

Corrigé type de l'interrogation 1 en apprentissage automatique

Exercice 01 :

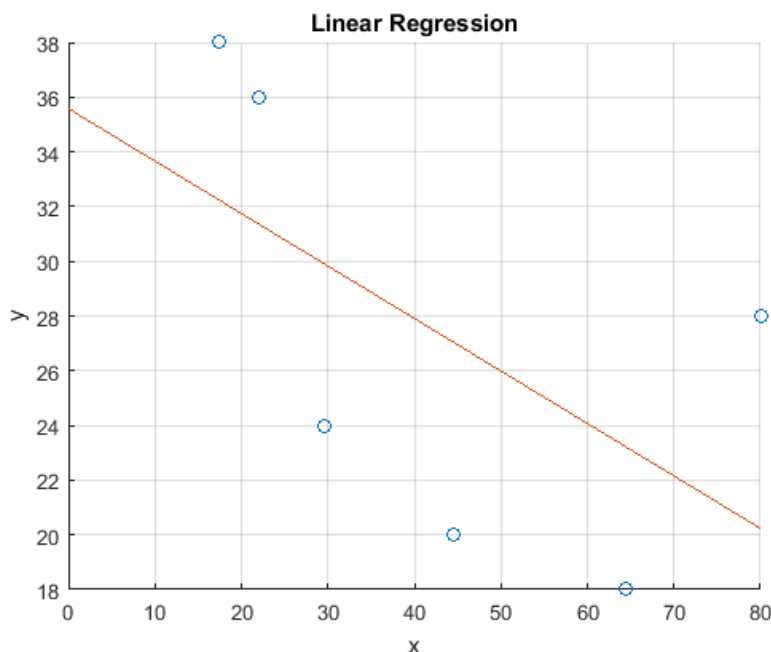
Soit l'ensemble de données représentant le nombre de décès de chauffeurs par 100 000 groupés par âges :

Age	Nombre de décès de chauffeurs par 100 000
17.5	38
22	36
29.5	24
44.5	20
64.5	18
80	28

- 1- Représenter les données dans un plan 2D.
- 2- Donner la formule mathématique qui permet de calculer les paramètres de régression linéaire pour ces données.
- 3- Prédire le nombre de décès pour les 40 et 60 ans.

Solution :

- 1- Représentations des données dans un plan 2D



- 2- Application du modèle de régression linéaire :

Dans le modèle de la régression linéaire, l'ensemble des paramètres est calculé par la formule suivante :

Application numérique sur les données de l'exercice :

On a _____ , donc

D'où

3- Prédire le nombre de décès pour les 40 et 60 ans.

- Si l'âge = 40, alors
- Si l'âge = 60, alors

Programme Matlab (facultatif):

```
x = [17.5; 22; 29.5; 44.5; 64.5; 80] ;
y = [38; 36; 24; 20; 18; 28] ;
x = [x ones(size(x,1),1)];
w = inv(x'*x)*x'*y ;
scatter(x(:,1), y)
hold on
x1 = [0:0.5:80]';
x1 = [x1 ones(size(x1,1),1)];
y1 = w'*x1';
plot(x1(:,1), y1)
xlabel('x')
ylabel('y')
title('Linear Regression')
grid on
```

Exercice 02 :

Etant donné l'ensemble d'apprentissage dans le tableau ci-dessous, prédire la classe de la nouvelle donnée suivante en utilisant la classification de Bayes Naïve :

âge<30, le revenu=moyen, étudiant=oui, évaluation crédit=acceptable

ID	Age	Revenu	Étudiant	Évaluation du crédit	Classe :
	<=30	G			acheter_ordinateur