Une image contenant texte

Description générée automatiquementUne image contenant texte, Police, logo, Graphique

Description générée automatiquement

**Nom : NIANG**

# Master 1 SD Prénom : Souna

**Analyse et visualisation de données**

Séance de TP 3

Modèle de mélange Gaussien

1. **Génération de données**
   1. Rappeler les équations du modèle de génération de données
   2. Écrire le code python de la méthode my\_GMM\_generate( )
   3. Expliquer et justifier les étapes de calcul de la méthode logsumexp( )
   4. Que réalise la méthode LogSumExp()
   5. Visualiser 1000 exemples générés du mélange étudié et analyser le résultat obtenu. Les données générées se conforment-elles au modèle ?
2. **Apprentissage du modèle de mélange**
   1. *Rappeler les étapes de l’algorithme EM en les expliquant*
   2. *Quelles sont les formules de réestimation du modèle de mélange ?*
   3. *Compléter la méthode* ***my\_GMM\_p\_a\_posteriori()***
   4. *Que réalise la méthode* ***my\_GMM\_predict()***
   5. *Compléter la méthode* ***my\_GMM\_fit( )***
3. **Analyse des résultats**
4. Illustrer le comportement de l’algorithme GMM\_fit() sur les données du mélange généré. Visualiser les clusters du modèle simulé et les comparer aux clusters déterminés par le modèle GMM trouvé. La distribution des données vous semble-t-elle respectée ? Y-a-t-il des points mal classsés ? Examiner les paramètres du modèle trouvé et les comparer à ceux du modèle génératuer. Les modèles sont-ils semblables ? Discuter.
5. **Critères d’information**

Rappeler l’expression des critères d’information AIC et BIC.

Programmer ces deux critères d’information et déterminer le meilleur modèle de mélange déterminé par chacun d’eux en visualisant leurs valeurs en faisant varier K de 2 à 7. Ces deux critère sont-ils concordant ?