

IA : Machine (Deep) Learning

Méthode (naïve) de Bayes

A. Zemmari

LaBRI - Université de Bordeaux

Jan. 09, 2018

Un Exemple

S(ex)	H(eight) (m)	W(eight) (kg)	F(oot) size (cm)
M	1.82	82	30
M	1.80	86	28
M	1.70	77	30
M	1.80	75	25
F	1.52	45	15
F	1.65	68	20
F	1.68	59	18
F	1.75	68	23

► Question : Est-ce que (1.81, 59, 21) est M ou F ?

Un Exemple

S(ex)	H(eight) (m)	W(eight) (kg)	F(oot) size (cm)
M	1.82	82	30
M	1.80	86	28
M	1.70	77	30
M	1.80	75	25
F	1.52	45	15
F	1.65	68	20
F	1.68	59	18
F	1.75	68	23

- Question : Est-ce que (1.81, 59, 21) est M ou F ?

Autrement dit :

$$\begin{aligned} & \mathbb{P}r(S = M \mid (H, W, F) = (1.81, 59, 21)) \\ & > \mathbb{P}r(S = F \mid (H, W, F) = (1.81, 59, 21)) ? (1) \end{aligned}$$

Rappel : Formmule de Bayes

$$\mathbb{P}r(S = M \mid H, W, F) = \frac{\mathbb{P}r(S) \times \mathbb{P}r(H, W, F \mid S)}{\mathbb{P}r(H, W, F)}$$

Or $\mathbb{P}r(H, W, F)$ ne dépend pas de S , donc (1) devient :

$$\begin{aligned} & \mathbb{P}r(S = M) \times \mathbb{P}r(H, W, F \mid S = M) \\ & > \mathbb{P}r(S = F) \times \mathbb{P}r(H, W, F \mid S = F) \end{aligned}$$

Autrement dit :

$$\begin{aligned} & \mathbb{P}r(S = M \mid (H, W, F) = (1.81, 59, 21)) \\ & > \mathbb{P}r(S = F \mid (H, W, F) = (1.81, 59, 21)) ? \quad (1) \end{aligned}$$

Rappel : Formmule de Bayes

$$\mathbb{P}r(S = M \mid H, W, F) = \frac{\mathbb{P}r(S) \times \mathbb{P}r(H, W, F \mid S)}{\mathbb{P}r(H, W, F)}$$

Or $\mathbb{P}r(H, W, F)$ ne dépend pas de S , donc (1) devient :

$$\begin{aligned} & \mathbb{P}r(S = M) \times \mathbb{P}r(H, W, F \mid S = M) \\ & > \mathbb{P}r(S = F) \times \mathbb{P}r(H, W, F \mid S = F) \end{aligned}$$

Autrement dit :

$$\begin{aligned} & \mathbb{P}r(S = M \mid (H, W, F) = (1.81, 59, 21)) \\ & > \mathbb{P}r(S = F \mid (H, W, F) = (1.81, 59, 21)) ? \quad (1) \end{aligned}$$

Rappel : Formmule de Bayes

$$\mathbb{P}r(S = M \mid H, W, F) = \frac{\mathbb{P}r(S) \times \mathbb{P}r(H, W, F \mid S)}{\mathbb{P}r(H, W, F)}$$

Or $\mathbb{P}r(H, W, F)$ ne dépend pas de S , donc (1) devient :

$$\begin{aligned} & \mathbb{P}r(S = M) \times \mathbb{P}r(H, W, F \mid S = M) \\ & > \mathbb{P}r(S = F) \times \mathbb{P}r(H, W, F \mid S = F) \end{aligned}$$

- ▶ On peut facilement estimer $\mathbb{P}r(S)$.
- ▶ Mais qu'en est-il de $\mathbb{P}r(H, W, F \mid S)$?

Hypothèse "Naive Bayes"

On a

$$\mathbb{P}r(H, W, F | S) = \mathbb{P}r(H | S) \times \mathbb{P}r(W | S, H) \times \mathbb{P}r(F | S, W, H)$$

Naive Bayes :

$$\mathbb{P}r(W | S, H) = \mathbb{P}r(W | S) \text{ et } \mathbb{P}r(F | S, W, H) = \mathbb{P}r(F | S)$$

Hypothèse "Naive Bayes"

On a

$$\mathbb{P}r(H, W, F \mid S) = \mathbb{P}r(H \mid S) \times \mathbb{P}r(W \mid S, H) \times \mathbb{P}r(F \mid S, W, H)$$

Naive Bayes :

$$\mathbb{P}r(W \mid S, H) = \mathbb{P}r(W \mid S) \text{ et } \mathbb{P}r(F \mid S, W, H) = \mathbb{P}r(F \mid S)$$

Hypothèse "Naive Bayes"

Qu'est-ce que cela veut dire ?

