Bayesian Learning Methods

Parametric Learning

Assumptions can greatly simplify the learning process, but can also limit what can be learned. Algorithms that simplify the function to a known form are called parametric machine learning algorithms.

Naive Bayes

Naive Bayes is a parametric learning algorithm. It bases classification on the probability that something will happen on the relation between the attributes.

A classic example is whether or not to play golf.

Outlook	Temp	Humidity	Windy	Play golf
Rainy	Hot	High	False	No
Rainy	Hot	High	True	No
Overcast	Hot	High	False	Yes
Sunny	Mild	High	False	Yes
Sunny	Cool	Normal	False	Yes

(YES) Outlook: Sunny = 2, Rainy = 0, Overcast = 1

(NO) Outlook: Sunny = 0, Rainy = 2, Overcast = 0

Example (Play golf = YES): 2/3 * 1/3 * 2/3 * 3/3 * 3/5 = 0.0889 = 8.89% change to play golf.

Gaussian Distribution

Distinguish children from adults based on size:

Classes: {a,c}, attributes: height[cm], weight[kg]

Training examples: {h_i, w_i, y_i}, 4 adults, 12 children.

Class probabilities: P(a)=4/(4+12)=0.25; P(c) = 0.75

Model for adults:

Height ~ Gaussian

Weight ~ Gaussian

Assume height and weight to be independent.

Same model for children.

Plot the data and show them in a distribution model for X and Y. Then combine the two on the original data-system - this should create some elliptical shape. Do this for both adults and children. Now there should be two ellipses. They represent the normal distributions of children and adults.

Non-parametric Learning

Cumulative Distribution Function

Heads & Tails

f(q) = P(Q=q)

F(q) = P(Q<=q)

Flipping coin two times.

Q = number of heads {0,1,2}

f(0) = P(Q = 0) = P(t,t) = 1/4

F(0) = P(Q <= 0) = P(t,t) = 1/4

f(1) = P(Q = 1) = P(t,h) + P(h,t) = 2/4

F(1) = P(Q <= 1) = P(t,h) + P(h,t) + P(t,t) = 3/4

f(2) = P(Q = 2) = P(h,h) = 1/4

F(2) = P(Q <= 2) = P(t,h) + P(h,h) + P(h,t) + P(t,t) = 4/4 = 1