Techniques for Analyzing Stochastic Time-Series Data

Dennis Castleberry Brandon Oubre Haikou Yu

November 18, 2013

The Naive Bayes Classifier

- Reduce classification to probability. What is $P(class|attribute1, attribute2, ..., attributeN)$.
- Assumes that each attribute is independent of the others. (Hence the "Naive" nickname.)
- For example, let's consider if a car is stolen using $P(stolen|Color, Type)$. Naive Bayes will assume $color = red$ and $type = sportscar$ to be independent.
- Naive Bayes is not sensitive to irrelevant attributes, since the probabilities of such attributes will be similar for all classes.

Advantages and Disadvantages of Naive Bayes

Advantages

- Only requires a single scan to train.
- Fast classification.
- Handles real and discrete data.
- Not sensitive to irrelevant attributes.

Disadvantages

- Assumes all attributes to be independent.

Training the Classifier