Naïve Bayes Example

Recall: Naïve Bayes Classifier

Returns a **MAP prediction**, where the posterior (conditional) probabilities are computed under the **strong assumption of conditional independence** between features given the class label.

$$c_{MAP} = \text{argmax}_{c \in Classes} P(c) \prod_{i=1}^{d} P(x_i|c)$$

Day	Outlook	Humidity	Wind	Play
D1	Sunny	High	Weak	No
D2	Sunny	High	Strong	No
D3	Overcast	High	Weak	Yes
D4	Rain	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	High	Weak	No
D9	Sunny	Normal	Weak	Yes
D10	Rain	Normal	Weak	Yes
D11	Sunny	Normal	Strong	Yes
D12	Overcast	High	Strong	Yes
D13	Overcast	Normal	Weak	Yes
D14	Rain	High	Strong	No

Let's start by defining our random variables:

- T → Play Tennis
- W → Wind
- H → Humidity
- O → Outlook

Day	Outlook	Humidity	Wind	Play
D1	Sunny	High	Weak	No
D2	Sunny	High	Strong	No
D3	Overcast	High	Weak	Yes
D4	Rain	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	High	Weak	No
D9	Sunny	Normal	Weak	Yes
D10	Rain	Normal	Weak	Yes
D11	Sunny	Normal	Strong	Yes
D12	Overcast	High	Strong	Yes
D13	Overcast	Normal	Weak	Yes
D14	Rain	High	Strong	No

$$P(T = Yes) = \frac{9}{14}$$

$$P(T = No) = \frac{5}{14}$$

Day	Outlook	Humidity	Wind	Play
D1	Sunny	High	Weak	No
D2	Sunny	High	Strong	No
D3	Overcast	High	Weak	Yes
D4	Rain	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	High	Weak	No
D9	Sunny	Normal	Weak	Yes
D10	Rain	Normal	Weak	Yes
D11	Sunny	Normal	Strong	Yes
D12	Overcast	High	Strong	Yes
D13	Overcast	Normal	Weak	Yes
D14	Rain	High	Strong	No

$$P(W = Weak \mid T = Yes) = \frac{6}{9}$$

$$P(W = Strong \mid T = Yes) = \frac{3}{9}$$

Day	Outlook	Humidity	Wind	Play
D1	Sunny	High	Weak	No
D2	Sunny	High	Strong	No
D3	Overcast	High	Weak	Yes
D4	Rain	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	High	Weak	No
D9	Sunny	Normal	Weak	Yes
D10	Rain	Normal	Weak	Yes
D11	Sunny	Normal	Strong	Yes
D12	Overcast	High	Strong	Yes
D13	Overcast	Normal	Weak	Yes
D14	Rain	High	Strong	No

$$P(W = Weak \mid T = Yes) = \frac{6}{9} \qquad P(W = Weak \mid T = No) = \frac{2}{5}$$

$$P(W = Strong \mid T = Yes) = \frac{3}{9} \qquad P(W = Strong \mid T = No) = \frac{3}{5}$$

Day	Outlook	Humidity	Wind	Play
D1	Sunny	High	Weak	No
D2	Sunny	High	Strong	No
D3	Overcast	High	Weak	Yes
D4	Rain	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	High	Weak	No
D9	Sunny	Normal	Weak	Yes
D10	Rain	Normal	Weak	Yes
D11	Sunny	Normal	Strong	Yes
D12	Overcast	High	Strong	Yes
D13	Overcast	Normal	Weak	Yes
D14	Rain	High	Strong	No

$$P(H = High \mid T = Yes) = \frac{3}{9}$$

$$P(H = Normal \mid T = Yes) = \frac{6}{9}$$

Day	Outlook	Humidity	Wind	Play
D1	Sunny	High	Weak	No
D2	Sunny	High	Strong	No
D3	Overcast	High	Weak	Yes
D4	Rain	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	High	Weak	No
D9	Sunny	Normal	Weak	Yes
D10	Rain	Normal	Weak	Yes
D11	Sunny	Normal	Strong	Yes
D12	Overcast	High	Strong	Yes
D13	Overcast	Normal	Weak	Yes
D14	Rain	High	Strong	No

$$P(H = High \mid T = Yes) = \frac{3}{9} \qquad P(H = High \mid T = No) = \frac{4}{5}$$

$$P(H = Normal \mid T = Yes) = \frac{6}{9} \qquad P(H = Normal \mid T = No) = \frac{1}{5}$$

Day	Outlook	Humidity	Wind	Play
D1	Sunny	High	Weak	No
D2	Sunny	High	Strong	No
D3	Overcast	High	Weak	Yes
D4	Rain	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	High	Weak	No
D9	Sunny	Normal	Weak	Yes
D10	Rain	Normal	Weak	Yes
D11	Sunny	Normal	Strong	Yes
D12	Overcast	High	Strong	Yes
D13	Overcast	Normal	Weak	Yes
D14	Rain	High	Strong	No

$$P(O = Overcast \mid T = Yes) = \frac{4}{9}$$

$$P(O = Sunny \mid T = Yes) = \frac{2}{9}$$

$$P(O = Rain \mid T = Yes) = \frac{3}{9}$$

Day	Outlook	Humidity	Wind	Play
D1	Sunny	High	Weak	No
D2	Sunny	High	Strong	No
D3	Overcast	High	Weak	Yes
D4	Rain	High	Weak	Yes
D5	Rain	Normal	Weak	Yes
D6	Rain	Normal	Strong	No
D7	Overcast	Normal	Strong	Yes
D8	Sunny	High	Weak	No
D9	Sunny	Normal	Weak	Yes
D10	Rain	Normal	Weak	Yes
D11	Sunny	Normal	Strong	Yes
D12	Overcast	High	Strong	Yes
D13	Overcast	Normal	Weak	Yes
D14	Rain	High	Strong	No

