

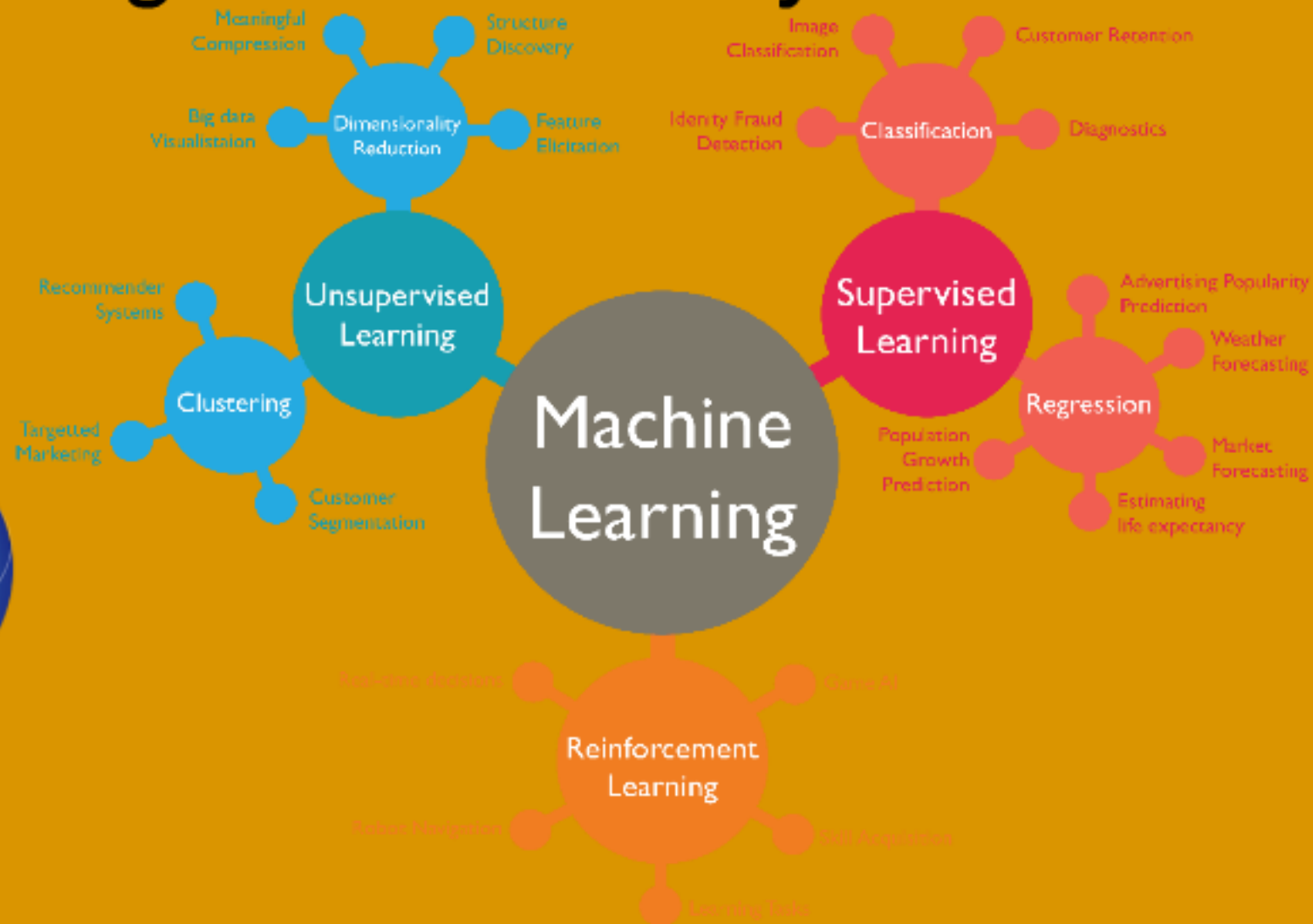
Machine learning

Classifieur Bayesien Naïf

Fatima YUCEF

Master 2 ADSI

Qu'elle est la relation entre le machine learning et classifieur bayésien naïf ?



hypothèse!

les descripteurs (X_j) sont deux à deux indépendants conditionnellement aux valeurs de la variable à prédire (Y).





chills	runny nose	headache	fever	flu?
Y	N	Mild	Y	N
Y	Y	No	N	Y
Y	N	Strong	Y	Y
N	Y	Mild	Y	Y
N	N	No	N	N
N	Y	Strong	Y	Y+
N	Y	Strong	N	N
Y	Y	Mild	Y	Y

Do I believe that a patient with the following symptoms has the flu?

chills	runny nose	headache	fever	flu?
Y	N	Mild	Y	?

