ARTIFICIAL INTELLIGENCE

LEARN RISE SHINE

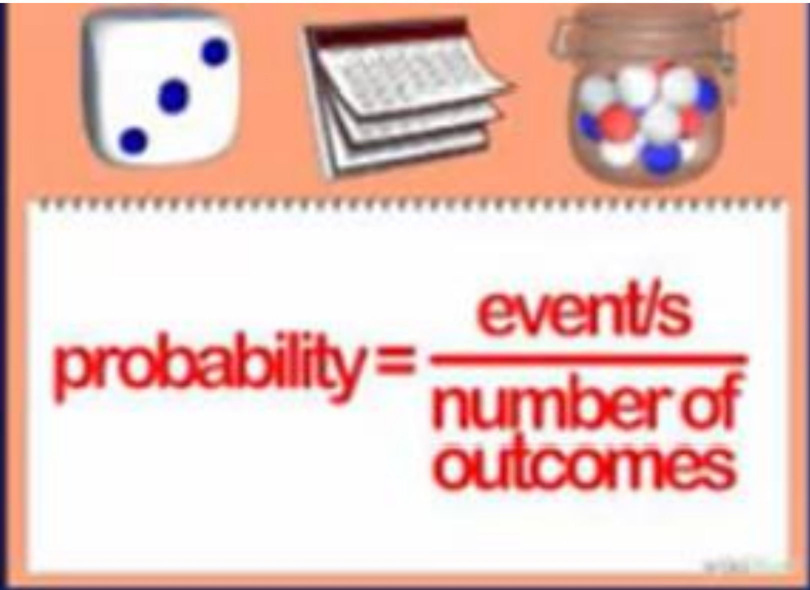
NAVTTC

Resource Person
Ms. Qurat Ul Ain Raja

Contents

- **Probability in Machine Learning**
- **Naïve Bayesian Algorithm**

Naïve Bayes Overview



Probabilistic approach to classification

- Relationships between input features and class expressed as probabilities
- Label for sample is class with highest probability given input

Naïve Bayes Classifier

Classification Using Probability

+

Bayes Theorem

+

Feature Independence Assumption

Probability of Event

Probability is measure of how likely an event is

Probability of Event ‘A’ Occurring

$$P(A) = \frac{\#\text{ ways for A}}{\#\text{ possible outcomes}}$$

Probability of Event

What is the probability of rolling a die and getting 6?

Probability of Rolling 6 on a Die

$$P(6) = \frac{\text{\# ways for getting 6}}{\text{\# possible outcomes}} = \frac{1}{6}$$

Bayes' Theorem

- Relationship between P(B | A) and P(A | B) can be expressed through Bayes' Theorem

