Bayes optimal classifier
Naïve Bayes
What's learning, revisited

Machine Learning – 10701/15781

Carlos Guestrin

Carnegie Mellon University

September 21st, 2009

Classification

GPA
Grade 10701

- **Learn**: h:**X** ↦ Y
  - **X** – features
  - Y – target classes

instead of continuous as in regression
↳ discrete
(hired, not hired)

- Suppose you know P(Y|**X**) exactly, how should you classify?
  - Bayes classifier: (optimal)

$$y^* = \arg\max_y P(Y=y \mid X=x)$$

$P(Y=\text{hired} \mid \text{GPA}=4.0, 10701=\text{B}) = 0.01$

- **Why?**

Optimal classification

- **Theorem:** Bayes classifier h_{Bayes} is optimal!

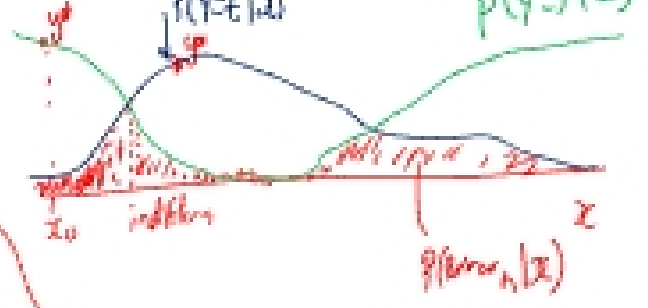
$$h_{Bayes}(x) = \arg\max_y P(Y=y|X=x)$$

  - That is $error_{true}(h_{Bayes}) \leq error_{true}(h), \ \forall h(\mathbf{x})$

- **Proof:** $p(error_h) = \int_x p(error_h|x)p(x)dx$

if I say $h(x) = hired$ ← small

$P(error_h|x) = P(Y = \text{not hired}|x)$ ← as small as possible ⇒ h_{Bayes}



Bayes Rule

likelihood prior

$$P(Y|X) = \frac{P(X|Y)P(Y)}{P(X)}$$

↗ we want this

$P(Ck=4.0, \text{logloc}B|hired)$ $P(hired)$

Which is shorthand for:

$$(\forall i, j) P(Y = y_i|X = x_j) = \frac{P(X = x_j|Y = y_i)P(Y = y_i)}{P(X = x_j)}$$

How hard is it to learn the optimal classifier?

- Data =

Sky	Temp	Humid	Wind	Water	Forest	EnjoySpt
Sunny	Warm	Normal	Strong	Warm	Same	Yes
Sunny	Warm	High	Strong	Warm	Same	Yes
Rainy	Cold	High	Strong	Warm	Change	No
Sunny	Warm	High	Strong	Cool	Change	Yes

- How do we represent these? How many parameters?
  - Prior, P(Y):
    - Suppose Y is composed of *k* classes
  - Likelihood, P(**X**|Y):
    - Suppose **X** is composed of *n* binary features

- Complex model ! High variance with limited data!!!

Conditional Independence

- X is **conditionally independent** of Y given Z, if the probability distribution governing X is independent of the value of Y, given the value of Z

$$(\forall i, j, k) P(X = i | Y = j, Z = k) = P(X = i | Z = k)$$

- e.g., $P(Thunder | Rain, Lightning) = P(Thunder | Lightning)$
- Equivalent to:

$$P(X, Y \mid Z) = P(X \mid Z) P(Y \mid Z)$$