Do I believe that a patient with the following symptoms has the flu?

chills	runny nose	headache	fever	flu?
Y	N	Mild	Y	?

$P(\text{Flu}=Y)$	0.625	$P(\text{Flu}=N)$	0.375
$P(\text{chills}=Y \text{flu}=Y)$	0.6	$P(\text{chills}=Y \text{flu}=N)$	0.333
$P(\text{chills}=N \text{flu}=Y)$	0.4	$P(\text{chills}=N \text{flu}=N)$	0.666
$P(\text{runny nose}=Y \text{flu}=Y)$	0.8	$P(\text{runny nose}=Y \text{flu}=N)$	0.333
$P(\text{runny nose}=N \text{flu}=Y)$	0.2	$P(\text{runny nose}=N \text{flu}=N)$	0.666
$P(\text{headache}=Mild \text{flu}=Y)$	0.4	$P(\text{headache}=Mild \text{flu}=N)$	0.333
$P(\text{headache}=No \text{flu}=Y)$	0.2	$P(\text{headache}=No \text{flu}=N)$	0.3333
$P(\text{headache}=Strong \text{flu}=Y)$	0.4	$P(\text{headache}=Strong \text{flu}=N)$	0.333
$P(\text{fever}=Y \text{flu}=Y)$	0.8	$P(\text{fever}=Y \text{flu}=N)$	0.333
$P(\text{fever}=N \text{flu}=Y)$	0.2	$P(\text{fever}=N \text{flu}=N)$	0.666

And then decide:

$\text{argmax} P(\text{flu} = Y)P(\text{chills} = Y|\text{flu} = Y)P(\text{runny nose} = N|\text{flu} = Y)P(\text{headache} = Mild|\text{flu} = Y)P(\text{fever} = N|\text{flu} = Y)$
 $= 0.006$

vs.

$\text{argmax} P(\text{flu} = N)P(\text{chills} = Y|\text{flu} = N)P(\text{runny nose} = N|\text{flu} = N)P(\text{headache} = Mild|\text{flu} = N)P(\text{fever} = N|\text{flu} = N)$
 $= 0.0185$

GAUSSIAN NAIVE BAYES CLASSIFIER

"Gaussian" because this is a normal distribution

This is our prior belief

$$P(\text{class} | \text{data}) = \frac{P(\text{data} | \text{class}) \times P(\text{class})}{P(\text{data})}$$

We don't calculate this in naive bayes classifiers

ChrisAlbon

$$\hat{y}(\omega) = y_{k^*} \Leftrightarrow y_{k^*} = \arg \max_k P[Y = y_k / \aleph(\omega)]$$

$$\hat{y}(\omega) = y_{k^*} \Leftrightarrow y_{k^*} = \arg \max_k P(Y = y_k) \times P[\aleph(\omega) / Y = y_k]$$





TEXT CLASSIFICATION IMPLEMENTATION

Texte

Catégorie

je fais la natation

sport

je cours chaque matin

sport

elle est occupé au
travail

pas de sport

Elle est fais du fitnes

sport

il part en voyage

pas de sport

Prédire

je cours chaque jeudi

$$P(\text{je} / \text{sport}) = 2/11$$

$$P(\text{cours} / \text{sport}) = 1/11$$

$$P(\text{chaque} / \text{sport}) = 1/11$$

$$P(\text{jeudi} / \text{sport}) = ?$$

Si n_k est le nombre d'individu de la modalité y_k dans un échantillon de n observations, nous utilisons

$$\hat{P}(Y = y_k) = p_k = \frac{n_k + \lambda}{n + \lambda \times K}$$

Lorsque $\lambda = 0$, nous avons la fréquence relative usuelle. Lorsque nous fixons $\lambda = 1$, nous obtenons l'estimateur laplacien des probabilités.

$$P(\text{jeudi} / \text{sport}) = \frac{0 + 1}{11 + 1 \times 14}$$



Thank you!