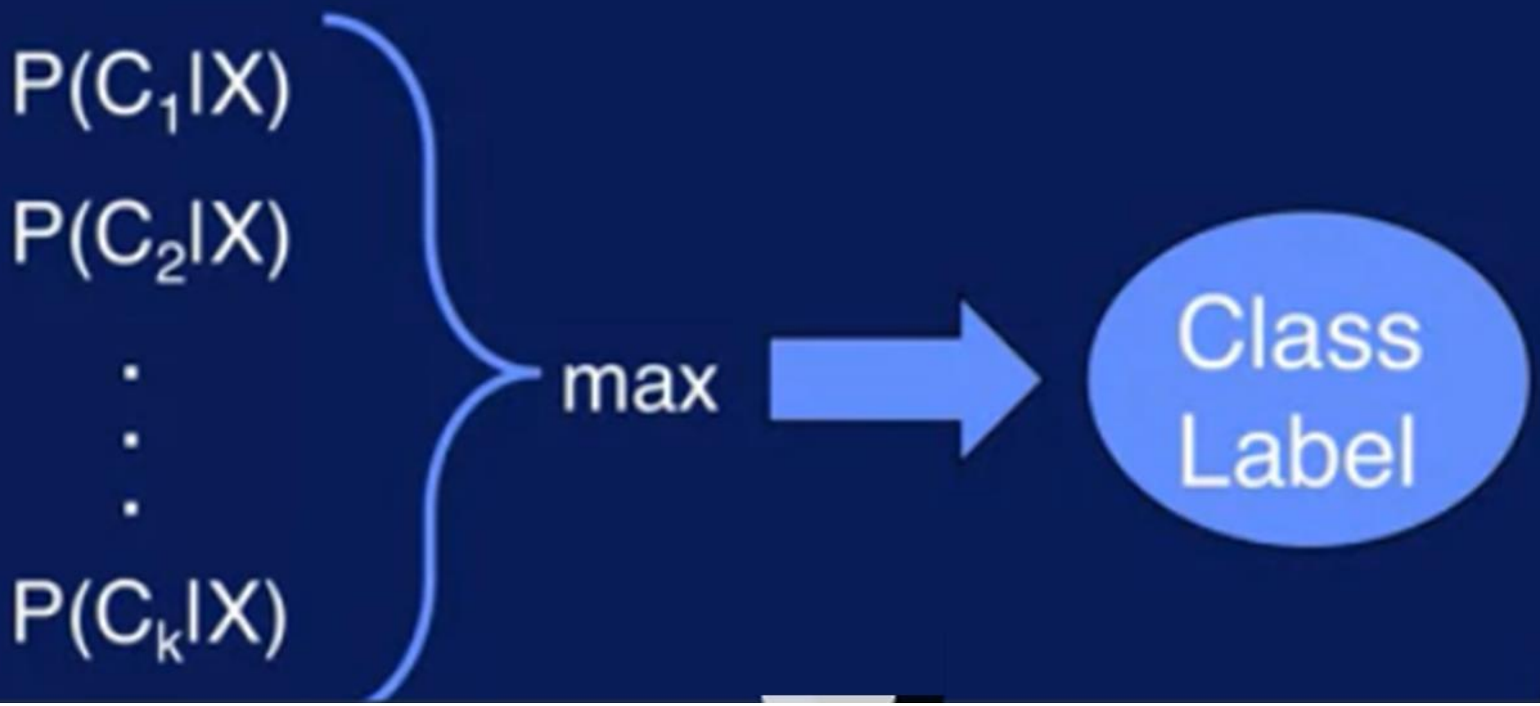
$$P(B \mid A) = \frac{P(A \mid B) * P(B)}{P(A)}$$

Bayes' Theorem

Classification with Probabilities

Given features X={X1,X2,…,Xn}, predict class C

Do this by finding value of C that maximizes P(C l X)



Bayes' Theorem for Classification

- **But estimating P(C|X) is difficult**
- **Bayes' Theorem to the rescue!**
  - Simplifies problem

Bayes' Theorem for Classification

Posterior Probability

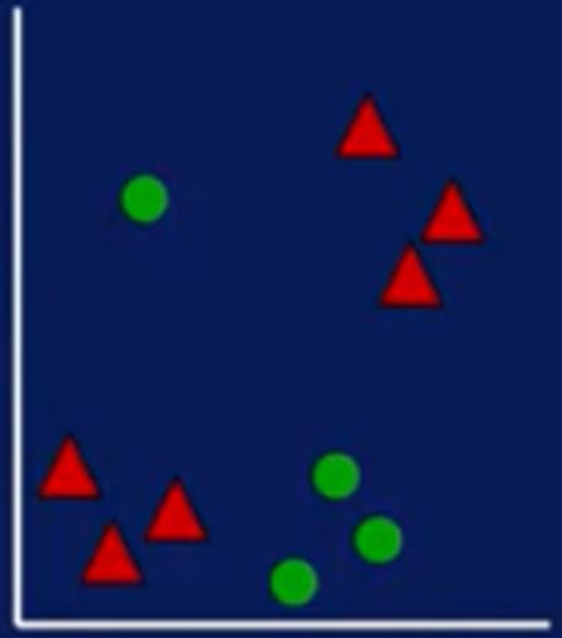
Class-Conditional Probability

Prior Probability

$$P(C \mid X) = \frac{P(X \mid C) * P(C)}{P(X)}$$

Probability of observing values for input features

Estimating P(C)



$P(\bullet) = 4/10 = 0.4$

$P(\blacktriangle) = 6/10 = 0.6$

To estimate P(C), calculate fraction of samples for class C in training data.

