

Devoir 1 - Partie théorique

$$a) P(X|Y) = \frac{P(X \cap Y)}{P(Y)}$$

$$b) \underbrace{\frac{1}{5} \cdot \frac{4}{5}}_{\text{pile} \rightarrow \text{face}} + \underbrace{\frac{4}{5} \cdot \frac{1}{5}}_{\text{face} \rightarrow \text{pile}} = \frac{8}{25}$$

$$c) \text{ i) } P(Y|X) = \frac{P(X,Y)}{P(X)} \Rightarrow \overset{(1)}{P(Y|X)P(X) = P(X,Y)}$$

$$d) \text{ ii) } P(X|Y) = \frac{P(X,Y)}{P(Y)} \Rightarrow \overset{(2)}{P(X,Y) = P(X|Y) \cdot P(Y)}$$

$$d) P(X|Y) = \frac{\overset{(2)}{P(X,Y)}}{P(Y)} = \frac{\overset{(1)}{P(Y|X)P(X)}}{P(Y)}$$

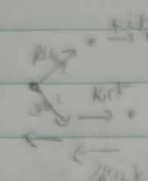
$$e) \text{ i) } 1 - 0.56 = 0.44$$

$$\text{ii) } \frac{P(Y|X)P(X)}{P(Y)} = \frac{0.4 \cdot 0.44}{0.56} = 0.309859$$

2. a) $12/1000$

b) $E = P(\text{congress} / \text{right: "sport"}) \cdot \text{occurrences}$
 $= \frac{4}{1000} \cdot 1000 = 4$

c) $\frac{1}{3} \cdot \frac{4}{100} + \frac{2}{3} \cdot \frac{12}{1000} = \frac{64}{3000} = \frac{8}{375} \approx \frac{7}{750}$

d)  $P(\text{kick} / \text{sport}) = \frac{P(\text{sport} / \text{kick}) \cdot P(\text{kick})}{P(\text{sport})}$

en baseant c, $\Rightarrow \frac{P(\text{kick} / \text{sport}) \cdot P(\text{sport})}{P(\text{kick})} = P(\text{sport} / \text{kick})$
 terme de Bayes

même formule $= \frac{\frac{15}{100} \cdot \frac{1}{3}}{\frac{15}{100} \cdot \frac{1}{3} + \frac{8}{1000} \cdot \frac{2}{3}} \approx 0.97$

e)

$P(\text{Sport} / \text{congress}) = \frac{\frac{4}{1000} \cdot \frac{1}{3}}{\frac{4}{1000} \cdot \frac{1}{3} + \frac{6}{100} \cdot \frac{2}{3}} = \frac{1}{31}$

$P(\text{politique} / \text{congress}) = \frac{\frac{6}{100} \cdot \frac{2}{3}}{\frac{4}{1000} \cdot \frac{1}{3} + \frac{6}{100} \cdot \frac{2}{3}} = \frac{30}{31}$

$P(\text{goal} / \text{congress}) = P(\text{Sport} / \text{congress}) \cdot P(\text{goal} / \text{Sport}) + P(\text{politique} / \text{congress}) \cdot P(\text{goal} / \text{politique})$
 $= \frac{4}{1000} \cdot \frac{1}{31} + \frac{12}{1000} \cdot \frac{30}{31} = \frac{2}{155}$

f). Je prendrais tous les rangs du sujet visé,
 puis je compte les occurrences du mot
 dans ce sujet et le nombre total de mots.

Pour les probabilités de sujets. On
 compte les sujets et on les divise
 par la gte totale de rangs.

3.

a)
$$\prod_{i=1}^n f_{\theta}(x_i)$$

b) On commence par remarquer que si un $x \notin \Theta$, alors tous les

termes seront absorbés par 0. Il faut donc

s'assurer que $\forall x: x \leq \theta$. x_{\max} est donc
 le θ le plus petit qui ne réduit pas $\prod_{i=1}^n f_{\theta}(x_i)$ à 0.

C'est évident car $\forall x: x_i \leq x_{\max}$.

On veut aussi minimiser θ car c'est $\frac{1}{\theta}$, inversement

proportionnel. Sachant que x_{\max} est le θ le

plus petit qu'on peut poser le maximisateur
 de vraisemblance est x_{\max}

4.

$$b) y=1 \quad \bar{x}_1 = \frac{5+2+3}{3} = 3 \quad \bar{x}_2 = \frac{-3+7}{3} = -4$$

$$y=2 \quad \bar{x}_1 = \frac{18}{3} = 6 \quad \bar{x}_2 = \frac{10}{3}$$

$$c) y=3 \quad \bar{x}_1 = \frac{5}{3} \quad \bar{x}_2 = \frac{-2}{3}$$

$$a) E[1_{\{x \in S\}}] = P(x \in S) \cdot \underset{\substack{\uparrow \\ 1}}{1_{\{x \in S\}}} + (1 - P(x \in S)) \cdot \underset{\substack{\uparrow \\ 0}}{1_{\{x \notin S\}}} \\ = P(x \in S)$$

b)

