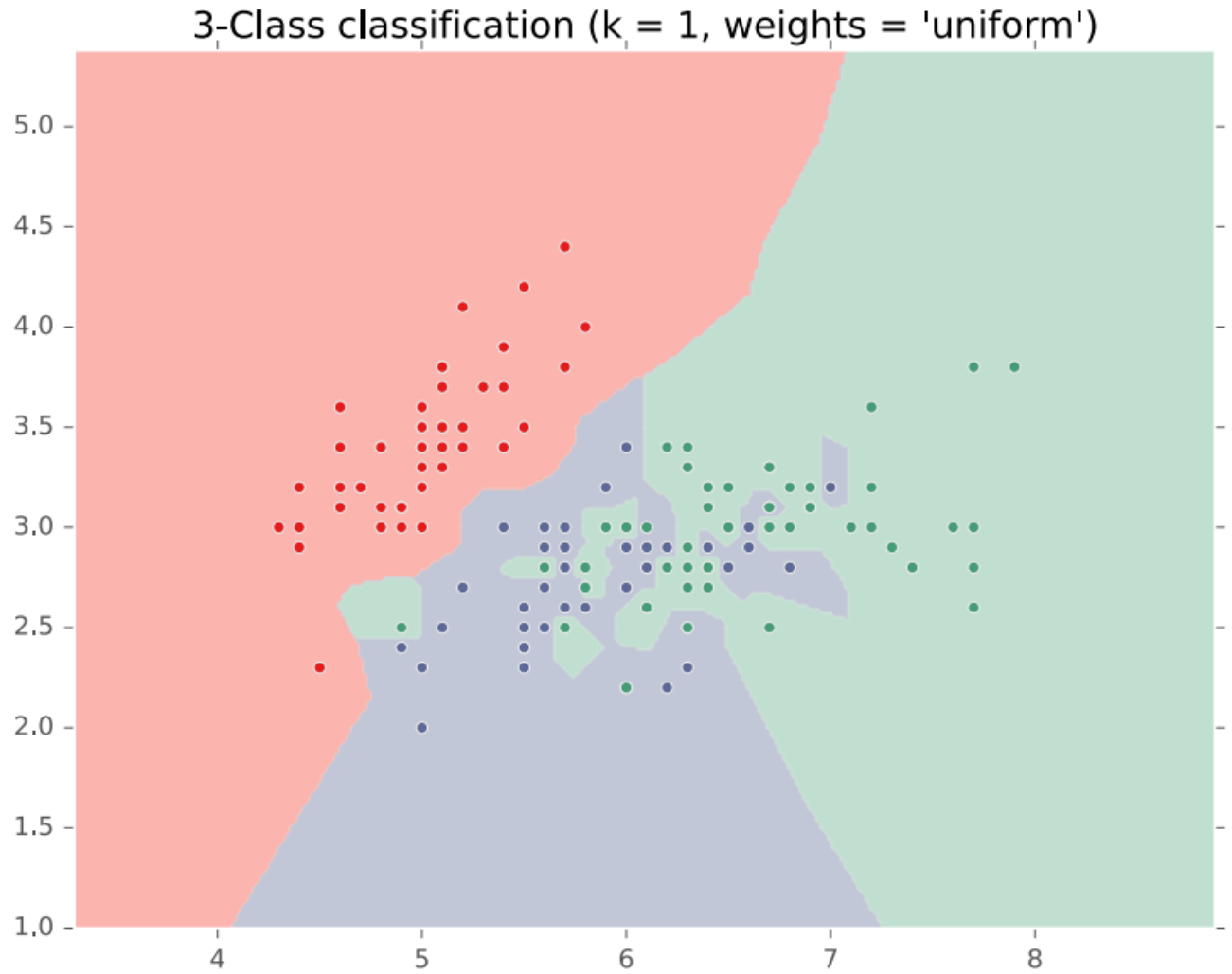
- Those edges across which the decision changes