Estimating $P(X_i \mid C)$

Home Owner	Marital Status	Loan Default
Yes	Single	No
No	Married	No
No	Single	No
Yes	Married	No
No	Divorced	Yes
No	Married	No
Yes	Divorced	No
No	Single	Yes
No	Married	No
No	Single	Yes

P(Home Owner = Yes l No) = 3/7 = 0.43

P(Marital Status = Singlel Yes) = 2/3 = 0.67

Naïve Bayes Classifier

- C_1: buys_computer = “yes”
- C_2: buys_computer = “no”

- You can use the given formula:

$$P(C_i|\mathbf{X}) = \frac{P(\mathbf{X}|C_i)P(C_i)}{P(\mathbf{X})}$$

RID	age	income	student	credit_rating	Class: buys_computer
1	youth	high	no	fair	no
2	youth	high	no	excellent	no
3	middle_aged	high	no	fair	yes
4	senior	medium	no	fair	yes
5	senior	low	yes	fair	yes
6	senior	low	yes	excellent	no
7	middle_aged	low	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	medium	yes	fair	yes
11	youth	medium	yes	excellent	yes
12	middle_aged	medium	no	excellent	yes
13	middle_aged	high	yes	fair	yes
14	senior	medium	no	excellent	no

- Data to be classified:

X = (age =youth, income = medium, student = yes, credit_rating = fair)

Naïve Bayes Classifier

1. **Compute class probabilities $P(C_i)$:**
   - ✓ P(buys_computer = "yes") = 9/14 = 0.643
   - ✓ P(buys_computer = "no") = 5/14= 0.357
2. **Compute $P(X|C_i)$ for each class:**
   - ✓ P(age = youth | buys_computer = "yes") = 2/9 = 0.222
   - ✓ P(age = youth | buys_computer = "no") = 3/5 = 0.6
   - ✓ P(income = "medium" | buys_computer = "yes") = 4/9 = 0.444
   - ✓ P(income = "medium" | buys_computer = "no") = 2/5 = 0.4
   - ✓ P(student = "yes" | buys_computer = "yes) = 6/9 = 0.667
   - ✓ P(student = "yes" | buys_computer = "no") = 1/5 = 0.2
   - ✓ P(credit_rating = "fair" | buys_computer = "yes") = 6/9 = 0.667
   - ✓ P(credit_rating = "fair" | buys_computer = "no") = 2/5 = 0.4

X = (age = youth, income = medium, student = yes, credit_rating = fair)

$P(X|C_i)$: P(X|buys_computer = "yes") = 0.222 x 0.444 x 0.667 x 0.667 = 0.044

P(X|buys_computer = "no") = 0.6 x 0.4 x 0.2 x 0.4 = 0.019

3. **Choose the class for which P(X|Ci)*P(Ci) is the maximum:**

$P(X|C_i)*P(C_i)$: P(X|buys_computer = "yes") * P(buys_computer = "yes") = 0.028

P(X|buys_computer = "no") * P(buys_computer = "no") = 0.007

Therefore, X belongs to class ("buys_computer = yes")

RID	age	income	student	credit_rating	Class: buys_compute
1	youth	high	no	fair	no
2	youth	high	no	excellent	no
3	middle_aged	high	no	fair	yes
4	senior	medium	no	fair	yes
5	senior	low	yes	fair	yes
6	senior	low	yes	excellent	no
7	middle_aged	low	yes	excellent	yes
8	youth	medium	no	fair	no
9	youth	low	yes	fair	yes
10	senior	medium	yes	fair	yes
11	youth	medium	yes	excellent	yes
12	middle_aged	medium	no	excellent	yes
13	middle_aged	high	yes	fair	yes
14	senior	medium	no	excellent	no

Naïve Bayes Classification

- Fast and simple
- Scales well
- Independence assumption may not hold true
  - In practice, still works quite well
- Does not model interactions between features