$$P(O = Overcast \mid T = Yes) = \frac{4}{9} \qquad P(O = Overcast \mid T = No) = \frac{0}{5}$$

$$P(O = Sunny \mid T = Yes) = \frac{2}{9} \qquad P(O = Sunny \mid T = No) = \frac{3}{5}$$

$$P(O = Rain \mid T = Yes) = \frac{3}{9} \qquad P(O = Rain \mid T = No) = \frac{2}{5}$$

Summary of Stats

$$P(T = Yes) = \frac{9}{14}$$

$$P(T = No) = \frac{5}{14}$$

$$P(W = Weak \mid T = Yes) = \frac{6}{9}$$

$$P(W = Strong \mid T = Yes) = \frac{3}{9}$$

$$P(W = Weak \mid T = No) = \frac{2}{5}$$

$$P(W = Strong \mid T = No) = \frac{3}{5}$$

$$P(H = High \mid T = Yes) = \frac{3}{9}$$

$$P(H = Normal \mid T = Yes) = \frac{6}{9}$$

$$P(H = High \mid T = No) = \frac{4}{5}$$

$$P(H = Normal \mid T = No) = \frac{1}{5}$$

$$P(O = Overcast \mid T = Yes) = \frac{4}{9}$$

$$P(O = Sunny \mid T = Yes) = \frac{2}{9}$$

$$P(O = Rain \mid T = Yes) = \frac{3}{9}$$

$$P(O = Overcast \mid T = No) = \frac{0}{5}$$

$$P(O = Sunny \mid T = No) = \frac{3}{5}$$

$$P(O = Rain \mid T = No) = \frac{2}{5}$$

Exercise Answers

- Will John play tennis if the outlook is sunny, humidity is high, and wind is strong?

$P(T = Yes \mid O = Sunny, H = High, W = Strong)$

$= P(T = Yes)P(O = Sunny \mid T = Yes)P(H = High \mid T = Yes)P(W = Strong \mid T = Yes)$

$= \left(\frac{9}{14}\right)\left(\frac{2}{9}\right)\left(\frac{3}{9}\right)\left(\frac{3}{9}\right)$ = 0.016

$P(T = No \mid O = Sunny, H = High, W = Strong)$

$= P(T = No)P(O = Sunny \mid T = No)P(H = High \mid T = No)P(W = Strong \mid T = No)$

$= \left(\frac{5}{14}\right)\left(\frac{3}{5}\right)\left(\frac{4}{5}\right)\left(\frac{3}{5}\right)$ = 0.103

Exercise Answers

- Will John play tennis if the outlook is sunny, humidity is high, and wind is strong?

$P(T = Yes \mid O = Sunny, H = High, W = Strong)$

$= P(T = Yes)P(O = Sunny \mid T = Yes)P(H = High \mid T = Yes)P(W = Strong \mid T = Yes)$

$= \left(\frac{9}{14}\right)\left(\frac{2}{9}\right)\left(\frac{3}{9}\right)\left(\frac{3}{9}\right) = 0.016$

$P(T = No \mid O = Sunny, H = High, W = Strong)$

$= P(T = No)P(O = Sunny \mid T = No)P(H = High \mid T = No)P(W = Strong \mid T = No)$

$= \left(\frac{5}{14}\right)\left(\frac{3}{5}\right)\left(\frac{4}{5}\right)\left(\frac{3}{5}\right) = 0.103$

Answer: No, John will not play tennis.

Exercise Answers

- Will John play tennis when it's rainy, highly humid, and not windy?

$P(T = Yes \mid O = Rain, H = High, W = Weak)$

$= P(T = Yes)P(O = Rain \mid T = Yes)P(H = High \mid T = Yes)P(W = Weak \mid T = Yes)$

$= \left(\frac{9}{14}\right)\left(\frac{3}{9}\right)\left(\frac{3}{9}\right)\left(\frac{6}{9}\right) = 0.048$

$P(T = No \mid O = Rain, H = High, W = Weak)$

$= P(T = No)P(O = Rain \mid T = No)P(H = High \mid T = No)P(W = Weak \mid T = No)$

$= \left(\frac{5}{14}\right)\left(\frac{2}{5}\right)\left(\frac{4}{5}\right)\left(\frac{2}{5}\right) = 0.046$

Exercise Answers

- Will John play tennis when it's rainy, highly humid, and not windy?

$P(T = Yes \,|O = Rain, H = High, W = Weak)$

$= P(T = Yes)P(O = Rain|T = Yes)P(H = High \,|T = Yes)P(W = Weak|T = Yes)$

$= \left(\frac{9}{14}\right)\left(\frac{3}{9}\right)\left(\frac{3}{9}\right)\left(\frac{6}{9}\right) = 0.048$ Answer: Yes, John will play tennis.

$P(T = No \,|O = Rain, H = High, W = Weak)$

$= P(T = No)P(O = Rain|T = No)P(H = High \,|T = No)P(W = Weak|T = No)$

$= \left(\frac{5}{14}\right)\left(\frac{2}{5}\right)\left(\frac{4}{5}\right)\left(\frac{2}{5}\right) = 0.046$