















> Course Shortcuts

> Student Lounge

> Q&A

Estimating Distributions from Data

Classification with the NaiveBayes Algorithm

- Module Introduction:
 Classification with the
 Naive Bayes Algorithm
- Limitations of High
 Dimensional Space in MLE
- MLE Curse of Dimensionality
- Formalize the Curse of Dimensionality
- Capturing All Possibilities in d Dimensions
- Naive Bayes Assumption
- Data Sets Where the Naive Bayes Assumption Holds
- Naive Bayes Classifier
- Derivation of Naive Bayes
 Classifier
- Naive Bayes Cheat Sheet
- Determine Probability
 With Categorical Naive
 Bayes
- Categorical Naive Bayes
 Classifiers
- Naive Bayes in Action
- Module Wrap-up:
 Classification with the
 Naive Bayes Algorithm
- > Building a Baby NameClassifying System
- > Course Resources

Estimating Probability Distributions >

☼ Naive Bayes in Action

Suzie is about to go on a blind date. Her friends want to help her to distinguish good from bad guys. Suzie's friends train a Naïve Bayes classifier to predict if someone is good or bad. As training data, they collect six people who they know are definitely good (Batman, Superman, Spiderman) or bad (Riddler, Penguin, Joker). Suzie's friends define the following features and labels:

Binary features:

C = "Guy wears a cape."

M = "Guy wears a mask."

U = "Guy wears his underwear outside his pants"

Labels:

G = "Guy is good"

B = "Guy is bad"

Features	Spiderman	Superman	Batman	Joker	Penguin	Riddler
Cape (C)	F	Т	Т	F	Т	F
Mask (M)	Т	Т	F	F	F	Т
Underwear outside of pants (U)	F	Т	Т	F	F	F
Label	G	G	G	В	В	В

Recall these definitions from previous sections:

$$\mathrm{P}(y=c)=rac{\sum_{i=1}^{n}I(y_{i}=c)}{n}=\hat{\pi}_{c}$$

$$\mathrm{P}(x=j|y=c) = rac{ ext{\# of samples with label c that have feature $lpha$ with value j}{ ext{\# of samples with label c}}$$

Fill in the following probabilities based on the data from the table above. The format of your input should be a decimal to the hundredths place (e.g., 0.65)

Class Prior Probabilites

P(G) 0.5

P(B) 0.5 **✓**

Probabilities of Features for "Good" Class

P(C|G) 0.67 ✔

P(M|G) 0.67 ✓

P(U|G) 0.67 ✓

Probabilities of Features for "Bad" Class

P(C|B) 0.33 ✓

P(M|B) 0.33 **✓**

P(U|B) 0.00 ✓



Hint: P(M,C|G)

0.25

✓ 0.44

The Naïve Bayes assumption says that the joint probability of features given label is independent, and thus equal to the probabilities of each feature given label multiplied.

Therefore, $P(M,C\mid G)=P(M\mid G)\cdot P(C\mid G)=0.66\cdot 0.66=0.44.$

0.66

0.33





>