Non-Parametric Learners: KNN, Decision Trees

Isaac Laughlin

May 9, 2016

Objectives

At the end of today's lecture you should:

1. Be able to describe the KNN algorithm.
2. Describe the curse of dimensionality.
3. Recognize the conditions under which the curse may be problematic.
4. Enumerate strengths and weaknesses of KNN.

KNN

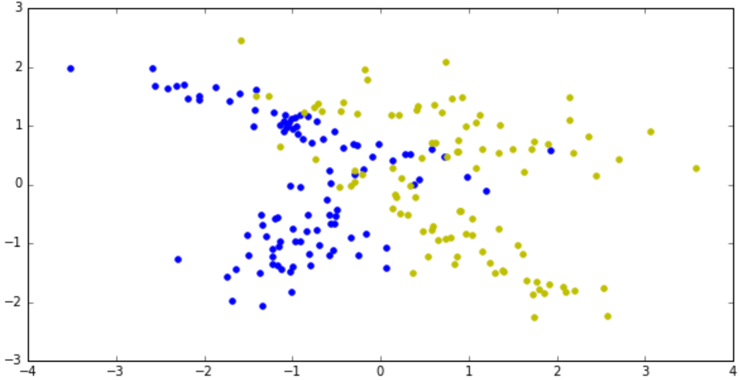

Figure 1:A classification problem

New data point

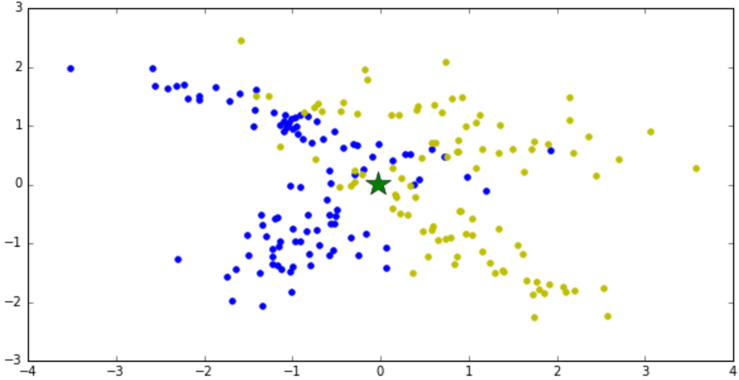

Figure 2:An uknown point

KNN

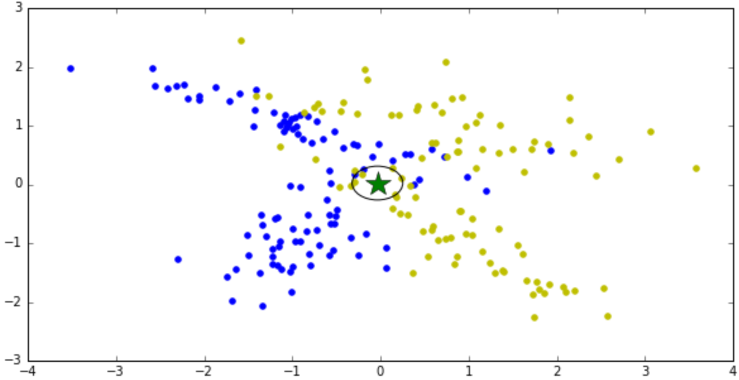

Figure 3:The approach

The KNN algorithm

Training:

1. Store all data.

Prediction:

1. Calculate the distance from new point to all points in dataset.
2. Keep the k nearest points to new point.
3. Predict the majority label.

What's k?