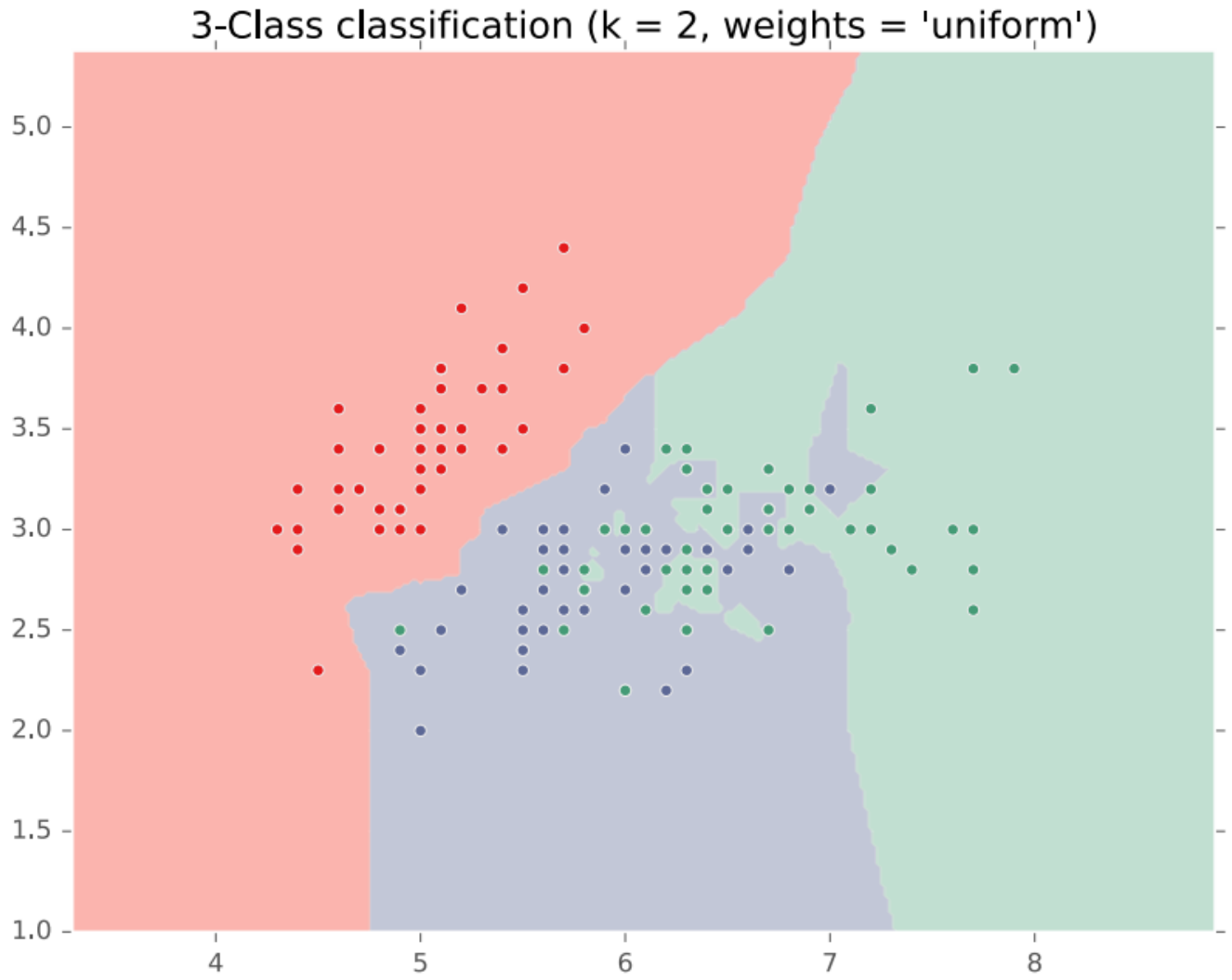
KNN on Fisher Iris Dataset



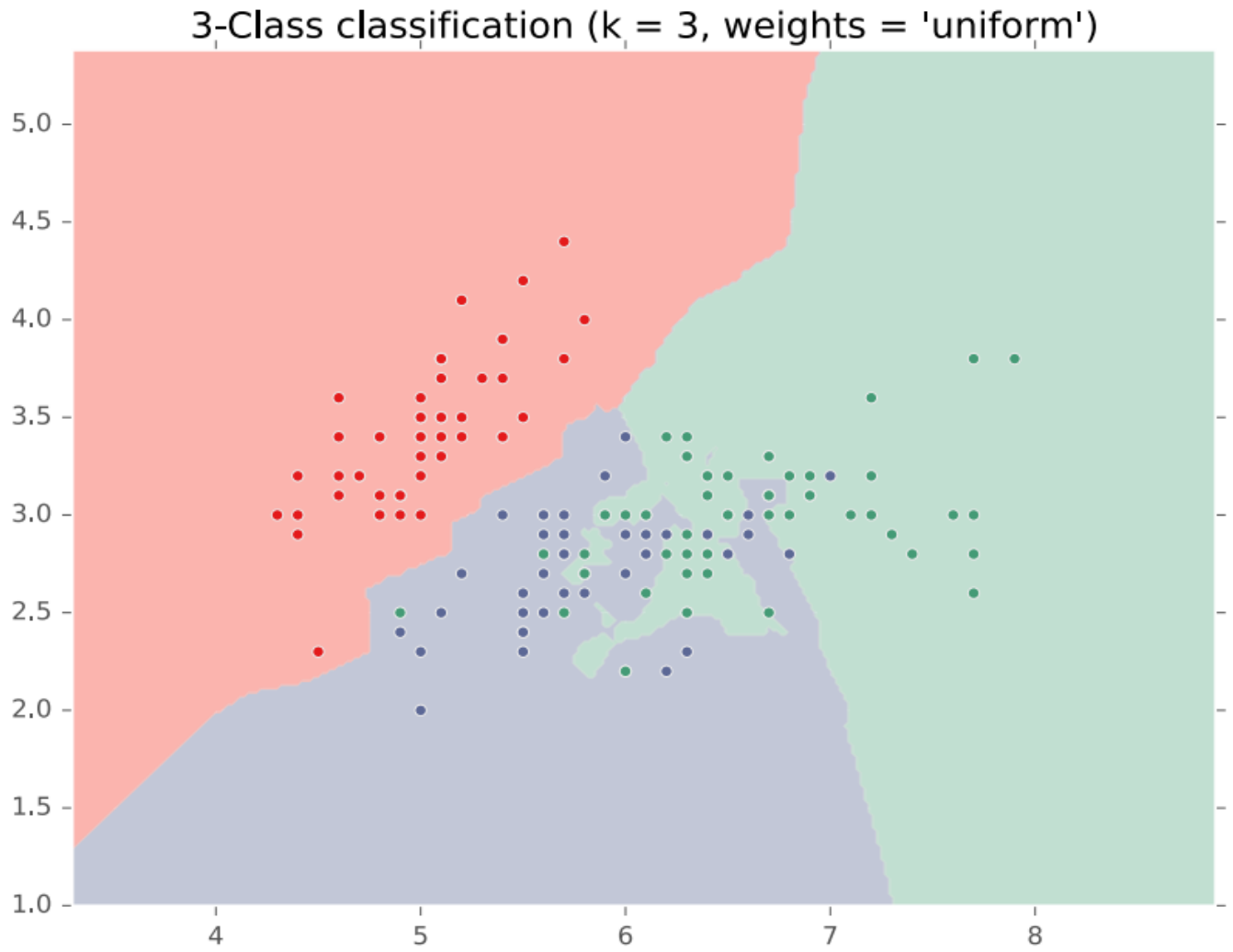
KNN on Fisher Iris Dataset



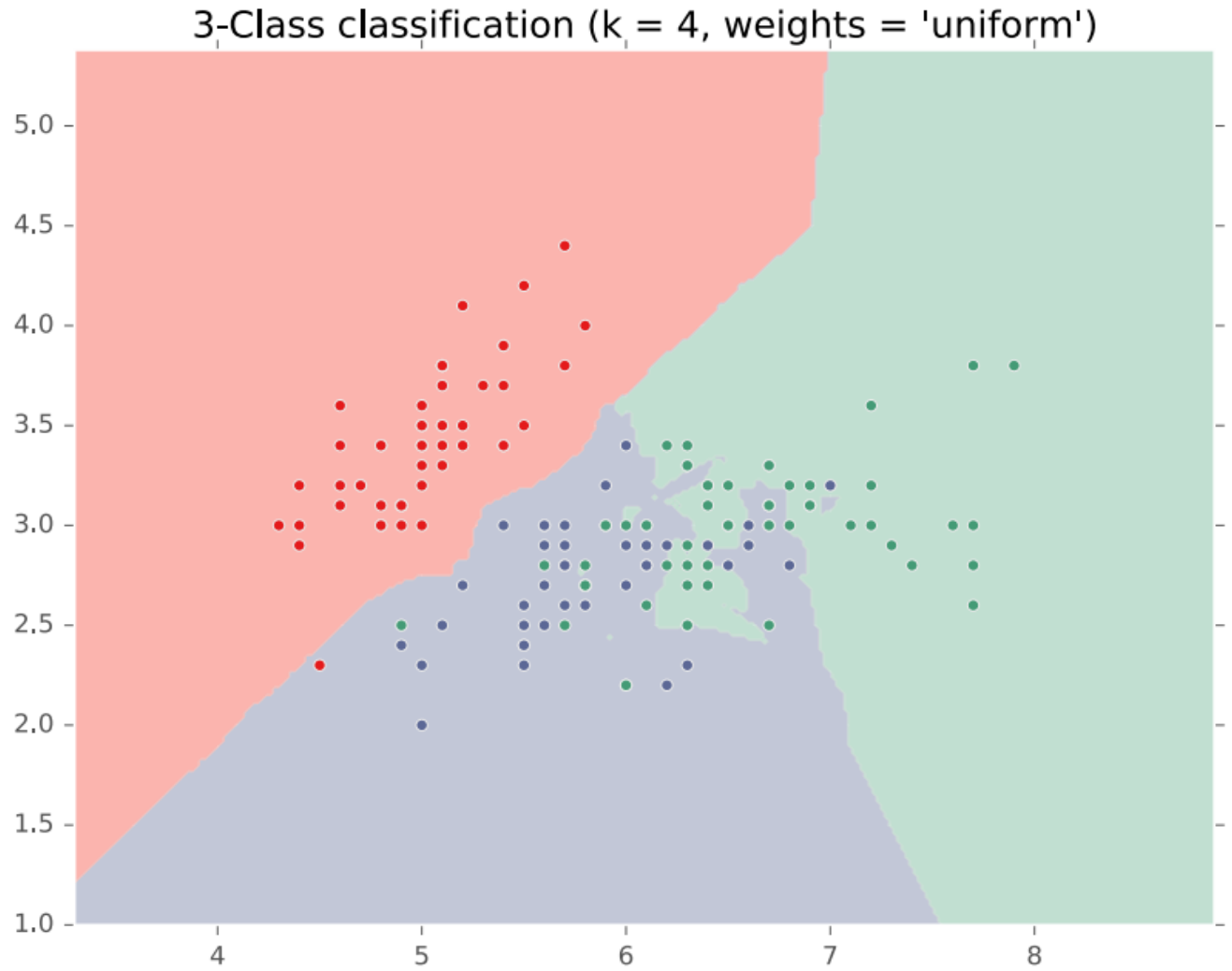
