CS 383: Machine Learning

Prof Adam Poliak

Fall 2024

09/26/2024

Lecture 10

Outline

Probability

Naive Bayes

Confusion Matrix

Probability & Bayes Derivation

Bayes Rule

Conditional Probability

Marginal Probability

Outline

Probability

Naive Bayes

Confusion Matrix

Components of a Bayes Model

Identify the evidence, prior, posterior, and likelihood in the equation below

$$\boxed{p(y=k|\boldsymbol{x}) = \frac{p(y=k)p(\boldsymbol{x}|y=k)}{p(\boldsymbol{x})}}$$

Components of a Bayes Model

Identify the evidence, prior, posterior, and likelihood in the equation below

$$p(y = k|\boldsymbol{x}) = \frac{p(y = k)p(\boldsymbol{x}|y = k)}{p(\boldsymbol{x})}$$

Evidence: this is the data (features) we observe, which we think will help us predict the outcome we're interested in

Components of a Bayes Model

Identify the evidence, prior, posterior, and likelihood in the equation below

$$p(y = k|\boldsymbol{x}) = \frac{p(y = k)p(\boldsymbol{x}|y = k)}{p(\boldsymbol{x})}$$

Prior: without seeing any evidence (data), what is our prior believe about each outcome (intuition: what is the outcome in the population as a whole?)

Components of a Bayes Model

Identify the evidence, prior, posterior, and likelihood in the equation below

$$p(y = k|\boldsymbol{x}) = \frac{p(y = k)p(\boldsymbol{x}|y = k)}{p(\boldsymbol{x})}$$

Posterior: this is the quantity we are actually interested in. **Given** the evidence, what is the probability of the outcome?

Components of a Bayes Model

Identify the evidence, prior, posterior, and likelihood in the equation below

$$p(y = k|\boldsymbol{x}) = \frac{p(y = k)p(\boldsymbol{x}|y = k)}{p(\boldsymbol{x})}$$

Likelihood: given an outcome, what is the probability of observing this set of features?

Naive Bayes Example

	Cat	Documents
Training	-	just plain boring
	-	entirely predictable and lacks energy
	-	no surprises and very few laughs
	+	very powerful
	+	the most fun film of the summer

Test: powerful very fun

$p(+) = ?$

$p(-) = ?$

Naive Bayes Example

	Cat	Documents
Training	-	just plain boring
	-	entirely predictable and lacks energy
	-	no surprises and very few laughs
	+	very powerful
	+	the most fun film of the summer

Test: predictable with no fun

$p(+) = ?$ $\qquad$ $p(-) = ?$

Laplacian Smoothing

$$\hat{P}(w_i|c) = \frac{count(w_i,c)+1}{\sum_{w\in V}(count(w,c)+1)} = \frac{count(w_i,c)+1}{\left(\sum_{w\in V} count(w,c)\right)+|V|}$$

Naive Bayes Example

	Cat	Documents
Training	-	just plain boring
	-	entirely predictable and lacks energy
	-	no surprises and very few laughs
	+	very powerful
	+	the most fun film of the summer

Test: predictable with no fun

$p(+) = ?$ $p(-) = ?$