Linear models for conditional probability estimation

Topics we'll cover

1. Sources of uncertainty in prediction
2. Linear functions for conditional probability estimation
3. The logistic regression model

Uncertainty in prediction

Can we usually expect to get a perfect classifier, if we have enough training data?

Problem 1: Inherent uncertainty

The available features x do not contain enough information to perfectly predict y, e.g.,

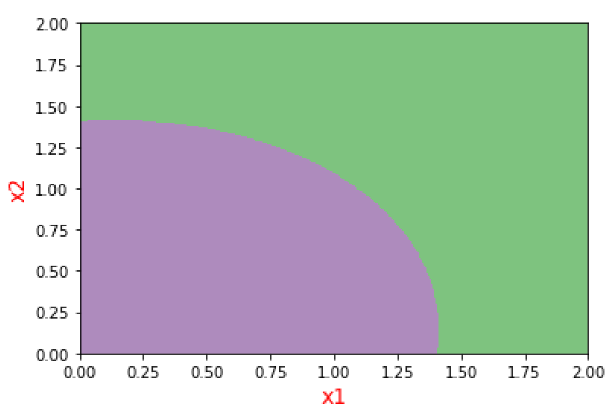
- $x =$ complete medical record for a patient at risk for a disease
- $y =$ will he/she contract the disease in the next 5 years?

Uncertainty in prediction, cont'd

Can we usually expect to get a perfect classifier, if we have enough training data?

Problem 2: Limitations of the model class

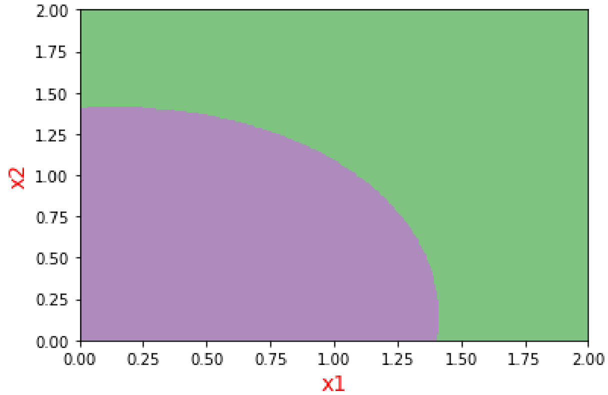
The type of classifier being used does not capture the decision boundary, e.g. using linear classifiers with:



Conditional probability estimation for binary labels

- Given: a data set of pairs (x, y), where $x \in \mathbb{R}^d$ and $y \in \{-1, 1\}$
- Return a classifier that also gives probabilities $\Pr(y = 1|x)$

Simplest case: using a linear function of x.



A linear model for conditional probability estimation

For data $x \in \mathbb{R}^d$, classify and return probabilities using a linear function

$$w_1 x_1 + w_2 x_2 + \cdots + w_d x_d + b \;=\; w \cdot x + b$$

where $w = (w_1, \ldots, w_d)$.

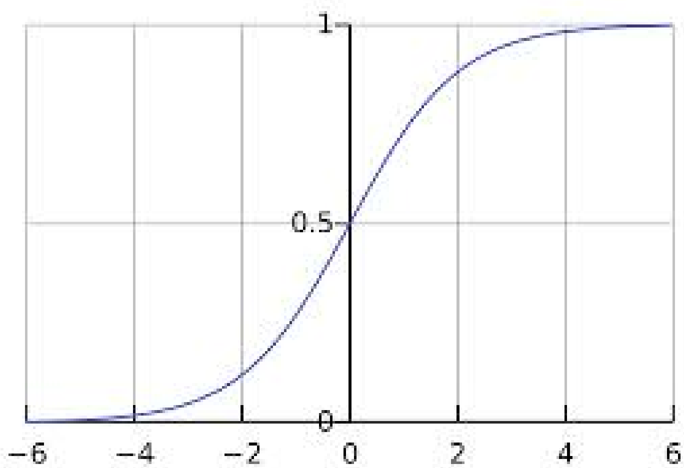
The probability of $y = 1$:

- Increases as the linear function grows.
- Is 50% when this linear function is zero.

How can we convert $w \cdot x + b$ into a probability?

The squashing function

$$s(z) \;=\; \frac{1}{1 + e^{-z}}$$



The logistic regression model

Binary labels $y \in \{-1, 1\}$. Model:

$$\Pr(y = 1|x) = \frac{1}{1 + e^{-(w \cdot x + b)}}$$

What is $\Pr(y = -1|x)$?

Summary: logistic regression for binary labels

- Data $x \in \mathbb{R}^d$
- Binary labels $y \in \{-1, 1\}$

Model parametrized by $w \in \mathbb{R}^d$ and $b \in \mathbb{R}$:

$$\Pr_{w,b}(y|x) = \frac{1}{1 + e^{-y(w \cdot x + b)}}$$

Learn parameters w, b from data