KNN on Fisher Iris Dataset



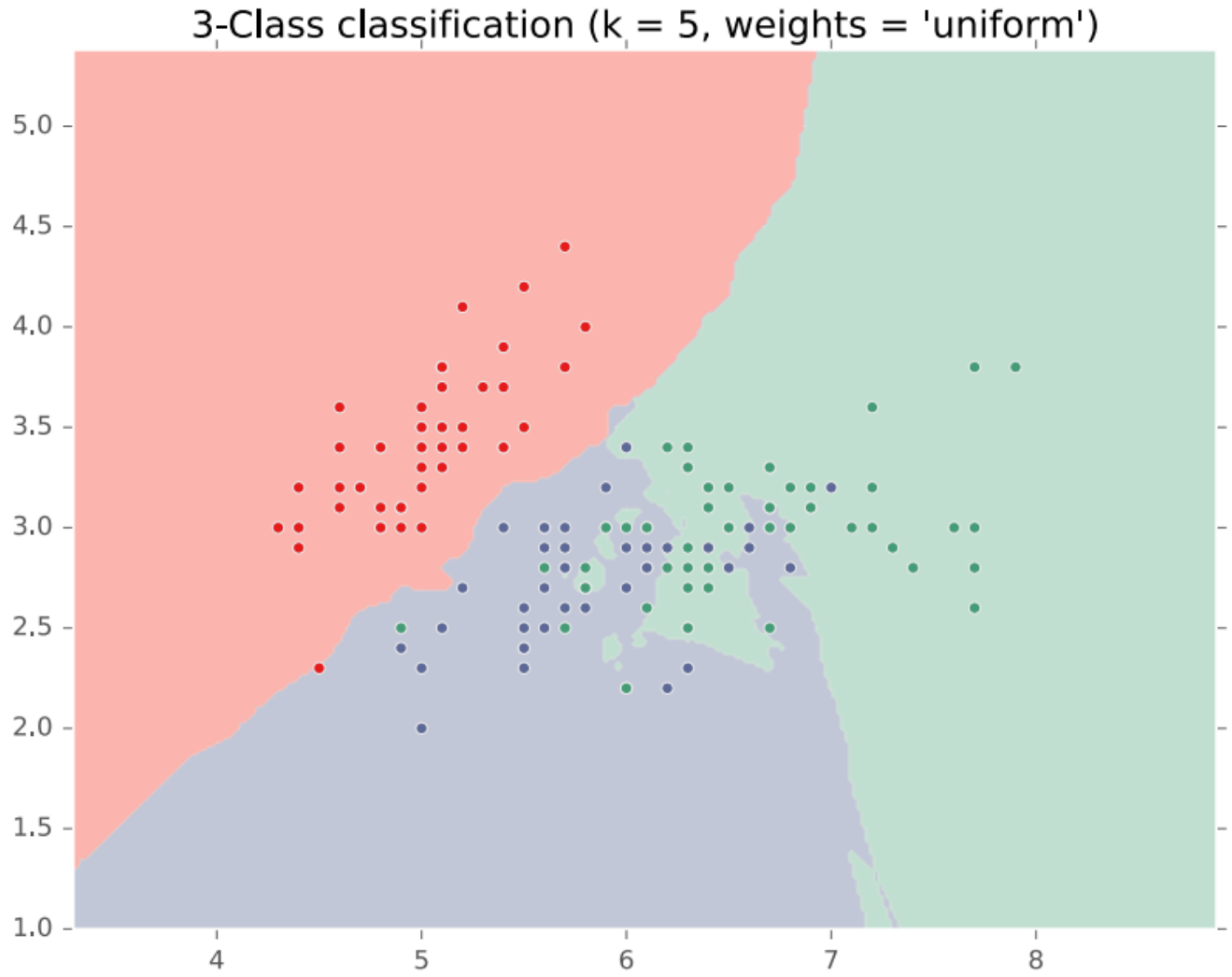
KNN on Fisher Iris Dataset



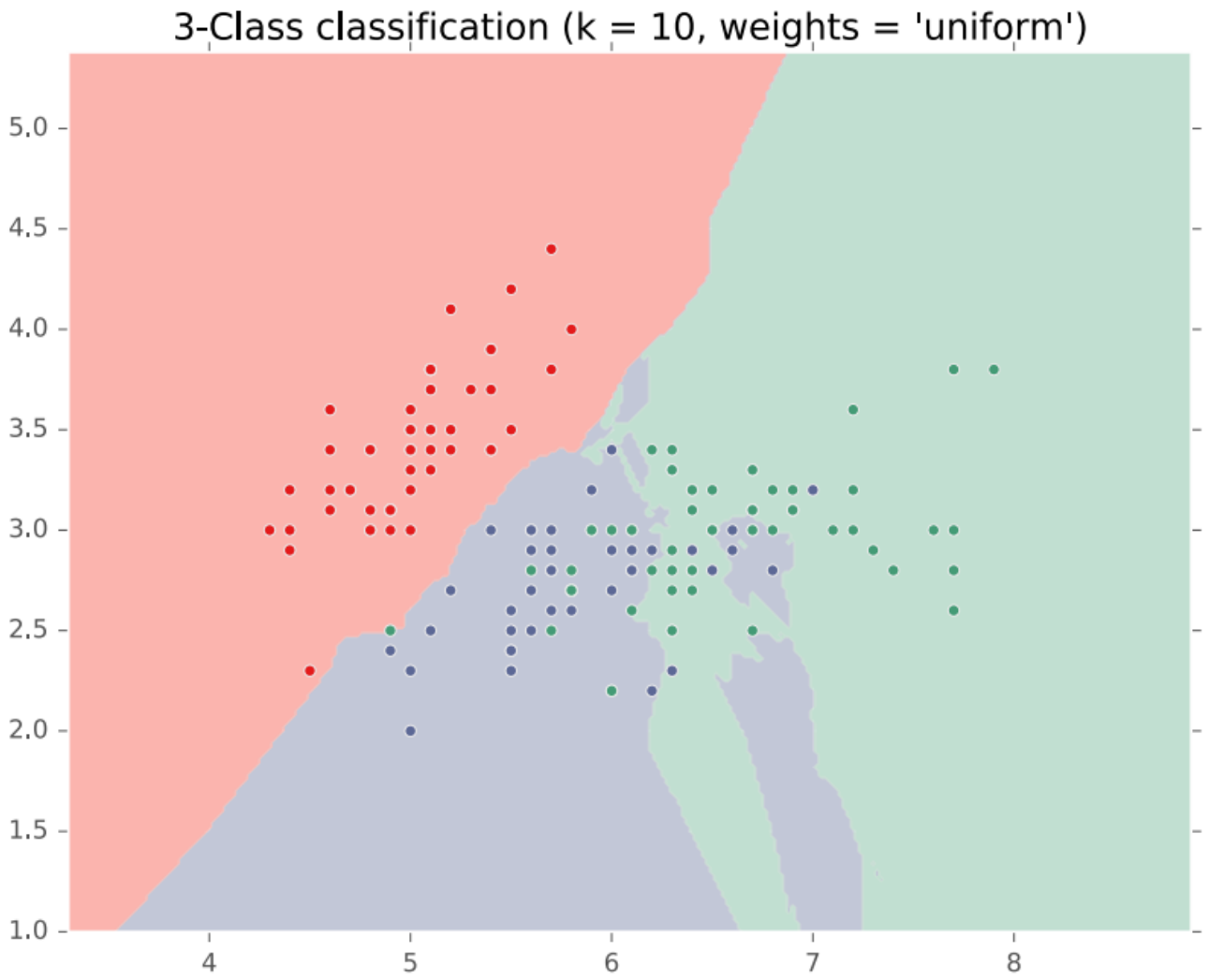
KNN on Fisher Iris Dataset



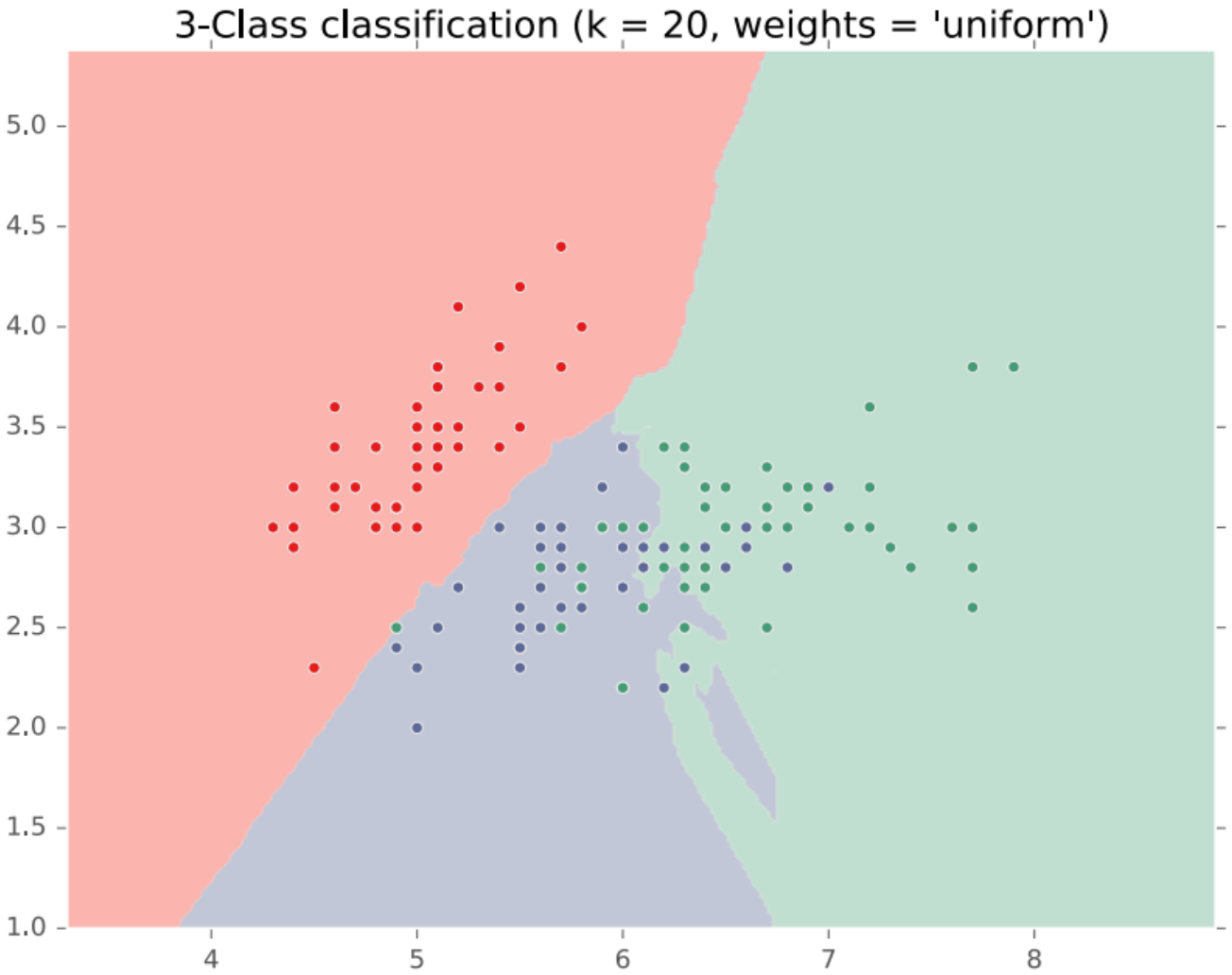
