

## Bayesian Classifier

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad - (1)$$

$$P(B|A) = \frac{P(B \cap A)}{P(A)} \quad - (2)$$

$$P(A \cap B) = P(B|A)P(A)$$

$$P(A|B) = \frac{P(B|A)P(A)}{P(A)} \quad - (3)$$

Divide (1) by (3)

$$\frac{P(A|B)}{P(B|A)} = \frac{P(A \cap B)}{P(B)} \times \frac{P(A)}{P(A \cap B)}$$

$$\frac{P(A|B)}{P(B|A)} \Rightarrow \frac{P(A)}{P(B)}$$

$$\frac{P(B|A)}{P(A|B)} = \frac{P(B)}{P(A)}$$

Bayes theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Day: \_\_\_\_\_

Date: \_\_\_\_\_

hypothesis

$$P(H/X) = \frac{P(X/H)P(H)}{P(X)}$$

Likelihood

(probability after evidence came)

Posterior probability

$$P(H/X) = \frac{P(X/H)P(H)}{P(X)}$$

evidence

Prior probability

Hypothesis

1. Orange

2. Apple

3. Cherry

Posterior likelihood

Prior probability  
(already known)

$$P(\text{orange} | \text{red n round}) = ①$$

$$P(\text{apple} | \text{red n round}) = ②$$

$$P(\text{cherry} | \text{red n round}) = ③$$

$$P(\text{orange} | \text{red n round}) = P(\text{red n round} | \text{orange})$$

maximize

$$\times P(\text{orange})$$

$$P(\text{red n round})$$

$$P(\text{orange}) = \frac{3}{10}$$

$$P(\text{red n round}) = \frac{5}{10}$$

Day: \_\_\_\_\_

Date: \_\_\_\_\_

MAP

maximize a posterior probability

Naive Bayesian Classifier



(Simple Bayesian Classifier)

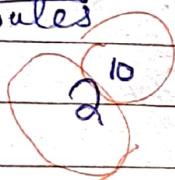
$P(\text{red round orange})$  is  
computationally expensive

We make an assumption

Store all the possible  
possibilities

12 possibilities      2  $\rightarrow$  4 (red round  $\rightarrow$  orange)  
 $\rightarrow$  4 (red round apple)  
 $\rightarrow$  4 (red round - day)

10 attributes



10

$2 \times 2 \times 2$

10 attributes

$$5^{10} \rightarrow 9765625$$

Day: \_\_\_\_\_

Date: \_\_\_\_\_

P(redn sound | orange)

Class Conditional Independence

6x2

$$\begin{matrix} & 3 \times 2 \\ & D \quad C \\ & | \quad | \\ & 3 \quad 1 \\ & | \quad | \\ & 1 \quad 1 \end{matrix}$$

Independent events

$$P(D=3 \cap C=1)$$

$$= P(D=3) \cdot P(C=1)$$

$$= \frac{1}{6} \times \frac{1}{2}$$

$$= \frac{1}{12}$$

$$3 \times 3 \times 3 = 27$$

3

orange top

4/3

2/3

2/3 x 2

orange sound

4/3

2/3

1

3

2

1

0

-1

-2

-3

-4

-5

-6

-7

-8

-9

-10

-11

-12

-13

-14

-15

-16

-17

-18

-19

-20

-21

-22

-23

-24

-25

-26

-27

-28

-29

-30

-31

-32

-33

-34

-35

-36

-37

-38

-39

-40

-41

-42

-43

-44

-45

-46

-47

-48

-49

-50

-51

-52

-53

-54

-55

-56

-57

-58

-59

-60

-61

-62

-63

-64

-65

-66

-67

-68

-69

-70

-71

-72

-73

-74

-75

-76

-77

-78

-79

-80

-81

-82

-83

-84

-85

-86

-87

-88

-89

-90

-91

-92

-93

-94

-95

-96

-97

-98

-99

-100

-101

-102

-103

-104

-105

-106

-107

-108

-109

-110

-111

-112

-113

-114

-115

-116

-117

-118

-119

-120

-121

-122

-123

-124

-125

-126

-127

-128

-129

-130

-131

-132

-133

-134

-135

-136

-137

-138

-139

-140

-141

-142

-143

-144

-145

-146

-147

-148

-149

-150

-151

-152

-153

-154

-155

-156

-157

-158

-159

-160

-161

-162

-163

-164

-165

-166

-167

-168

-169

-170

-171

-172

-173

-174

-175

-176

-177

-178

-179

-180

-181

-182

-183

-184

-185

-186

-187

-188

-189

-190

-191

-192

-193

-194

-195

-196

-197

-198

-199

-200

-201

-202

-203

-204

-205

-206

-207

-208

-209

-210

-211

-212

-213

-214

-215

-216

-217

-218

-219

-220

-221

-222

-223

-224

-225

-226

-227

-228

-229

-230

-231

-232

-233

-234

-235

-236

-237

-238

-239

-240

-241

-242

-243

-244

-245

-246

-247

-248

-249

-250

-251

-252

-253

-254

-255

-256

-257

-258

-259

-260

-261

-262

-263

-264

-265

-266

-267

-268

-269

-270

-271

</

Day:

Independent event

Date:

$$P(\text{red} \wedge \text{round} \mid \text{orange}) =$$

$$P(\text{red} \mid \text{orange}) \times$$

$$P(\text{round} \mid \text{orange})$$

3

10 attributes

every one is binary

$$2^{10} = 1024 \times 3$$

$$2^{10} = 3076$$

10 attributes

$$2 + 2 + 2 + \dots + 2$$

red/orange

not red/orange

sweet/orange

bitter/orange

10 values

$$2 \times 10$$

$$\therefore 20 \text{ for each class}$$

$$20 \times 3$$

$$\therefore 60$$

Day:

Date:

training  
✓  
stored probabilities

testing  
result on  
these stored  
probabilities

100 attributes

binary  
 $2^{100}$   
 $2 \times 2$

$\rightarrow 2+2 \dots$   
200  
 $200 \times 2$   
 $= 400$

A	B	C	D	E	Class
T	T	T	T	T	X
F	F	F	F	F	X
T	T	F	T	T	Y
					X
					X

Independent

$P(A=T|X)$   
 $P(A=F|X)$

$P(A=T|Y)$   
 $P(A=F|Y)$

$2^5 = 32$  for one class

$32 \times 2 = 64$

Independent  $2 \times 5 = 10$

$10 \times 2 = 20$

Date: \_\_\_\_\_

Date: \_\_\_\_\_

A

B

C

Class

a<sub>1</sub>b<sub>1</sub>c<sub>1</sub>

Y

a<sub>2</sub>b<sub>2</sub>c<sub>2</sub>

Y

a<sub>3</sub>a<sub>1</sub>b<sub>1</sub>c<sub>3</sub>a<sub>2</sub>b<sub>2</sub>c<sub>1</sub>

X

$$= 3 \times 2 \times 3$$

$$= 18 \text{ for } X.$$

~~$$= 3 \times 2 \times 3$$~~

$$= 18 \text{ for } Y.$$

36

$$\text{for } X \quad 3+2+3 = 8$$

$$\text{for } Y \quad 3+2+3 = 8$$

16

Day: \_\_\_\_\_

Date: \_\_\_\_\_

Buy computer class

$$\begin{array}{r} 3 \times 3 \times 2 \times 2 \\ \hline 0 \\ 36 \\ \hline 72 \end{array}$$

$$\begin{array}{r} 3 + 3 + 2 + 2 \\ = 10 \end{array}$$

$$\begin{array}{r} \leq 10 \times 2 \\ = 20 \end{array}$$

$X$  = Age  $\leq 30$ , income = medium,

student says yes, credit says fair,

$$P(\text{Buy computer} = \text{yes} | X) =$$

$$P(\text{Buy computer} = \text{no} | X) =$$

$$X \cdot P(\text{Yes}) \stackrel{?}{=} P(\text{Yes} | X)$$

Day: \_\_\_\_\_

Date: \_\_\_\_\_

$$\frac{P(X/H)}{P(X)} \cdot P(H)$$

$$P(\text{<30/Yes}) \times P(\text{medium/Yes}) \times P(\text{Yes/No}) \times P(\text{Fair/Yes}) \\ \times P(\text{Yes})$$

P(X) No need same  
in both

$$P(\text{By-comb/Yes}) =$$

$$P(\text{<30/Yes}) \times P(\text{Yes/No}) \times \dots$$

(medium n Yes)

$$\frac{1}{P(Yes)}$$

$$(Yes) \quad \frac{2}{9} \times \frac{4}{9} \times \frac{6}{9} \times \frac{6}{9} = 0.64$$

$$(No) \quad \frac{3}{5} \times \frac{2}{5} \times \frac{1}{5} \times \frac{2}{5} = 0.08$$

