















Naïve Bayes Classifier $Y^{\text{predict}} = \operatorname{argmax} P(Y = v X_1 = u_1 \cdots X_m = u_m)$	
$Y^{\text{predict}} = \underset{v}{\operatorname{argmax}} \frac{P(Y = v^{\wedge} X_1 = u_1 \cdots X_m = u_m)}{P(X_1 = u_1 \cdots X_m = u_m)}$	
$Y^{\text{predict}} = \underset{v}{\operatorname{argmax}} \frac{P(X_{1} = u_{1} \cdots X_{m} = u_{m} Y = v)P(Y = v)}{P(X_{1} = u_{1} \cdots X_{m} = u_{m})}$)
$Y^{\text{predict}} = \underset{v}{\operatorname{argmax}} P(X_1 = u_1 \cdots X_m = u_m \mid Y = v) P(Y = v)$	
Because of the struct the Bayes Net	ure of
$Y^{\text{predict}} = \underset{v}{\operatorname{argmax}} P(Y = v) \prod_{j=1}^{n_{Y}} P(X_{j} = u_{j} Y = v)$	
Copyright © 2004, Andrew W. Moore	Slide 9

More Facts About Naïve Bayes Classifiers

- Naïve Bayes Classifiers can be built with real-valued inputs*
- Rather Technical Complaint: Bayes Classifiers don't try to be maximally discriminative---they merely try to honestly model what's going on*
- Zero probabilities are painful for Joint and Naïve. A hack (justifiable with the magic words "Dirichlet Prior") can help*.
- Naïve Bayes is wonderfully cheap. And survives 10,000 attributes cheerfully!

*See future Andrew Lectures

Copyright © 2004, Andrew W. Moore

Slide 10

What you should know • How to build a Bayes Classifier • How to predict with a BC

Slide 11