KNN on Fisher Iris Dataset



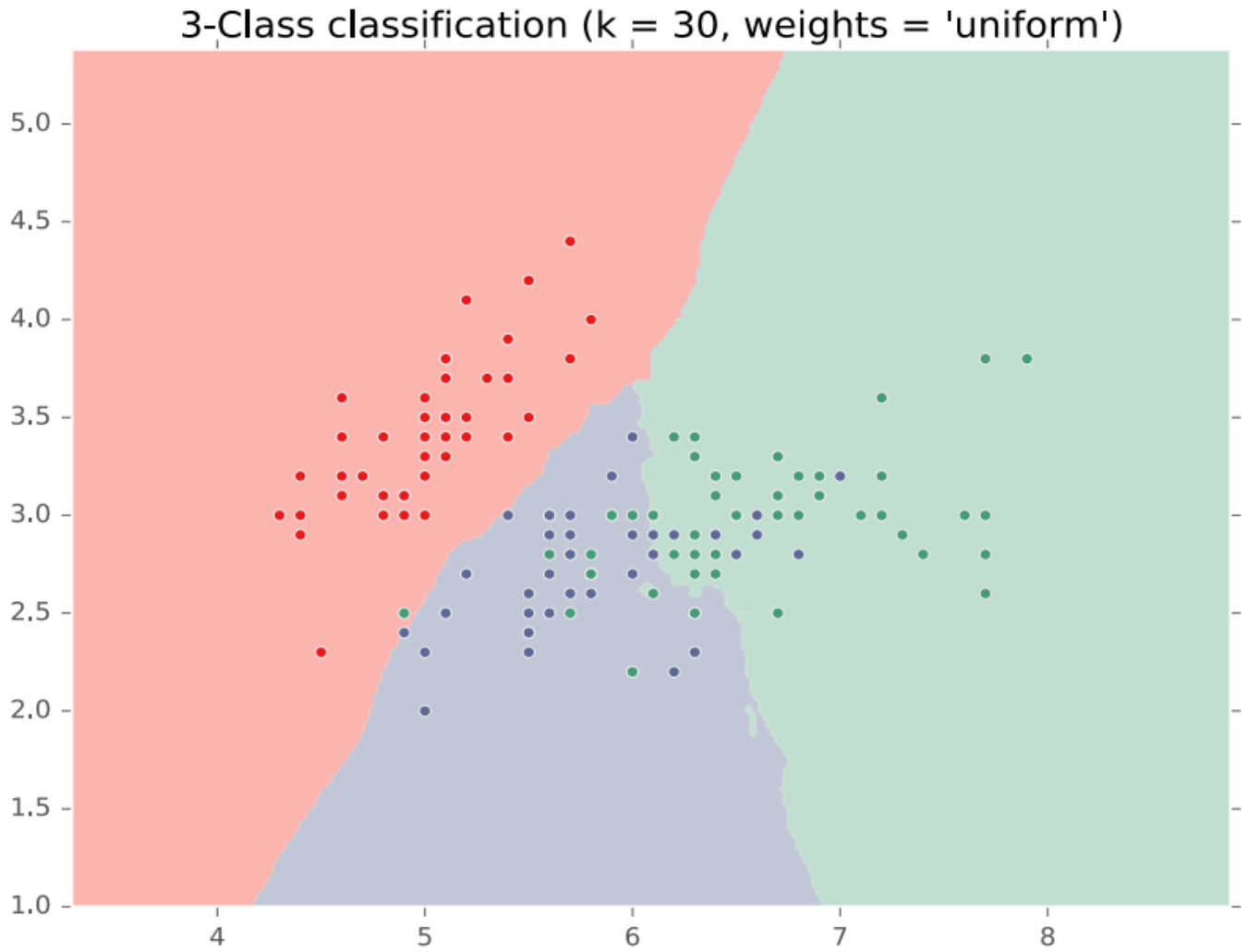
KNN on Fisher Iris Dataset



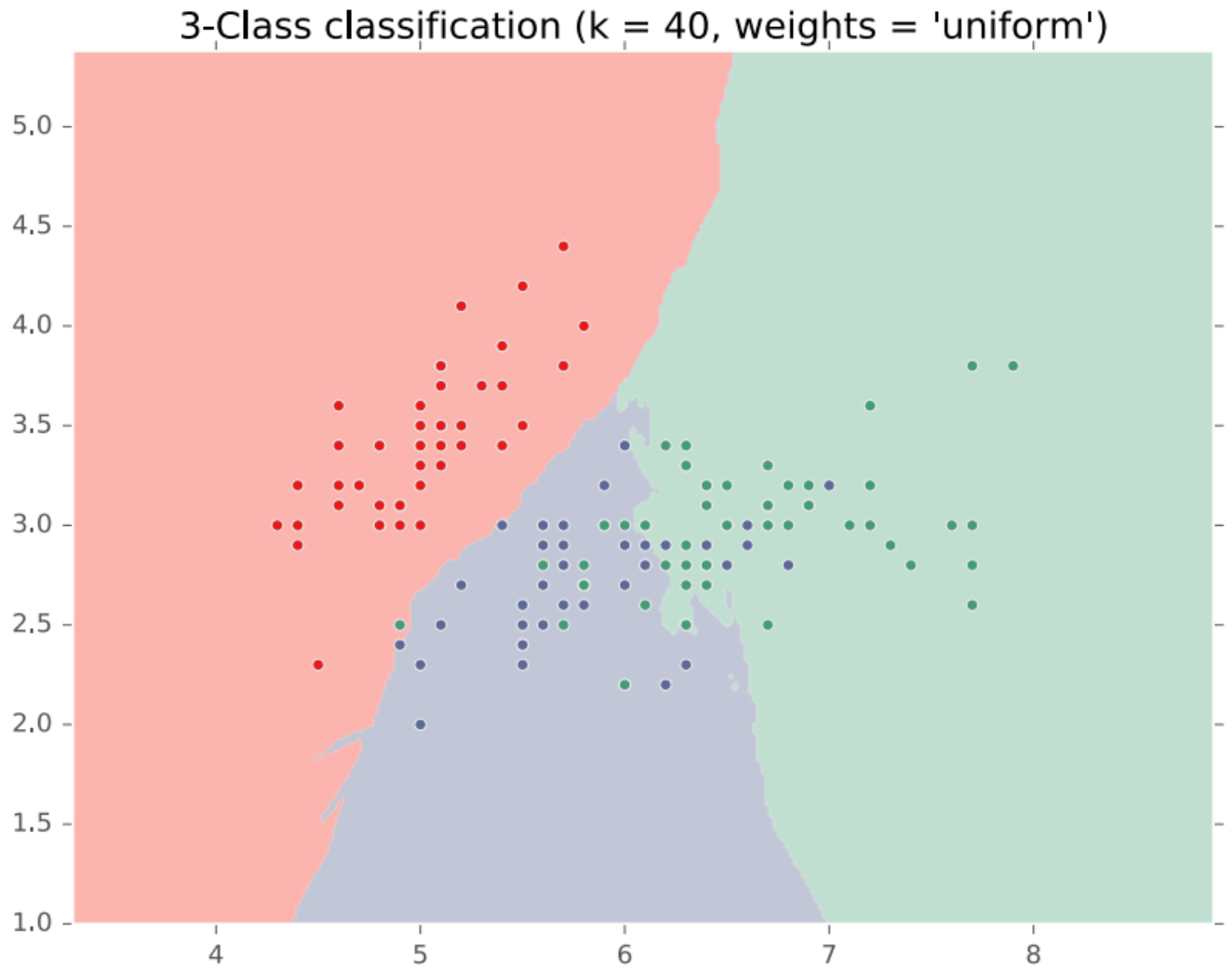
KNN on Fisher Iris Dataset