Now  $\rightarrow$  Which one is greater will win

5

Day: \_\_\_\_\_

Date: \_\_\_\_\_

## Normalize

$$\frac{0.028}{0.028 + 0.007} = 0.8$$

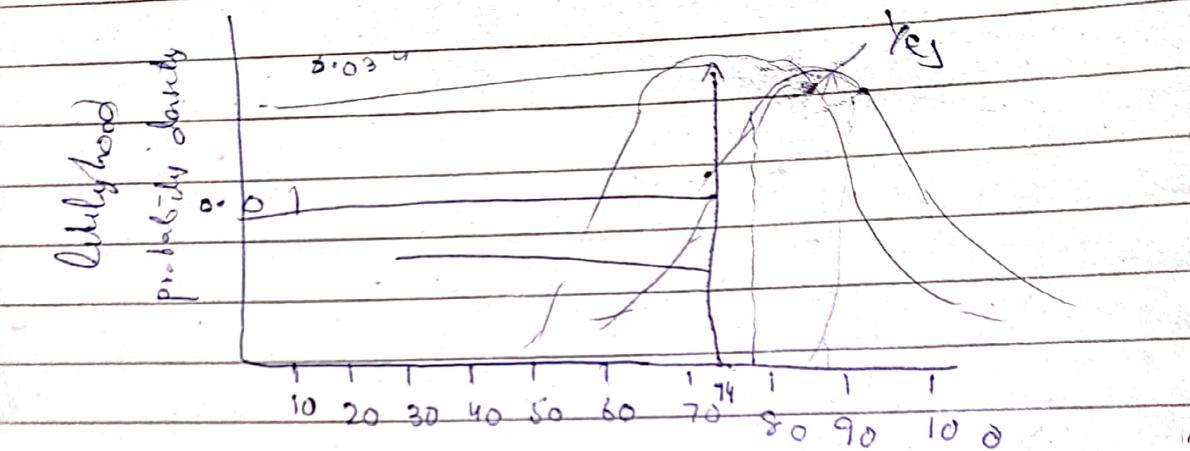
$$\frac{0.007}{0.028 + 0.007} \rightarrow 0.2$$

$$\text{Yes- } 1 \times 2 \times 3 \times \text{Yes}$$
$$No = 1 \times 2 \times 3 \times No$$

$$P(\text{Yes}) \times P(X_5 \text{ vs } v_0)$$

Day: \_\_\_\_\_

Gaussian



yes 86 96 80 65 70 80 70 90 85

Mean

std dev

79.1

10.2

 $N_0$ 

86

74

 $74 \rightarrow \text{peak} \rightarrow \text{greater}$ 

Gaussian

For every column

mean and median

must be

similar

Day: \_\_\_\_\_

L X Date: \_\_\_\_\_

Smoothing X

laplace

$c(x, y) + l$

$c(j) + l(3)$

Day: \_\_\_\_\_

## Multinomial Naïve Bayes Classifier

o fit  $\rightarrow$  find columns

transform  $\rightarrow$  set values from in  
columns

What to use

Features documents have  
frequency  
counts

$$P(+ | GD) =$$

$$\frac{P(\text{good} | +) * P(\text{nice} | +)}{P(+)} \cdot$$

$$P(- | GD) = P(\text{good} | -) *$$

$$P(\text{nice} | -)$$

$$P(-) = \frac{1}{12} \times \frac{1}{12} \times \frac{2}{5}$$

$$P(+)$$

Day:

Date:

→ Normalization

test data

training

new

mean

just adding  
into the  
columns

fit transform

hot dog

on the basis  
of training.

because of  
mean

can't

hot  
known  
the word above  
frequency.

transform

using

sentence  
so the data has  
no words  
so training  
already use  
those so.  
isilge

but fit

in rule  
so transform into  
what rule  
words

Bernoulli: NB  
 $(a-1)$

$$\frac{d}{2N}$$

$$P(+ | TD)_{s_1}$$

$$P(\text{nice} | +) *$$

$$P(\text{good} | +) * P(+)$$

$$P(- | TD)_{s_2}$$

$$P(\text{nice} | -) *$$

$$P(\text{good} | -) * P(-)$$

Bernoulli:

$$P(+ | TD)_{s_1} = P(\text{good} | +) *$$

$$*(1 - P(\text{bad} | +)) * (1 - P(\text{bad} | +))$$

$$*(1 - P(\text{unpleasant} | +)) *$$

$$P(\text{nice} | +) * (1 - P(\text{pleasant} | +)) *$$

$$P(- | TD)_{s_2} \quad (1 - P(\text{rude} | +)),$$

$$P(- | TD)_{s_2}$$