

TD4 : Introduction à l'apprentissage machine (2h)

Exercice 1 –Regression linéaire

A l'aide d'un système à base d'arduino et capteur ultrasons, nous cherchons à mesurer l'évolution de la hauteur d'une tige d'une plante en fonction de la largeur de son tronc. Au laboratoire nous avons utilisé deux capteurs ultrason pour mesurer la hauteur de la tige en fonction du tronc pendant 03 mois pour prédire la hauteur de la tige en fonction de la largeur du tronc une fois la plante boisée dans le jardin. Pendant 03, mois nous avons collecté les valeurs suivantes :

Date de mesure	X_i (cm)	Y_i (cm)/10
Mars 2021	1	3
Juin 2021	4	4,5
Aout 2021	6	7
Septembre	7	???
Novembre	???	6

Q 1.1 Représenter graphiquement sous forme de nuage de points l'évolution de la tige de la plante en fonction de la largeur de son tronc.

Q 1.2 Trouver l'expression de l'hypothèse pour décrire l'évolution de la tige en fonction de la largeur de tronc?

Q 1.3 Ecrire le critère J pour décrire l'évolution de la tige de la plante en fonction de la largeur du tronc.

Q 1.4 Ecrire l'algorithme de descente de gradient pour calculer les paramètres (θ_1, θ_0) pour l'hypothèse $h_\theta(x)$.

Q 1.5 Calculer les valeurs de (θ_1, θ_0) en itérant l'algorithme de descente de gradient 7 fois avec $\alpha = 1/40$.

Q 1.6 Quelle est l'erreur commise entre les valeurs prédites et mesurées sur le site de plantation.

Exercice 2 –Regression logistique

Nous utilisons d'une LDR attache à un AOP en mode comparateur pour décider du seuil d'éclairage d'une salle. Les deux sortie de la LDR et l'AOP sont attachés à l'entrée analogique et numérique pour la lecture des valeurs mesures. Nous décidons de mettre en place un système intelligent pour prédire

l'éclairage. Nous collectons de la LDR et l'AOP les données pour écrire la fonction une hypothèse « éclairage ». Nous avons mis les valeurs dans la table ci-dessous :

Tension sortie de LDR (V/1000)	Sortie AOP (Etat 1=5V) (Etat 0=0V)
3	1
2,3	1
1,5	0
1	?
0,5	0
2,5	?
2,4	?
4	?

Notes : Les tensions mesurées sont augmentés de 500mv par le circuit connecté au système.

Nous choisissons de modeliser l'évolution de l'éclairage de la chambre sous forme d'une regression logistique et aider notre système à décider quand'il faut augmenter ou non l'éclairage de la salle.

Q 2.1 Représenter sur un graphe les points collectés pour la fonction hypothèse « éclairage » (tension de l'AOP en fonction de la tension en sortie du circuit LDR) ?.

Q 2.2 Trouver l'expression de l'hypothèse pour predire la fonction eclaireage.

Q 2.3 Comment la régression logistique peut aider à décider du seuil d'éclairage et prédire une augmentation d'éclairage de la salle.

Q 2.4 Donner l'expression du critère J en fonction de l'hypothèse de regression logistique.

Q 2.5 Utiliser l'algorithme de descente de gradient pour calculer les paramètres de la regression logistique. Iterer l'algorithme 5 fois pour avoir un θ_1, θ_0 .($\alpha = 1/5$).

Q 2.6 Tracer la fonction d'éclairage et déduire le seuil d'éclairage du graphe.

$$h_{\theta}(x) = \frac{1}{1 + e^{-\theta_1(x-500)-\theta_0}}$$

Exercice 3 –SVM

Supposons que nous ayons un ensemble de données avec deux caractéristiques (x_1 et x_2) et deux classes (-1 et 1). Nous avons 5 exemples d'entraînement :

X_1	X_2	Y
1	2	-1
2	3	-1
3	3	1
2	1	1
1	3	-1

Nous voulons entraîner un SVM pour classer ces exemples d'entraînement.

Q3.1 Calculer la fonction de perte de SVM

Q3.2 Calculer les première des dérivées de la fonction perte

Exercice 4 –SVM

Nous décidons de concevoir un système de ventilation qui se déclenche à chaque fois que la température et le taux d'humidité varie dans une salle confinée. Nous cherchons à mettre en place une aide à la décision pour activation système de la ventilation. Nous avons mesuré la température et le taux humidité. Nous avons mis des étiquettes +1 pour la fonction activer et -1 pour la fonction désactiver. Nous avons classé les les mesures dans 03 tables avec les étiquettes.

Température° ($(X)*+30$)	Humidité ($(Y)*10\%+10$ %	Activation du ventilateur
0	1	+1
2	1	+1
-1	0	+1
-2	2	+1

Température° ($(X)*+30$)	Humidité ($(Y)*10\%+10$ %	Activation du ventilateur
3	3	+1
1	0	+1

1	-2	-1
-2	-2	-1

Température° (X)*+ 30)	Humidité (Y)*10%+10 %	Activation du ventilateur
0	-1	-1
3	-2	-1
-4	-3	-1
0	-4	-1

Note: Les données sont simplifiées pour faciliter le calcul de la SVM

Q4.1 Calculer l'hyperplan optimal au sens de SVM en utilisant les points suivants de données suivant :

Points positives ($y=1$) $\{(2,2), (3,3), (3,4), (4,3)\}$

Points négatives ($y=-1$) $\{(1,1), (0,-1), (-1,1), (-1,-1)\}$

Q4.2 Calculer la marge maximale nécessaire.

Q4.3 Tracer l'hyperplan .

Exercice 5 –SVM

Nous désirons entraîner un SVM sur un ensemble de données avec deux classes. Nous avons les données suivantes :

- 3 exemples d'entraînement dans la première classe ($y=-1$) : (1, 2), (2, 3), (3, 3)
- 3 exemples d'entraînement dans la deuxième classe ($y=+1$) : (4, 4), (4, 5), (5, 5)

Nous utilisons la fonction de perte de charnière suivante:

$$J(\theta_2, \theta_1, \theta_0) = \frac{1}{m} \sum_{i=1, \dots, m} \max(0, 1 - y_i * (\theta_2 x_2 + \theta_1 x_1 + \theta_0))$$

- 1) Calculer la perte totale à pour le vecteur des poids initiale $[\theta_2, \theta_1, \theta_0] = [0, 0, 0]$. m est la taille du vecteur d'entraînement.
-