Classification: Bayesian Classification

Dr. Chandranath Adak

Dept. of CSE, Indian Institute of Technology Patna

September 15, 2025

Different Models of Classification

Two models of classification

- A discriminative model
  - Models the decision boundary between the classes
  - Learns the conditional probability distribution $p(y|x)$, where x is the set of features and y is the class level
  - Example: Logistic Regression
- A generative model
  - Models the actual distribution of each class
  - Learns the joint probability distribution $p(x, y)$
  - Finally, predicts the conditional probability $p(y|x)$
  - Example: Naïve Bayes Classifier
- Both of them are supervised learning techniques

Bayesian Classification

- A generative model of classification
- Few terminologies
  - $X \in \mathbb{R}^k$ is a k dimensional feature vector
  - $C = \{C_1, C_2, \dots, C_m\}$ is a set of classes
  - **Unconditional density function / Evidence** $P(X)$: Probability distribution of the feature X in the entire population
  - **Prior Probability** $P(C_i)$: Probability that a random sample belongs to class C_i
  - **Class conditional probability / Likelihood** $P(X|C_i)$: Probability of obtaining a feature value X, given that the sample is from C_i
  - Objective of classification is to measure the **conditional probability** $P(C_i|X)$: Probability of a given sample X belongs to C_i

Bayes' theorem

Bayes' theorem

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}; \quad Given \quad P(B) \neq 0$$

- Naïve Bayes classifier measures the conditional probability $P(C_i|X)$ using Bayes rules
- Assumptions:
  - The set of features are independent to each other, i.e., the value of a particular feature is independent of the value of any other feature
  - The prior probability and the class conditional probability are known or can be measured from training dataset

Bayesian Classification

$\forall\, C_i \in \{C_1, C_2, \ldots, C_m\};$

$$P(C_i|X) = P(C_i|x_1, x_2, \ldots, x_k)$$

$$P(C_i|X) = \frac{P(x_1, x_2, \ldots, x_k|C_i)P(C_i)}{P(x_1, x_2, \ldots, x_k)}$$

$$P(C_i|X) = \frac{P(x_1|C_i)P(x_2|C_i)\ldots P(x_k|C_i)P(C_i)}{P(x_1)P(x_2)\ldots P(x_k)}$$

$$P(C_i|X) = \frac{P(C_i)\prod_{j=1}^{k} P(x_j|C_i)}{\prod_{j=1}^{k} P(x_j)}$$

$$P(C_i|X) \propto P(C_i)\prod_{j=1}^{k} P(x_j|C_i)$$

Finally, $X \in C_i$, for some $C_i \in \{C_1, C_2, \ldots, C_m\}$, such that $argmax_{i=1,2,\ldots,m} P(C_i)\prod_{j=1}^{k} P(x_j|C_i)$ corresponds to C_i

An example

	Outlook	Temperature	Humidity	Windy	Play Golf
0	Rainy	Hot	High	False	No
1	Rainy	Hot	High	True	No
2	Overcast	Hot	High	False	Yes
3	Sunny	Mild	High	False	Yes
4	Sunny	Cool	Normal	False	Yes
5	Sunny	Cool	Normal	True	No
6	Overcast	Cool	Normal	True	Yes
7	Rainy	Mild	High	False	No
8	Rainy	Cool	Normal	False	Yes
9	Sunny	Mild	Normal	False	Yes
10	Rainy	Mild	Normal	True	Yes
11	Overcast	Mild	High	True	Yes
12	Overcast	Hot	Normal	False	Yes
13	Sunny	Mild	High	True	No

[1]

- In the above example, we have
  - 4 features (Outlook, Temperature, Humidity, Windy) and 2 classes (Yes, No)
  - 14 samples (each row corresponding to a sample with class level)
- Objective: Classify a sample (Sunny, Hot, Normal, False)

[1] Tom M. Mitchell, "Machine Learning", McGraw Hill, 1997.

An example (contd.)

Class Conditional Probabilities:

Outlook	Class labels		$P(Outlook\|Yes)$	$P(Outlook\|No)$
	Yes	No		
Sunny	3	2	$\frac{3}{9}$	$\frac{2}{5}$
Overcast	4	0	$\frac{4}{9}$	$\frac{0}{5}$
Rainy	3	2	$\frac{3}{9}$	$\frac{2}{5}$
Total	9	5		

Humidity	Class labels		$P(Humidity\|Yes)$	$P(Humidity\|No)$
	Yes	No		
High	3	4	$\frac{3}{9}$	$\frac{4}{5}$
Normal	6	1	$\frac{6}{9}$	$\frac{1}{5}$
Total	9	5		

Temperature	Class labels		$P(Temp.\|Yes)$	$P(Temp.\|No)$
	Yes	No		
Hot	2	2	$\frac{2}{9}$	$\frac{2}{5}$
Mild	4	2	$\frac{4}{9}$	$\frac{2}{5}$
Cool	3	1	$\frac{3}{9}$	$\frac{1}{5}$
Total	9	5		

Wind	Class labels		$P(Wind\|Yes)$	$P(Wind\|No)$
	Yes	No		
False	6	2	$\frac{6}{9}$	$\frac{2}{5}$
True	3	3	$\frac{3}{9}$	$\frac{3}{5}$
Total	9	5		

Prior probabilities:

- $P(Yes) = \frac{9}{14}$
- $P(No) = \frac{5}{14}$

An example (contd.)

