BAYESIAN ANALYSIS

MOHAMMAD GHODDOSI

BAYESIAN CLASSIFIERS

- Statistical classifiers
- Predict class membership probability
- Based on Bayes theorem
- Naïve Bayesian classifier
- Good for large datasets

BAYES THEOREM

- X is a data tuple (“evidence”)
- H is a predicted class (“hypothesis”)
- $P(H|X)$ is the probability that the hypothesis holds given the evidence
- $P(H|X)$ is the probability that data tuple X belongs to class C

BAYES THEOREM

- Posterior probability: $P(A|B)$
- Prior probability: $P(A)$
- $P(X)$ VS $P(H)$
- $P(X|H)$ VS $P(H|X)$

BAYES THEOREM

$$P(H|X) = \frac{P(X|H)\,P(H)}{P(X)}$$

EXAMPLE OF BAYESIAN THEOREM

- Consider XOR problem
- We have 4 inputs
- [0, 0] -> 0
- [0, 1] -> 1
- [1, 0] -> 1
- [1, 1] -> 0

EXAMPLE OF BAYESIAN THEOREM

- For example we want to test if [0, 0] is 0 or 1

$$P(y=0|x=[0,0]) = \frac{P(x=[0,0]|y=0) * P(y=0)}{P(x=[0,0])} = \frac{\frac{1}{2} * \frac{1}{2}}{\frac{1}{4}} = 1$$

$$P(y=1|x=[0,0]) = \frac{P(x=[0,0]|y=1) * P(y=1)}{P(x=[0,0])} = \frac{0 * \frac{1}{2}}{\frac{1}{4}} = 0$$

NAÏVE BAYESIAN CLASSIFIER

- Simple Bayesian classifier
- input tuple X
- X belongs to class with highest posterior probability

NAÏVE BAYESIAN CLASSIFIER

- $P(X)$ is constant for all classes
  - We chose class with maximum $P(X|C_i)\,P(C_i)$
- Naïve assumption
  - Attributes are conditionally independent
  - $P(X|C_i) = P(x_1|C_i) \times P(x_2|C_i) \times \cdots \times P(x_m|C_i)$
  - $P(X|C_i) = \prod_{k=1}^{m} P(x_k|C_i)$

$$P(C_i|X) = \frac{P(X|C_i)\,P(C_i)}{P(X)}$$

$P(x_k|C_i)$

- If x_k is categorical
  - Fraction
- If x_k is continues
  - Assume values of attribute have a Gaussian distribution
  - Find μ and σ of values

$$P(x_k|C_i) = \frac{1}{\sqrt{2\pi}\sigma_{C_i}} e^{-\frac{\left(x_k - \mu_{C_i}\right)^2}{2\sigma_{C_i}^2}}$$

EXAMPLE

- Consider OR problem
- We have 4 inputs
- [0, 0] -> 0
- [0, 1] -> 1
- [1, 0] -> 1
- [1, 1] -> 1

EXAMPLE

- For example we want to test if [0, 0] is 0 or 1

$$P(C_1) = \frac{3}{4} \quad P(C_0) = \frac{1}{4}$$

$$P(X = [0,0]|C_0) = P(x_1 = 0|C_0) \times P(x_2 = 0|C_0) = 1 \times 1 = 1$$

$$P(X = [0,0]|C_1) = P(x_1 = 0|C_1) \times P(x_2 = 0|C_1) = \frac{1}{3} \times \frac{1}{3} = \frac{1}{6}$$

$$P(C_0|X = [0,0]) = \alpha \times P(X = [0,0]|C_0) \times P(C_0) = \frac{1}{4}$$

$$P(C_1|X = [0,0]) = \alpha \times P(X = [0,0]|C_1) \times P(C_1) = \frac{1}{6} \times \frac{3}{4} = \frac{1}{8}$$