

Première partie : Sans document (durée : 45 minutes)

Rédigez vos réponses avec soin et de façon parfaitement lisible.

- 1) Expliquer le principe de *Early stopping* (arrêt précoce) d'un algorithme d'apprentissage.
- 2) Expliquer le principe de « Validation croisée » quand on n'a pas beaucoup de données.
- 3) Quelle est la différence entre les méthodes de gradient « déterministe » et « stochastique » ? Quels sont les avantages et les inconvénients de chacune ? Quel compromis entre ces 2 méthodes est souvent utilisé en pratique ?
- 4) Expliquer la méthode KNN (*K-Nearest Neighbors* ou « K plus proches voisins »). S'agit-il d'une méthode d'apprentissage supervisée ou non-supervisée ? Pourquoi ? Que se passe-t-il lorsque la valeur choisie pour K est trop petite ou trop grande ?
- 5) Quelles sont les principales motivations pour l'utilisation de la convolution dans un réseau de neurones ? Détaillez votre réponse.
- 6) Quels sont les avantages de *pooling* ? Détaillez votre réponse.
- 7) Expliquer les différences entre les options « *valid* », « *same* » et « *full* » pour le calcul d'une convolution. Quels sont les avantages et les inconvénients de chacune ?
- 8) Comment peut-on utiliser un classifieur SVM (*Support Vector Machine*) lorsque les données ne sont pas linéairement séparables ?
- 9) Que représente une courbe ROC (*Receiver Operating Characteristic*) ? Quelle mesure de l'évaluation de performances associée à cette courbe est couramment utilisée en pratique ?

Seconde partie : Documents de cours-TD-TP et calculatrices non-programmables autorisés (durée : 45 minutes)

Exercice 1 :

On souhaite estimer les valeurs des variables x et y qui minimisent la fonction suivante :

$$f(x,y) = (x-2)^2 + (2y-1)^2 + 4$$

(a) On utilise un algorithme du gradient à pas fixe (avec un pas $\varepsilon = 0.2$). Partant d'un point initial $x=0$ et $y=0$, calculer les valeurs obtenues pour x et y dans les 3 premières itérations de l'algorithme du gradient.

(b) On utilise l'algorithme de Newton. Partant du même point initial, calculer les valeurs obtenues pour x et y après la 1^{ère} itération. Conclusion ? Quels sont les inconvénients de l'algorithme de Newton en pratique ?

Exercice 2 :

Dans le programme Python suivant, qu'est-ce qui sera affiché après l'exécution de chaque instruction « `print` » ?

```
import numpy as np

tab=[[1, -2, 3, -4, 5],
      [-9, 8, -7, 6, -5],
      [5, 4, 3, -1, 0]]

A=np.array(tab)

print(A[1,2])

print(A[A<1])

print(A.cumsum(axis=1))

B=A

B[2,4]=4

print(B)

print(B[B[:,1]>B[:,2]])

print(A[-1:,-3:])

A=A.reshape(-1,3,5)

print(A.shape)
```

Exercice3 :

Pour un problème de classification d'images, le modèle suivant a été utilisé :

```

model = models.Sequential()

model.add(layers.Convolution2D(40, (3, 3), padding='same', input_shape=(48, 48, 1), activation='relu'))

model.add(layers.MaxPooling2D(pool_size=(2, 2), strides=(2, 2)))

model.add(layers.MaxPooling2D(pool_size=(3, 3), strides=(1, 1)))

model.add(layers.Flatten())

model.add(layers.Dense(100, activation='relu'))

model.add(layers.Dense(1, activation='sigmoid'))

model.compile(optimizer='adam', loss='BinaryCrossentropy', metrics=['accuracy'])

model.summary()

```

(a) Les images utilisées sont en noir et blanc ou en couleur ? Quel est le nombre de classes ? Justifier vos réponses.

(b) La dernière ligne du programme a affiché le tableau suivant où nous avons remplacé certaines valeurs par des points d'interrogation. Compléter le tableau (sur votre copie et non pas sur le sujet) avec les valeurs appropriées en justifiant votre réponse.

Layer (type)	Output Shape	Number of Parameters
=====		
conv2d (Conv2D)	(None, 48, 48, 40)	?
max_pooling2d (MaxPooling2D)	(None, ?, ?, ?)	?
max_pooling2d_1 (MaxPooling2D)	(None, ?, ?, ?)	?
flatten (Flatten)	(None, ?)	?
dense (Dense)	(None, ?)	?
dense_1 (Dense)	(None, ?)	?
=====		

Total number of parameters: ?

(c) Voici un rapport de classification sur la base de test généré à la fin de l'apprentissage du modèle :

Precision=0.9, Recall=0.8, False_Positive_Rate (FPR)=0.1

- Calculer *F-score* et *Accuracy*.
- S'il y a 1000 images dans la base de test, fournir la matrice de confusion (avec des valeurs arrondies).