KNN on Fisher Iris Dataset

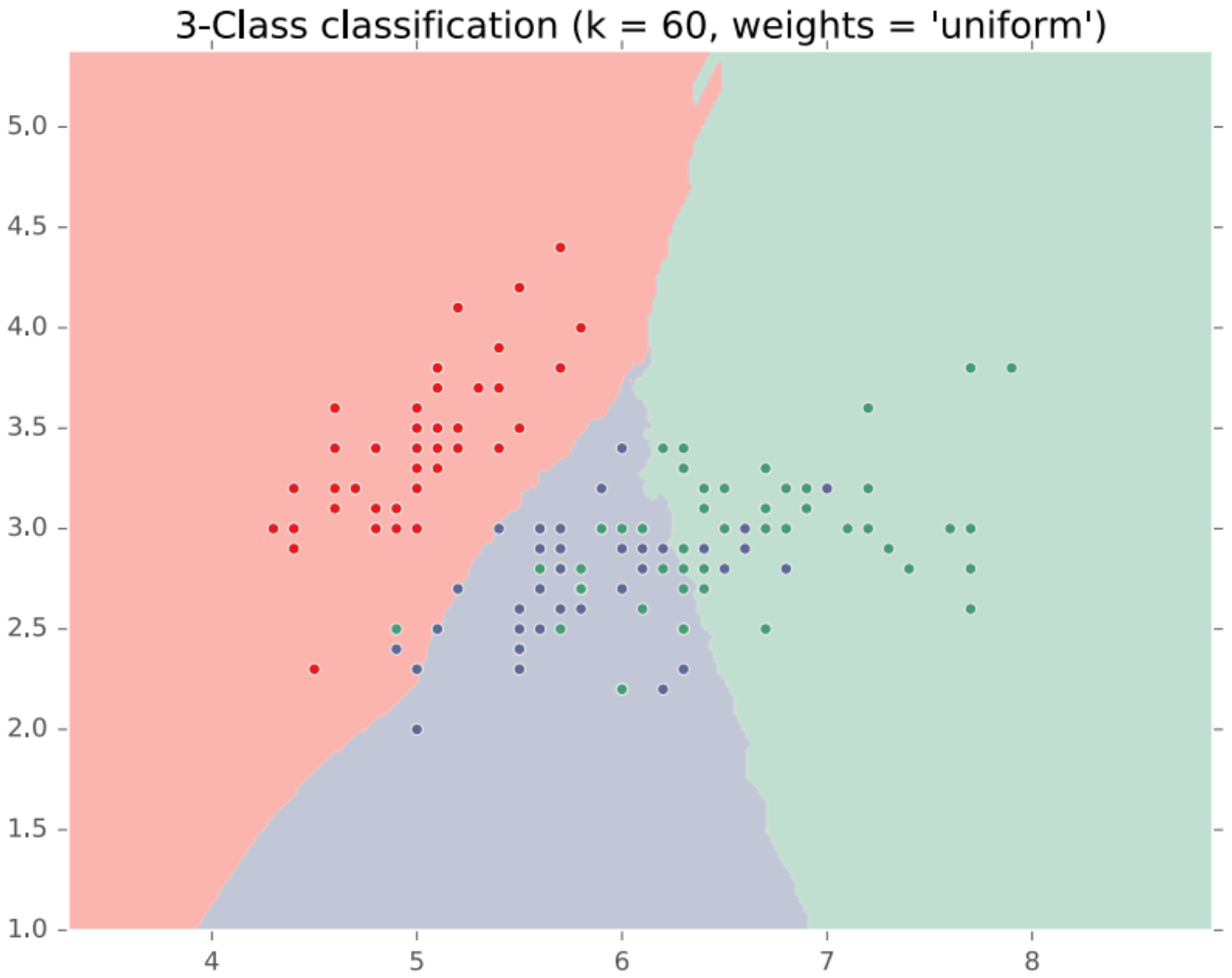


KNN on Fisher Iris Dataset

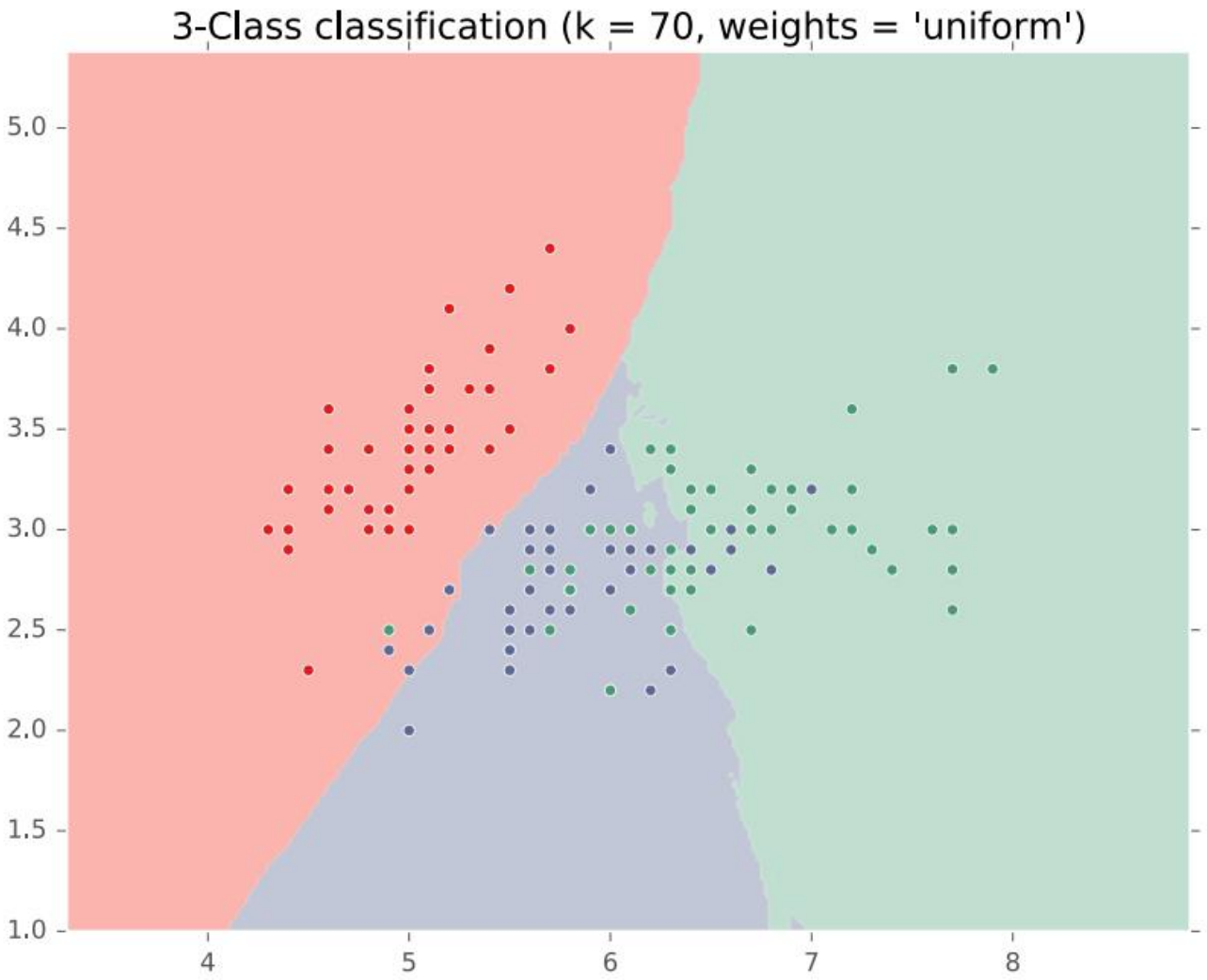


KNN on Fisher Iris Dataset

