TECHNIQUES D'APPRENTISSAGE

MASTER UDES

IFT 712 : Devoir 1

JOUFFROY Emma || ADOLPHE Maxime

Mr. JODOIN Pierre-Marc

Septembre 2019

1 Question 1

[1 point] : Démontrez la propriété de l'entropie suivante :

$$H\left[x,y\right] = H\left[y|x\right] + H\left[x\right]$$

D'après la définition de l'entropie :

$$H\left[x,y\right] = -\sum_x \sum_y p[x,y] log_2(p[x,y])$$

D'après la formule de Bayes :

$$H\left[x,y\right] = -\sum_x \sum_y p[x,y] log_2(p[y|x]p[x])$$

$$H\left[x,y\right] = -\sum_x \sum_y p[x,y] log_2(p[y|x]) - \sum_x \sum_y p[x,y] log_2(p[x])$$

D'après la formule de l'entropie conditionnelle :

$$H\left[x,y\right] = H[y|x] - \sum_x \sum_y p[x,y] log_2 p[x]$$

$$H\left[x,y\right] = H[y|x] - \sum_x log_2 p[x] \sum_y p[x,y]$$

D'après la formule de Bayes :

$$H\left[x,y\right] = H[y|x] - \sum_x log_2 p[x] \sum_y p[x]p[y,x]$$

$$H\left[x,y\right] = H[y|x] - \sum_x p[x] log_2 p[x] \sum_y p[y,x]$$

Comme :

$$\sum_y p[y,x] = 1$$

Alors :

$$H\left[x,y\right] = H\left[y|x\right] + H\left[x\right]$$

2 Question 2

[2 points] : Démontrez la propriété de l'information mutuelle suivante :

$$I\left[x,y\right] = H\left[x\right] - H\left[x|y\right]$$

D'après la définition de l'information mutuelle :

$$I[x,y] = KL[p[x,y]||p[x]p[y]]$$

D'après la définition de la divergence de Kullback-Leibder :

$$I[x,y] = \sum_x \sum_y p[x,y] log_2(\frac{p[x,y]}{p[x]p[y]})$$

D'après les propriétés de la fonction logarithme et la définition de l'entropie :

$$I[x,y] = -H[x,y] - \sum_x \sum_y p[x,y] log_2(p[x]p[y])$$

$$I[x,y] = -\sum_x p[x] log_2(p[x]) \sum_y p[y|x] - \sum_y p[y] log_2(p[y]) \sum_x p[x|y] - H[x,y]$$

D'après la formule démontrée précédemment (question 1) et la définition de l'entropie :

$$I[x,y] = H[x] + H[y] - H[x|y] - H[y]$$

Finalement :

$$I[x,y] = H[x] - H[x|y]$$

3 Question 3

[1 point] : Démontrez la propriété suivante :

$$cov\left[x,y\right] = E_{xy}\left[x,y\right] - E\left[x\right]E\left[y\right]$$

D'après la définition de la covariance :

$$cov[x,y] = E_{xy}[(x - E_x[x])(y - E_y[y])]$$

$$cov[x,y] = E_{xy}[xy - xE_y[y] - yE_x[x] + E_x[x]E_y[y]]$$

La linéarité de l'espérance donne ensuite :

$$cov[x,y] = E_{xy}[xy] - E_{xy}[xE_y[y]] - E_{xy}[yE_x[x]] - E_{xy}[E_x[x]E_y[y]]$$

$$cov[x,y] = E_{xy}[xy] - E_{xy}[xE_y[y]] - E_{xy}[yE_x[x]] + E_x[x]E_y[y]$$

Or on peut remarquer que :

$$E_{xy}[x] = \textstyle\sum_x \sum_y xp[x,y]$$

$$E_{xy}[x] = \textstyle\sum_x x \sum_y p[y|x]p[x]$$

$$E_{xy}[x] = \textstyle\sum_x xp[x] \sum_y p[y|x]$$

$$E_{xy}[x] = \textstyle\sum_x xp[x]$$

$$E_{xy}[x] = E_x[x]$$

Donc :

$$cov[x,y] = E_{xy}[xy] - E_x[x]E_y[y] - E_y[y]E_x[x] + E_x[x]E_y[y]$$

D'où :

$$cov\left[xy\right] = E_{xy}\left[xy\right] - E_x\left[x\right]E_y\left[y\right]$$

4 Question 4

[6 points] : Programmez les algorithmes de regression linéaire et non lineaire polynomiales vues aux chapitres 1 et 3 ainsi que la recherche d'hyperparametre « k-fold cross-validation » vu au chapitre 1.e Pour ce faire, vous devez telecharger le fichier devoir1.zip du site web du cours.

Pour cette question, se référer au code envoyé dans le dossier Turninweb, aux commentaires présents et au README.md associé.