

## **Bilan projet scala:**

### **Début du projet:**

J'ai tout d'abord réfléchi au nombre de classe que j'allais utiliser ensuite donc j'ai commencé par faire un diagramme de classe et j'ai choisi de faire 4 classes qui sont:

Kmeans, Donnees, MatriceDonnees et Cluster

### **Difficulté rencontré:**

J'ai pris du temps sur la conception du diagramme de classe qu'il fallait d'abord effectuer pour ensuite commencer à coder. J'ai commencé par faire la lecture du fichier iris.data qui au final nous avons pu prendre sur internet.

Je n'ai pas pu aller jusqu'au bout de ce projet car je n'est pas réussi à faire l'affiche avec les graphiques je n'ai pas compris comment utiliser la bibliothèque Ploty.

La condition de sortie du Kmeans, j'ai mis 300 itérations pour être sûr que l'algorithme a le temps de tourner et de tout bien faire.

### **Les points positifs du projet:**

Cela m'a permis d'apprendre un nouveau langage. mais cette fois c'est différents car nous avons eu un projet concret. C'est beaucoup plus facile d'apprendre un langage en faisant un projet plutôt que de lire ou d'assister à des cours sans réel

### **Partie technique:**

#### **Class MatriceDonnee:**

La classe est la matrice donc toutes les lignes. Il y a tout d'abord mon toString pour l'affichage, ensuite deux façons de récupérer la matrice avec les deux guetteurs et un guetteur pour avoir la taille de la matrice

#### **Class Donnees:**

Tout d'abord les données sont les lignes de la matrice donc les coordonnées plus le type. Il y a la fonction distance pour calculer la distance entre deux données. Deux getteurs pour récupérer les données que j'ai appelé des points, un guetteur pour la longueur d'une ligne et un guetteur pour récupérer le type.

#### **Class Cluster:**

Un cluster est un groupe de données avec un centroïde donc qui est le centre. Il y a deux guetteurs pour récupérer l'id d'une ligne, un guetteur pour récupérer un centroïde, un setteur pour en créer un, une fonction pour ajouter un indice et une fonction pour remettre à zéro une ligne.

#### **Class Kmeans:**

Kmeans est l'algorithme. Tout d'abord il y a l'initialisation de la matriceDonnees donc pour récupérer à partir du fichier iris.data. L'algorithme Kmeans qui possède plusieurs fonctions: tableauDeClusters qui met dans un Tableau les k clusters, la condition de sortie, updateCentroide qui permet de mettre à jour les centroïdes en recalculant la distance entre

les clusters et chaque données, ensuite le résultat qui permet de mètres le résultat de quelles données va avec quel cluster dans un tableau.

Une fonction affichage pour afficher quelles données va avec quel cluster.

Une fonction min pour avoir le plus petit indice.

Les fonctions moyennes, variance, écart-Type, covariance et une fonction commentaire des résultats

### **Pour exécuter le programme:**

Il faut tout d'abord installer scala via le site web officiel, il faut suivre les étapes.

Ensuite il faut extraire le dossier KMeans.

Puis aller dans un terminal, se déplacer via le terminal dans le dossier Kmeans.

Ecrire dans le terminal "scala Test.scala"