Class Conditional Probabilities:

Outlook	Class labels		$P(Outlook\|Yes)$	$P(Outlook\|No)$
	Yes	No		
Sunny	3	2	$\frac{3}{9}$	$\frac{2}{5}$
Overcast	4	0	$\frac{4}{9}$	$\frac{0}{5}$
Rainy	3	2	$\frac{3}{9}$	$\frac{2}{5}$
Total	9	5		

Humidity	Class labels		$P(Humidity\|Yes)$	$P(Humidity\|No)$
	Yes	No		
High	3	4	$\frac{3}{9}$	$\frac{4}{5}$
Normal	6	1	$\frac{6}{9}$	$\frac{1}{5}$
Total	9	5		

Temperature	Class labels		$P(Temp.\|Yes)$	$P(Temp.\|No)$
	Yes	No		
Hot	2	2	$\frac{2}{9}$	$\frac{2}{5}$
Mild	4	2	$\frac{4}{9}$	$\frac{2}{5}$
Cool	3	1	$\frac{3}{9}$	$\frac{1}{5}$
Total	9	5		

Wind	Class labels		$P(Wind\|Yes)$	$P(Wind\|No)$
	Yes	No		
False	6	2	$\frac{6}{9}$	$\frac{2}{5}$
True	3	3	$\frac{3}{9}$	$\frac{3}{5}$
Total	9	5		

Prior probabilities:

- $P(Yes) = \frac{9}{14}$
- $P(No) = \frac{5}{14}$

Conditional probabilities:

- $P(Yes|(Sunny, Hot, Normal, False)) = \frac{P((Sunny, Hot, Normal, False)|Yes)P(Yes)}{P((Sunny, Hot, Normal, False))} = \frac{P(Sunny|Yes)P(Hot|Yes)P(Normal|Yes)P(False|Yes)P(Yes)}{P((Sunny, Hot, Normal, False))}$

An example (contd.)

Class Conditional Probabilities:

Outlook	Class labels		$P(Outlook\|Yes)$	$P(Outlook\|No)$
	Yes	No		
Sunny	3	2	$\frac{3}{9}$	$\frac{2}{5}$
Overcast	4	0	$\frac{4}{9}$	$\frac{0}{5}$
Rainy	3	2	$\frac{3}{9}$	$\frac{2}{5}$
Total	9	5		

Humidity	Class labels		$P(Humidity\|Yes)$	$P(Humidity\|No)$
	Yes	No		
High	3	4	$\frac{3}{9}$	$\frac{4}{5}$
Normal	6	1	$\frac{6}{9}$	$\frac{1}{5}$
Total	9	5		

Temperature	Class labels		$P(Temp.\|Yes)$	$P(Temp.\|No)$
	Yes	No		
Hot	2	2	$\frac{2}{9}$	$\frac{2}{5}$
Mild	4	2	$\frac{4}{9}$	$\frac{2}{5}$
Cool	3	1	$\frac{3}{9}$	$\frac{1}{5}$
Total	9	5		

Wind	Class labels		$P(Wind\|Yes)$	$P(Wind\|No)$
	Yes	No		
False	6	2	$\frac{6}{9}$	$\frac{2}{5}$
True	3	3	$\frac{3}{9}$	$\frac{3}{5}$
Total	9	5		

Prior probabilities:

- $P(Yes) = \frac{9}{14}$
- $P(No) = \frac{5}{14}$

Conditional probabilities:

- $P(Yes|(Sunny, Hot, Normal, False)) = \frac{P(Sunny|Yes)P(Hot|Yes)P(Normal|Yes)P(False|Yes)P(Yes)}{P((Sunny, Hot, Normal, False))}$
  $\propto (\frac{3}{9}\frac{2}{9}\frac{6}{9}\frac{6}{9}\frac{9}{14}) \approx 0.0211$

An example (contd.)