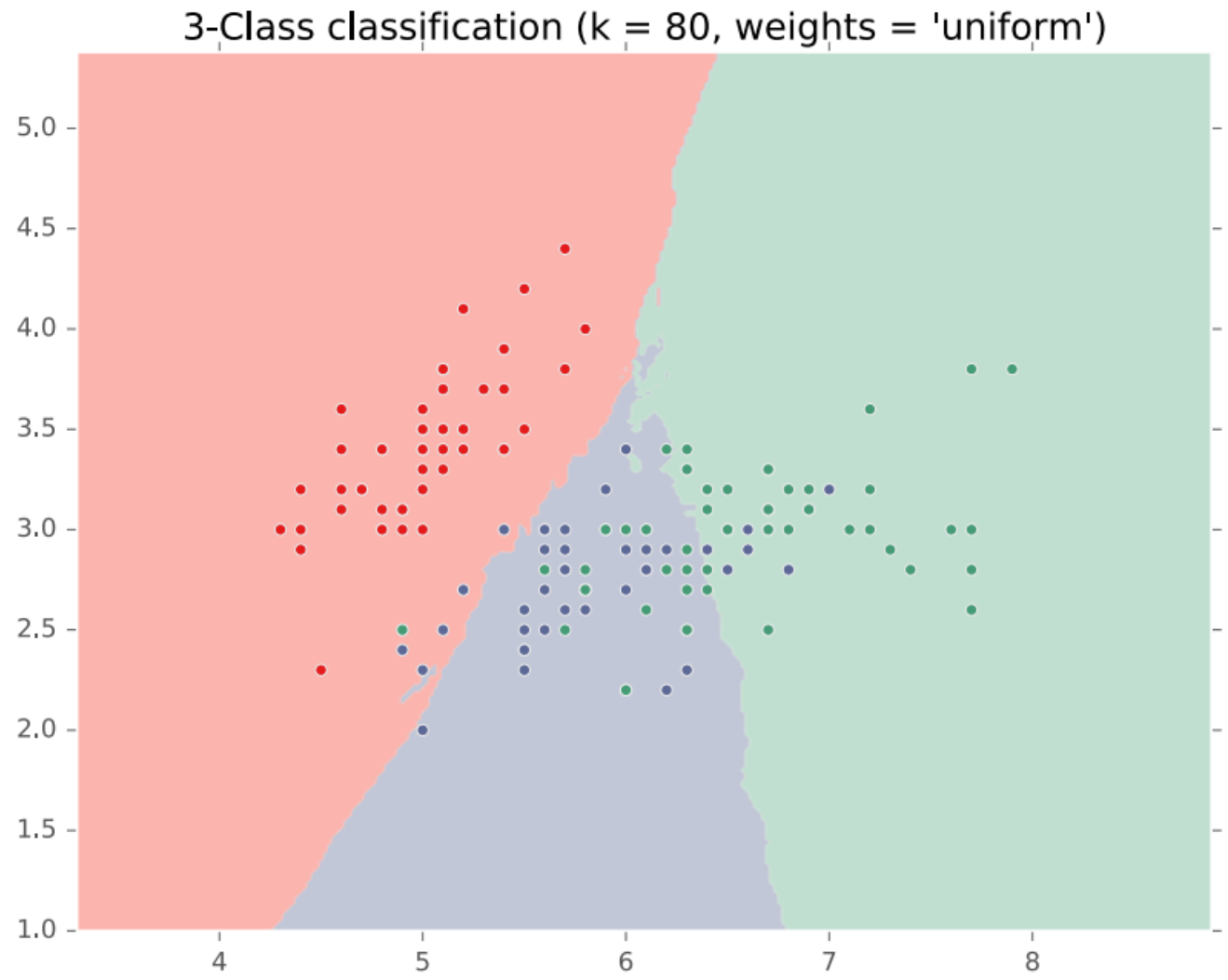
KNN on Fisher Iris Dataset



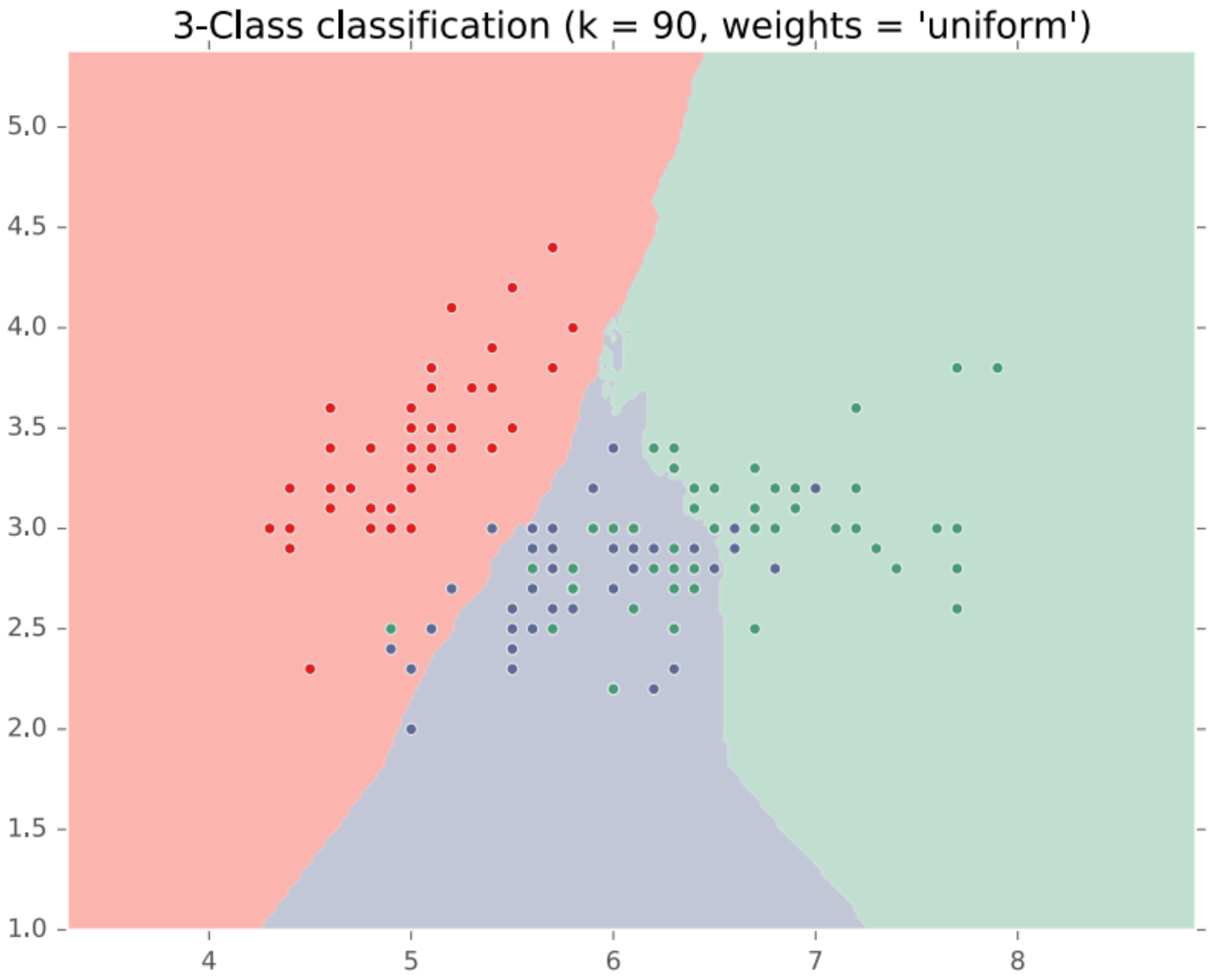
KNN on Fisher Iris Dataset



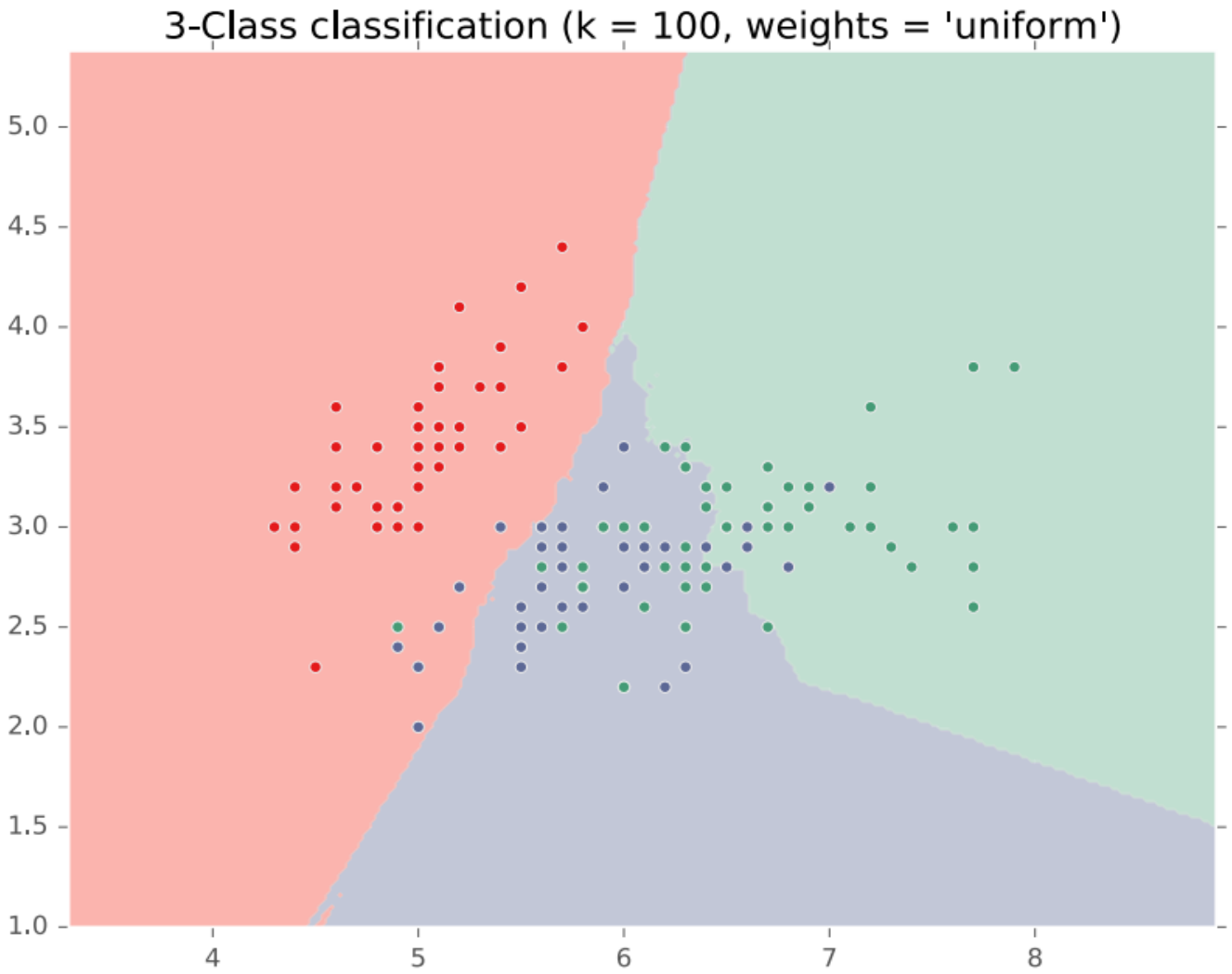
KNN on Fisher Iris Dataset