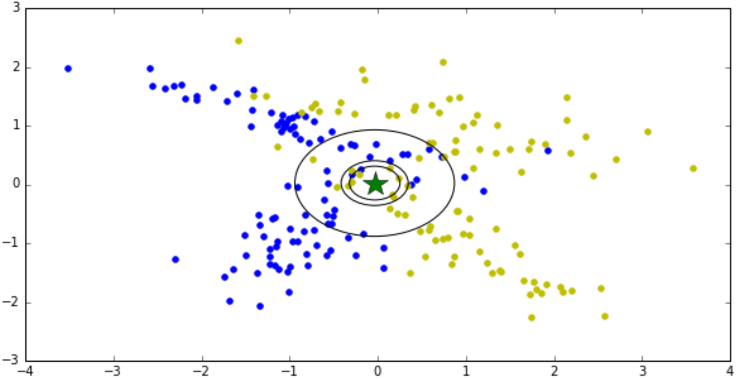

Figure 4:k = 5, 10, 40

1-neighbor

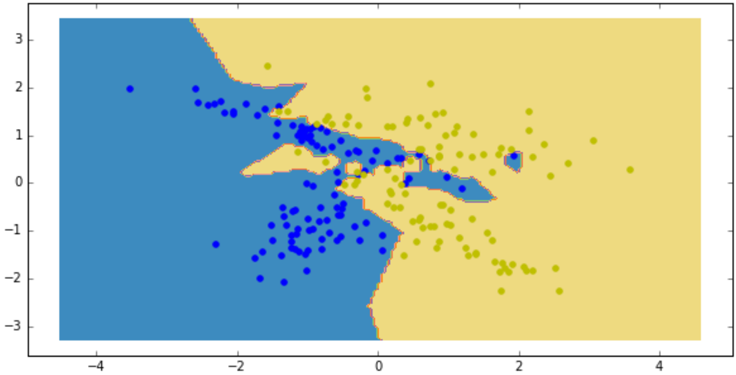

Figure 5:k = 1

3-neighbor

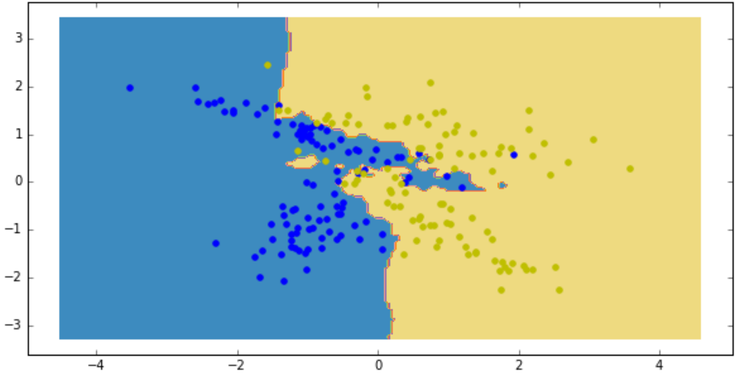

Figure 6:k = 3

10-neighbor

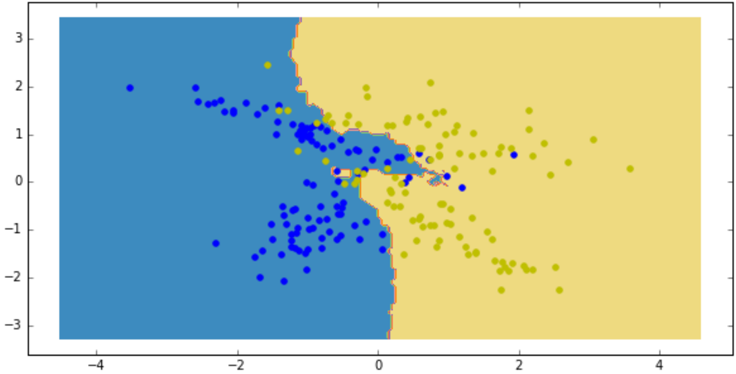

Figure 7:k = 10

100-neighbor

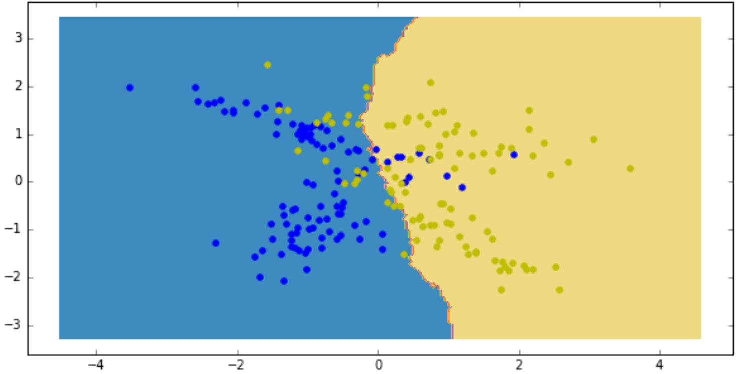

Figure 8:k = 100

Distance Metrics

So far we haven't been explicit about what distance metrics we're using. Some choices:

- Euclidean:

$$\sqrt{\sum_i (a_i - b_i)^2}$$

- Manhattan:

$$\sum_i |a_i - b_i|$$

- Cosine:

$$1 - \frac{a \cdot b}{||a||||b||}$$

Variants

- One variant is to weight the votes by $\frac{1}{d_i}$ so closer points get more weight.
- Use for regression, take (optionally, weighted) mean of continuous target rather than vote.
- Approximate nearest neighbors, overcomes performance issues.

Curse of Dimensionality

Curse of dimensionality

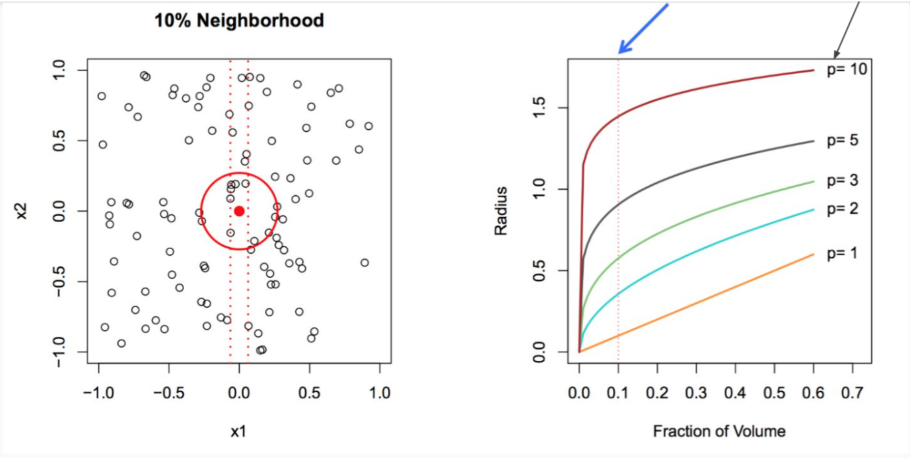

Figure 9:Curse of Dimensionality

Another view

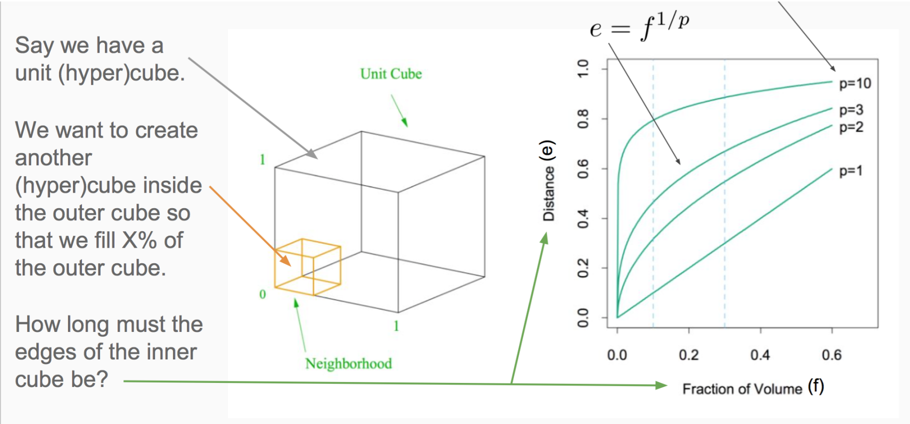

Figure 10:Hypercubes

Still another view

You have a dataset with 100 samples and one predictor.

You decide 1 predictor isn't enough so you decide to measure 10 predictors instead.

How many samples do you need to achieve the same sample density you originally had?

$$100^{1}0 = 1,000,000,000,000,000,000$$

Last one

$$lim_{d\to\infty}\frac{V_{sphere}(R,d)}{V_{cube}(R,d)} = lim_{d\to\infty}\frac{\frac{\pi^{d/2}R^d}{\Gamma(d/2+1)}}{(2R)^d} = lim_{d\to\infty}\frac{\pi^{d/2}}{2^d\Gamma(d/2+1)} = 0$$