Class Conditional Probabilities:

Outlook	Class labels		$P(Outlook \mid Yes)$	$P(Outlook \mid No)$
	Yes	No		
Sunny	3	2	$\frac{3}{9}$	$\frac{2}{5}$
Overcast	4	0	$\frac{4}{9}$	$\frac{0}{5}$
Rainy	3	2	$\frac{3}{9}$	$\frac{2}{5}$
Total	9	5		

Humidity	Class labels		$P(Humidity \mid Yes)$	$P(Humidity \mid No)$
	Yes	No		
High	3	4	$\frac{3}{9}$	$\frac{4}{5}$
Normal	6	1	$\frac{6}{9}$	$\frac{1}{5}$
Total	9	5		

Temperature	Class labels		$P(Temp. \mid Yes)$	$P(Temp. \mid No)$
	Yes	No		
Hot	2	2	$\frac{2}{9}$	$\frac{2}{5}$
Mild	4	2	$\frac{4}{9}$	$\frac{2}{5}$
Cool	3	1	$\frac{3}{9}$	$\frac{1}{5}$
Total	9	5		

Wind	Class labels		$P(Wind \mid Yes)$	$P(Wind \mid No)$
	Yes	No		
False	6	2	$\frac{6}{9}$	$\frac{2}{5}$
True	3	3	$\frac{3}{9}$	$\frac{3}{5}$
Total	9	5		

Prior probabilities:

- $P(Yes) = \frac{9}{14}$
- $P(No) = \frac{5}{14}$

Conditional probabilities:

- $P(No \mid (Sunny, Hot, Normal, False)) = \frac{P((Sunny, Hot, Normal, False) \mid No)P(No)}{P((Sunny, Hot, Normal, False))} = \frac{P(Sunny \mid No)P(Hot \mid No)P(Normal \mid No)P(False \mid No)P(No)}{P((Sunny, Hot, Normal, False))}$

Class Conditional Probabilities:

Outlook	Class labels		$P(Outlook\|Yes)$	$P(Outlook\|No)$
	Yes	No		
Sunny	3	2	$\frac{3}{9}$	$\frac{2}{5}$
Overcast	4	0	$\frac{4}{9}$	$\frac{0}{5}$
Rainy	3	2	$\frac{3}{9}$	$\frac{2}{5}$
Total	9	5		

Humidity	Class labels		$P(Humidity\|Yes)$	$P(Humidity\|No)$
	Yes	No		
High	3	4	$\frac{3}{9}$	$\frac{4}{5}$
Normal	6	1	$\frac{6}{9}$	$\frac{1}{5}$
Total	9	5		

Temperature	Class labels		$P(Temp.\|Yes)$	$P(Temp.\|No)$
	Yes	No		
Hot	2	2	$\frac{2}{9}$	$\frac{2}{5}$
Mild	4	2	$\frac{4}{9}$	$\frac{2}{5}$
Cool	3	1	$\frac{3}{9}$	$\frac{1}{5}$
Total	9	5		

Wind	Class labels		$P(Wind\|Yes)$	$P(Wind\|No)$
	Yes	No		
False	6	2	$\frac{6}{9}$	$\frac{2}{5}$
True	3	3	$\frac{3}{9}$	$\frac{3}{5}$
Total	9	5		

Prior probabilities:

- $P(Yes) = \frac{9}{14}$
- $P(No) = \frac{5}{14}$

Conditional probabilities:

- $P(No|(Sunny, Hot, Normal, False)) = \frac{P(Sunny|No)P(Hot|No)P(Normal|No)P(False|No)P(No)}{P((Sunny, Hot, Normal, False))}$
  $\propto (\frac{2}{5}\frac{2}{5}\frac{1}{5}\frac{2}{5}\frac{5}{14}) \approx 0.0045$

An example (contd.)

Class Conditional Probabilities:

Outlook	Class labels		$P(Outlook\|Yes)$	$P(Outlook\|No)$
	Yes	No		
Sunny	3 2	2 3	$\frac{2}{9}$ 3/9	$\frac{3}{5}$ 2/5
Overcast	4	0	$\frac{4}{9}$	$\frac{0}{5}$
Rainy	3	2	$\frac{3}{9}$	$\frac{2}{5}$
Total	9	5		

Humidity	Class labels		$P(Humidity\|Yes)$	$P(Humidity\|No)$
	Yes	No		
High	3	4	$\frac{3}{9}$	$\frac{4}{5}$
Normal	6	1	$\frac{6}{9}$	$\frac{1}{5}$
Total	9	5		

Temperature	Class labels		$P(Temp.\|Yes)$	$P(Temp.\|No)$
	Yes	No		
Hot	2	2	$\frac{2}{9}$	$\frac{2}{5}$
Mild	4	2	$\frac{4}{9}$	$\frac{2}{5}$
Cool	3	1	$\frac{3}{9}$	$\frac{1}{5}$
Total	9	5		

Wind	Class labels		$P(Wind\|Yes)$	$P(Wind\|No)$
	Yes	No		
False	6	2	$\frac{6}{9}$	$\frac{2}{5}$
True	3	3	$\frac{3}{9}$	$\frac{3}{5}$
Total	9	5		

Prior probabilities:

- $P(Yes) = \frac{9}{14}$
- $P(No) = \frac{5}{14}$

Conditional probabilities:

- $P(Yes|(Sunny, Hot, Normal, False)) \propto 0.0211$
- $P(No|(Sunny, Hot, Normal, False)) \propto 0.0045$
- The sample is classified in *Yes* class

Thank You!