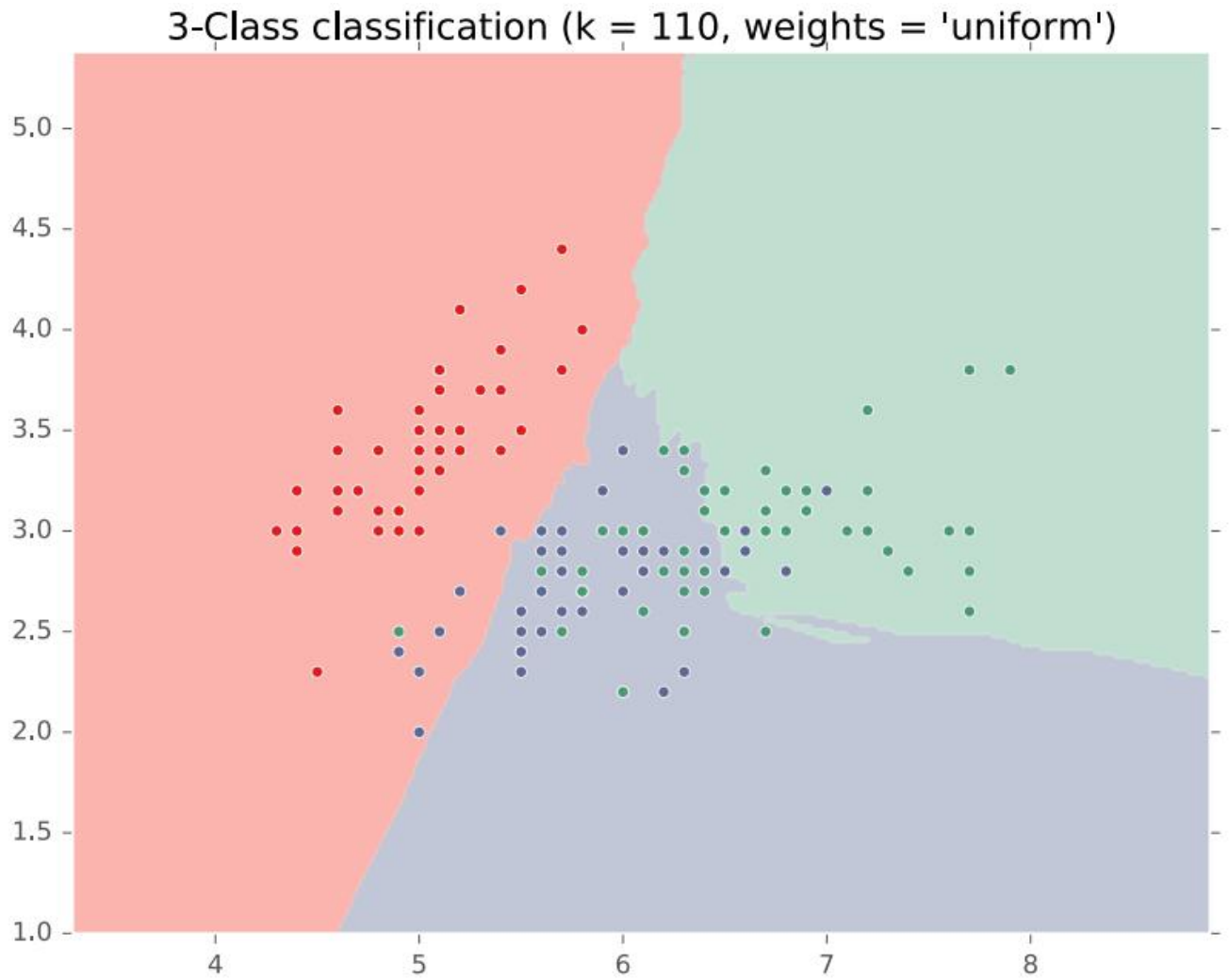
KNN on Fisher Iris Dataset



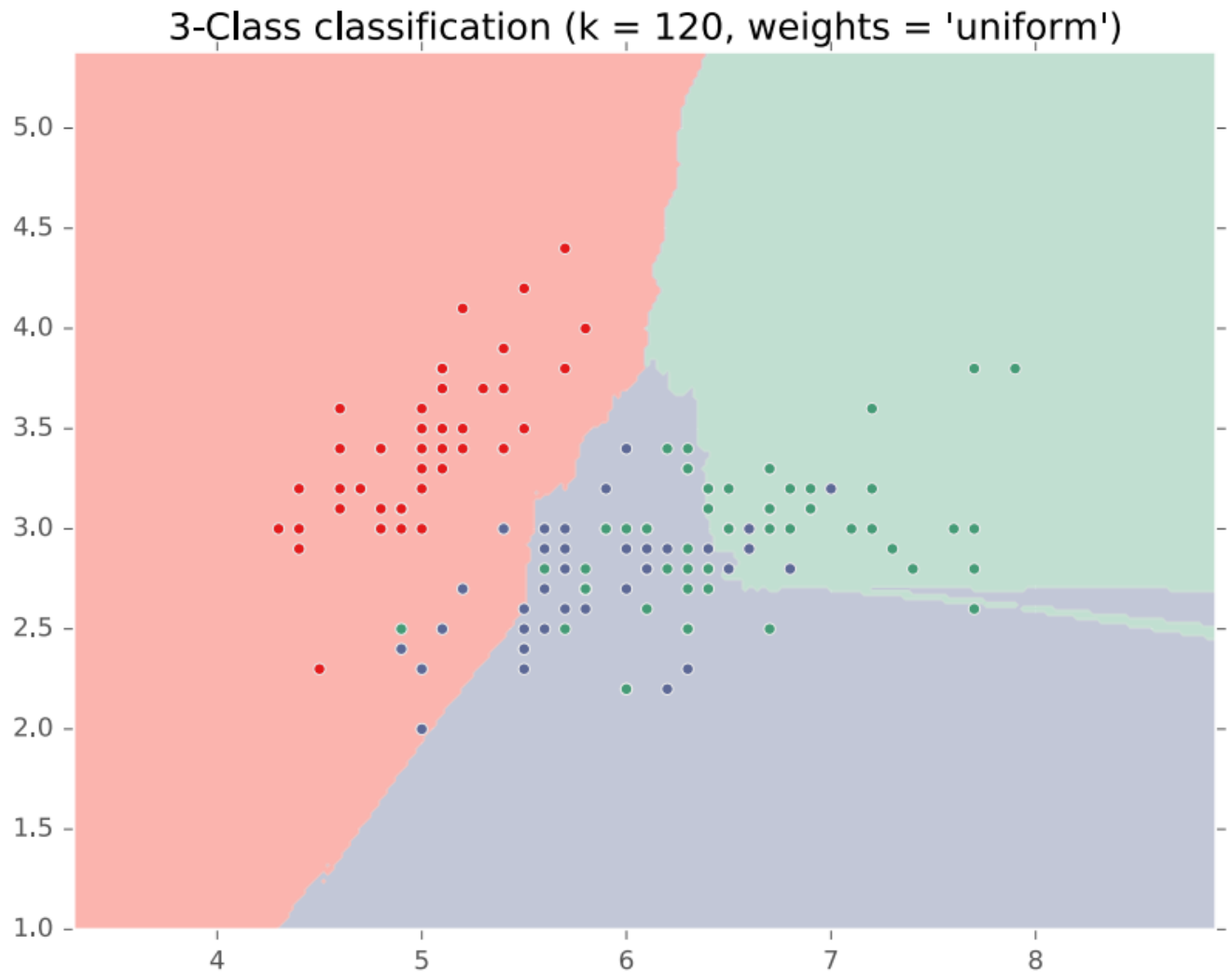
KNN on Fisher Iris Dataset



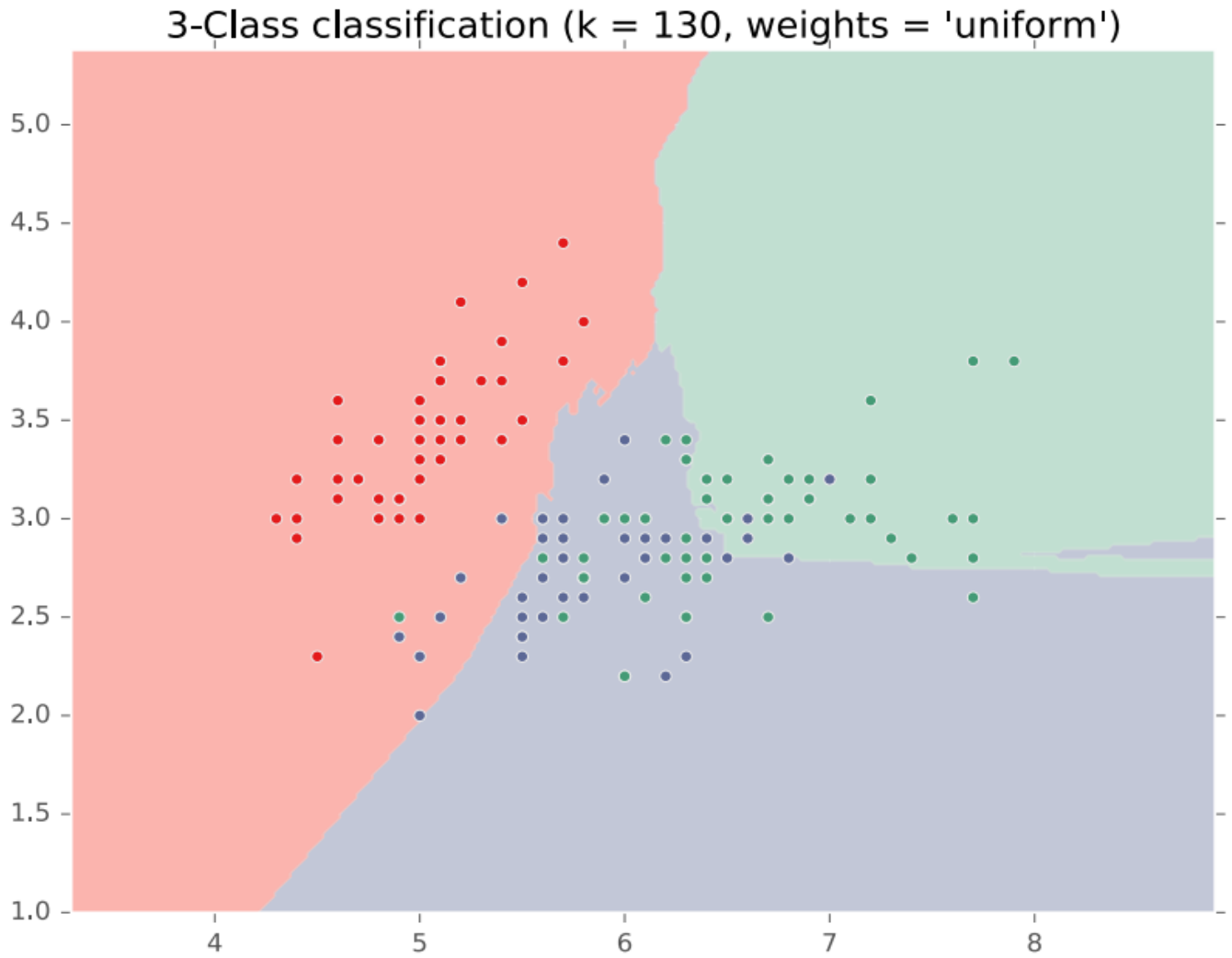
