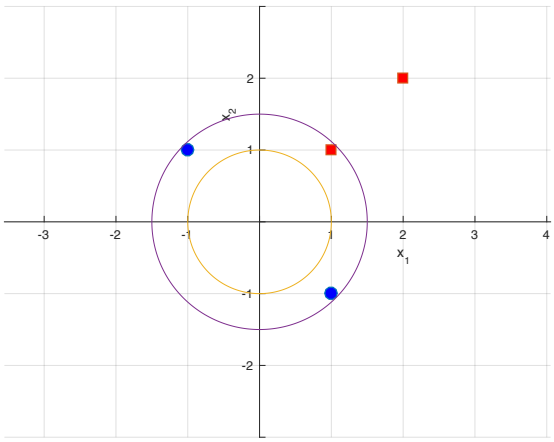
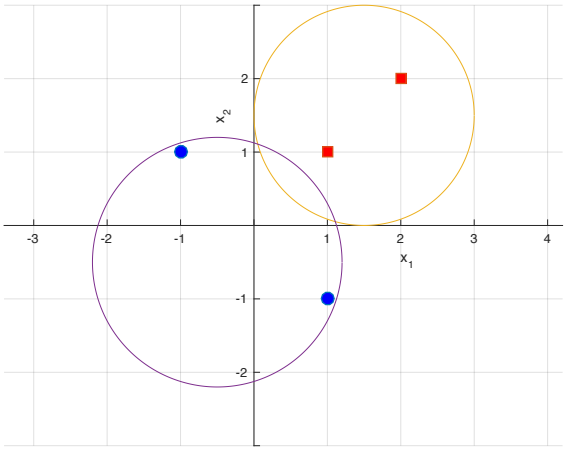
Discussion 1
Machine Learning, Spring 2019

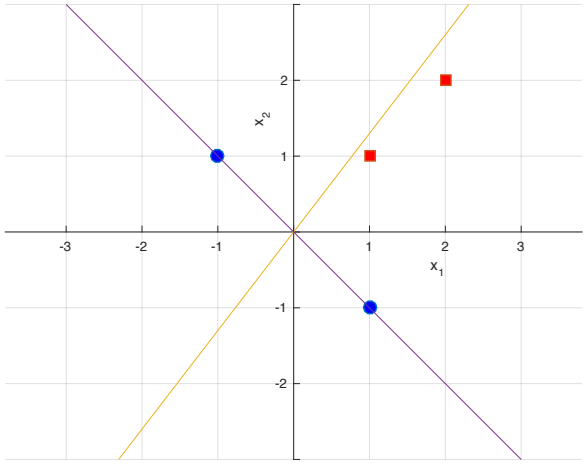
1.1, 1.2 Families of classifiers



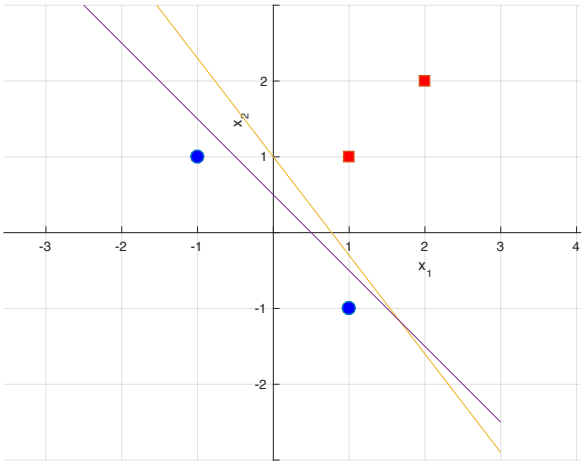
Circle – not separable



Circle with offset – separable



Line – not separable, linear



Line with offset – separable, linear

Figure 1: Circle and Line Classifiers.

1 Concepts of learning

1. Classification
2. Regression
3. Classification
4. Regression
5. Ranking
6. Clustering
7. Finding patterns
8. Density estimation
9. Conditional probability estimation
10. Classification

2.3 Linear Classifiers for Rain Prediction

Suppose we are given a linear classification model that predicts whether or not it is going to rain based upon the temperature (in degrees Celsius) and humidity (expressed as a percentage from 0-100). The model has weights defined such that if the sum of the temperature and the humidity exceeds 110, then it predicts rainfall instead of clear weather.

(a) Assume that an output of +1 corresponds to predicted rainfall. This model has a weight vector θ of length 2 and a nonzero offset θ_0. What are the values of θ and θ_0?

Answer:

$$y = sign(T + H - 110)$$

Thus, $\theta = [1, 1]^T$, $\theta_0 = -110$

(b) Consider what happens when we feed this model a data point from the planet Mercury (where it never rains) on which the temperature is observed to be 400° C with a humidity of zero. What does this model predict will happen on Mercury? What does this say about the generalization ability of this model?

Answer: When $T = 400, H = 0$ we have

$$y = sign(T + H - 110) > 0.$$

It predicts rainfall, which is false. It means the the classifier overfits the data on Earth, and can not be generalized to Mercury.

2.4 Higher dimensions

(a) Data is not separable

(b) See Figure 2

(c) Data is separable. Example of classifier: $x_2 - 3x_1 + 0.5 = 0$

(d) Advantages – separation is easier in higher dimensions; Disadvantages – overfitting, computational cost.

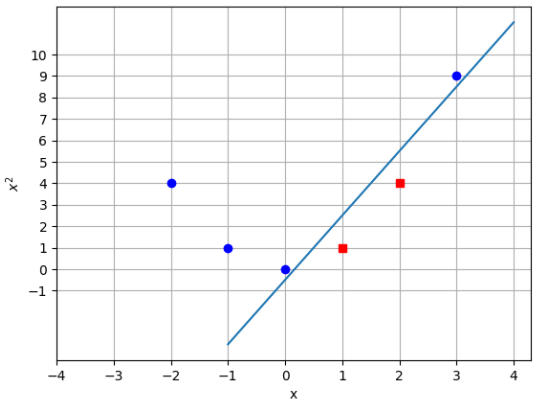


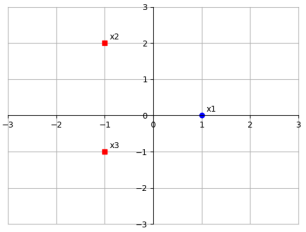
Figure 2: 1.4 Higher dimensions

3.3 Perceptron

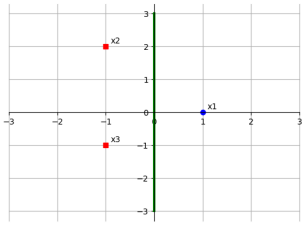
(a) 1 mistake

(b) 2 mistakes

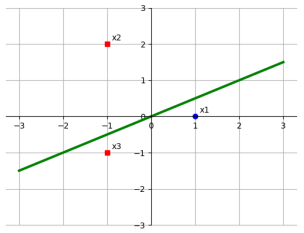
(c) $x_1^T x$ defines separating line.



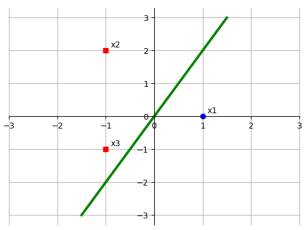
Dataset



Point x_1 – 1st mistake



Point x_2 – 1st mistake



Point x_2 – 2nd mistake

Figure 3: Perceptron