$$\mathbb{P}r(W \mid S, H) = \mathbb{P}r(W \mid S)$$

"Parmi les hommes, le poids est indépendant de la taille"

"Naive Bayes" de manière plus formelle

$$\mathbb{P}r(Y | X_1, \dots, X_n) = \frac{\mathbb{P}r(Y) \times \mathbb{P}r(X_1, \dots, X_n | Y)}{\mathbb{P}r(X_1, \dots, X_n)}$$

Hypothèse naïve d'indépendance conditionnelle :

$$\forall i \neq j, \mathbb{P}r(X_i | Y, X_j) = \mathbb{P}r(X_i | Y)$$

$$\Rightarrow \mathbb{P}r(Y | X_1, \dots, X_n) = \frac{1}{Z} \times \mathbb{P}r(Y) \times \prod_{i=1}^n \mathbb{P}r(X_i | Y)$$

Retour à l'exemple

$$\mathbb{P}r(S \mid H, W, F) = \frac{1}{Z} \times \mathbb{P}r(S) \times \mathbb{P}r(H \mid S) \times \mathbb{P}r(W \mid S) \times \mathbb{P}r(F \mid S)$$

S(ex)	H(eight) (m)	W(eight) (kg)	F(oot) size (cm)
M	1.82	82	30
M	1.80	86	28
M	1.70	77	30
M	1.80	75	25
F	1.52	45	15
F	1.65	68	20
F	1.68	59	18
F	1.75	68	23

$$\mathbb{P}r(S = M) = ? \quad (1)$$

$$\mathbb{P}r(H = 1.81 \mid S = M) = ?$$

$$\mathbb{P}r(W = 59 \mid S = M) = ?$$

$$\mathbb{P}r(F = 21 \mid S = M) = ?$$

Retour à l'exemple

$$\mathbb{P}r(S \mid H, W, F) = \frac{1}{Z} \times \mathbb{P}r(S) \times \mathbb{P}r(H \mid S) \times \mathbb{P}r(W \mid S) \times \mathbb{P}r(F \mid S)$$

S est discrète, H, W et F sont supposée gaussiennes

S	$\hat{p}_S$	$\hat{\mu}_{H S}$	$\hat{\sigma}_{H S}$	$\hat{\mu}_{W S}$	$\hat{\sigma}_{W S}$	$\hat{\mu}_{F S}$	$\hat{\sigma}_{F S}$
M	0.5	1.78	0.0469	80	4.3012	28.25	2.0463
F	0.5	1.65	0.0834	60	9.4074	19	2.9154

$$\begin{aligned}\mathbb{P}r(S = M \mid 1.81, 59, 21) &= \frac{1}{Z} \times 0.5 \times \frac{e^{-\frac{(1.78-1.81)^2}{2 \cdot 0.0469^2}}}{\sqrt{2\pi}0.0469} \times \frac{e^{-\frac{(80-59)^2}{2 \cdot 4.3012^2}}}{\sqrt{2\pi}4.3012} \times \frac{e^{-\frac{(28.25-21)^2}{2 \cdot 2.0463^2}}}{\sqrt{2\pi}2.0463} \\ &= \frac{1}{Z} \times 7.854 \cdot 10^{-10}\end{aligned}$$

$$\mathbb{P}r(S = F \mid 1.81, 59, 21) = \frac{1}{Z} \times 1.730 \cdot 10^{-3}$$

Retour à l'exemple

$$\mathbb{P}r(S \mid H, W, F) = \frac{1}{Z} \times \mathbb{P}r(S) \times \mathbb{P}r(H \mid S) \times \mathbb{P}r(W \mid S) \times \mathbb{P}r(F \mid S)$$

Conclusion : étant donné les données dont on dispose, Il est plus probable que $(1.81m, 59kg, 21cm)$ soit une femme

$$\mathbb{P}r(Y | X_1, \dots, X_n) = \frac{1}{Z} \times \mathbb{P}r(Y) \times \prod_{i=1}^n \mathbb{P}r(X_i | Y)$$

- ▶ X_i est continue $\rightarrow X_i | Y = y \sim \mathcal{N}(\mu_{iy}, \sigma_{iy})$
- ▶ X_i est binaire, $\rightarrow X_i | Y = y \sim \mathcal{B}(p_{iy})$

L'algorithme

- ▶ Entrainement :

Pour toutes les valeurs possibles de Y et de X_i , calculer $\hat{\mathbb{P}}_r(Y = y)$ et $\hat{\mathbb{P}}_r(X_i = x_i \mid Y = y)$

- ▶ Prédiction : étant donné $(x_1, x_2, \dots, x_n)$, retourner y qui maximise

$$\hat{\mathbb{P}}_r(Y = y) \hat{\mathbb{P}}_r(X_i = x_i \mid Y = y)$$