KNN on Fisher Iris Dataset



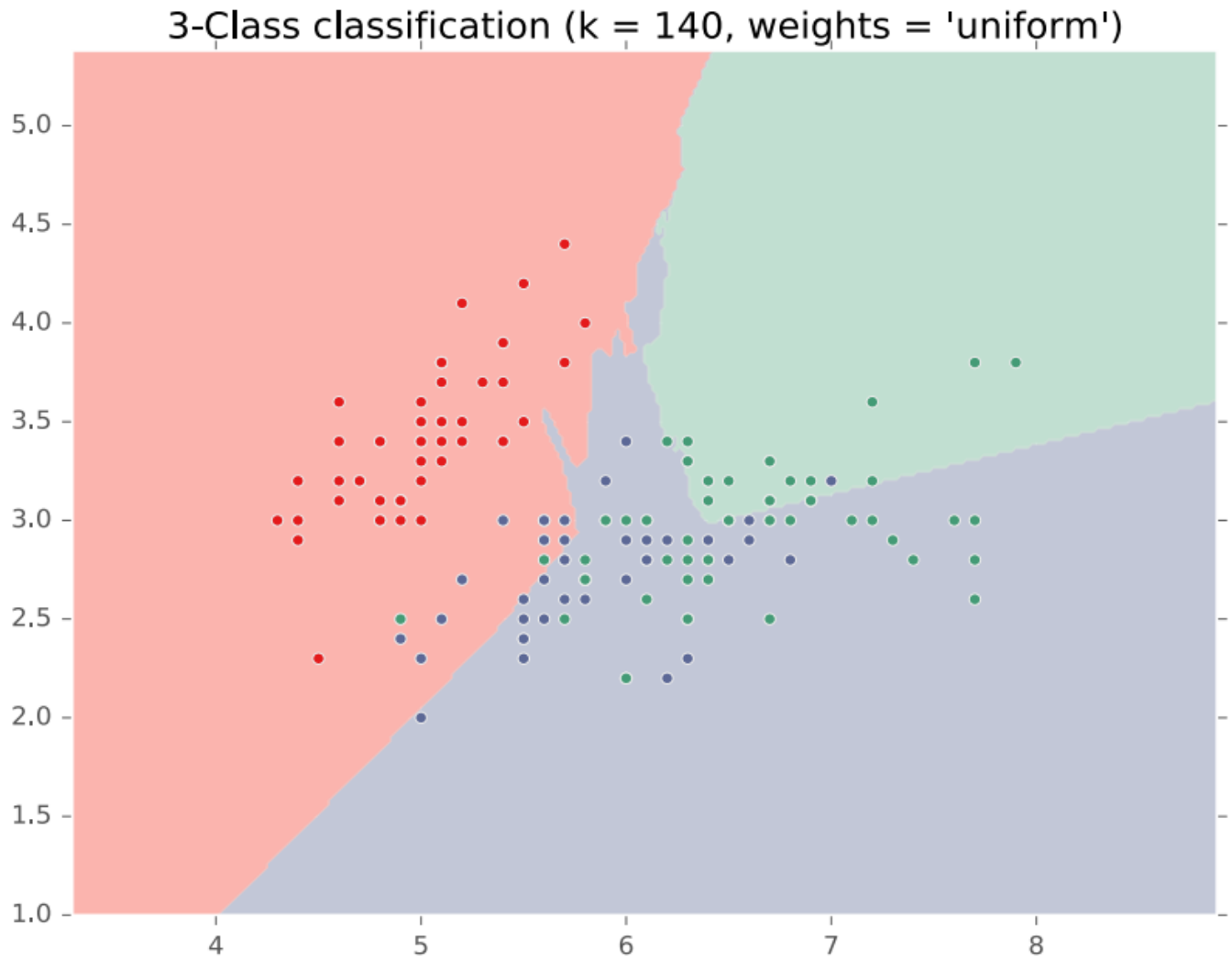
KNN on Fisher Iris Dataset



KNN on Fisher Iris Dataset



KNN on Fisher Iris Dataset



KNN on Fisher Iris Dataset