$$m^d = 2^{4611}$$

d) nombre d'interval de gains perissure =

$$\text{on commence à 10 donc } \frac{0.8}{0.05} = 16$$

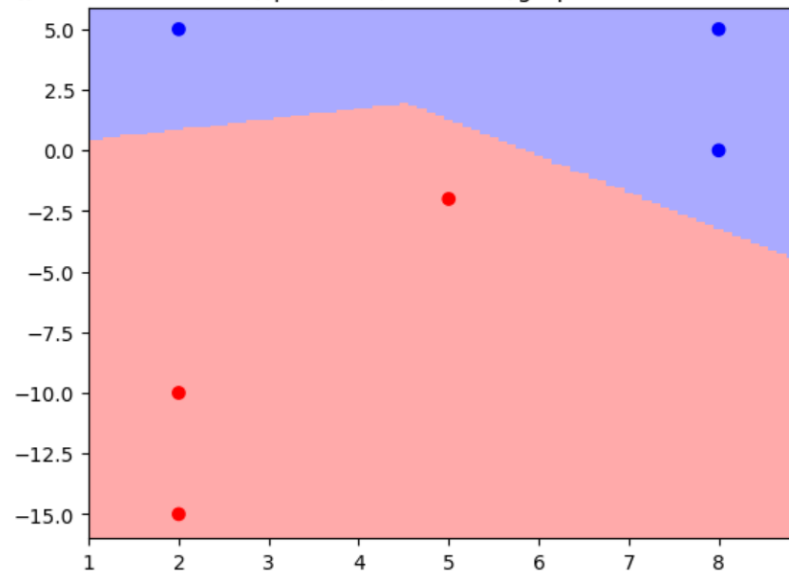
$$16 \cdot 4 = 64$$

↑
points needed for increment

$$e) \left(1 - \frac{1}{m^d}\right)^n$$

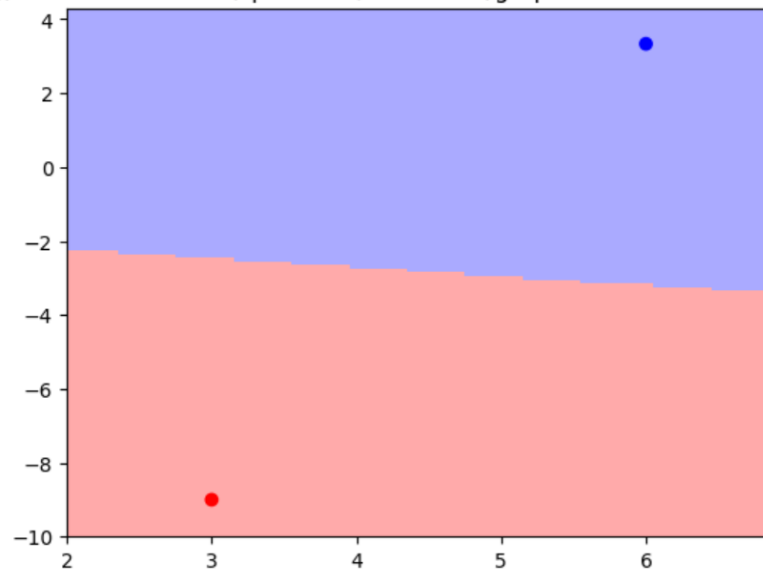
Plus proche voisin

CODE DE <https://stackoverflow.com/questions/45075638/graph-k-nn-decision-boundaries-in-matplotlib>



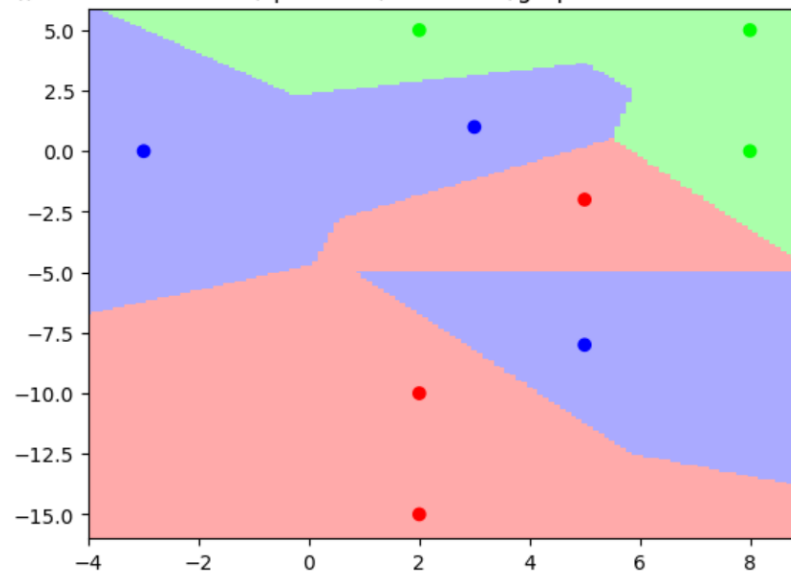
Plus proche point moyen

CODE DE <https://stackoverflow.com/questions/45075638/graph-k-nn-decision-boundaries-in-matplotlib>



Avec une 3eme classe

CODE DE <https://stackoverflow.com/questions/45075638/graph-k-nn-decision-boundaries-in-matplotlib>



Avec une 3 eme classe, avec points moyen moyen

CODE DE <https://stackoverflow.com/questions/45075638/graph-k-nn-decision-boundaries-in-matplotlib>

