8. State Bayes theorem and discuss how Bayesian Classifier works?

Bayes' Theorem is named after Thomas Bayes', a Non Conformist English clergyman who did early work in probability and decision theory during the 18th Century

Let x be a data tuple. In Bayesian terms, x is considered "evidence". As usual, it is described by measurements made on a set n attributes. Let H be some hypothesis such as that the data tuple x belongs to a specified class c. For classification problems, we want to determine P(H/x), the probability that the hypothesis H holds given the "evidence" or observed data tuple x. In other words, we are looking for the probability that tuple x belongs to class c, given that we know the attribute description of x.

P(H/x) is the posterior probability, or a posterior probabilety, of H conditioned on x. For Example, Suppose our world of data tuples is confined to Customers described by the attributes age and income, respectively, and that x is a 35-year-old customer with an income of $ 40,000. Suppose that H is the hypothesis that our Customer will buy a Computer.



Scanned with CamScanner

In Contrast, $P(H)$ is the prior probability, or a priori probability, of $H$. For our Example this is the probability that any given customer will buy a Computer, regardless of age, Income, or any other information, for that matter. The posterior probability, $P(H/X)$ is based on more information than the prior probability, $P(H)$, which is independent of $X$

Similarly, $P(X/H)$ is the posterior probability of $X$ conditioned on $H$. That is it is the probability that a Customer, $X$, is 35 years old and earns $40,000, given that we know the customer will buy a Computer.

$P(X)$ is the prior probability of $X$.

$P(X)$, $P(X/H)$, and $P(X)$ may be estimated from the given data, as we'll shall See next. Bayes theorem is useful in that it provides a way of calculating the posterior probability, $P(H/X)$ from $P(H)$, $P(X/H)$, $P(X)$

Bayes's theorem is

$$P(H/X) = \frac{P(X/H)P(H)}{P(X)}$$



Scanned with CamScanner