

# TP1: Implémentation de l'algorithme Fuzzy c-means

Laila Sabar  
Lailasb2@gmail.com  
Université de Paris  
VMI2

## Introduction

Dans ce travail, on va implémenter et utiliser l'algorithme de Fuzzy-Cmeans et l'appliquer sur deux différents types de données, sur des images et sur le dataset Iris.

L'implémentation est dans le fichier ipynb que vous trouverez ci-joint. Dans la première partie, on va décrire l'algorithme Fuzzy-Cmeans, puis la deuxième partie contient les résultats obtenus de l'algorithme. Enfin, on conclue ce travail par une discussion de ces résultats.

## Algorithme Fuzzy c-means

L'algorithme\*\* Fuzzy c-means (FCM) est une méthode de clustering qui permet à une donnée d'appartenir à deux ou plusieurs clusters. Cette méthode (développée par Dunn en 1973 et améliorée par Bezdek en 1981) est fréquemment utilisée en reconnaissance de formes. Voici la documentation officielle de l'algorithme :

[https://home.deib.polimi.it/matteucc/Clustering/tutorial\\_html/cmeans.html](https://home.deib.polimi.it/matteucc/Clustering/tutorial_html/cmeans.html)

Cette méthode repose sur la minimisation de la fonction objectif suivante :

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2, \quad 1 \leq m < \infty$$

- $u_{ij}$  est le degré d'appartenance d'une observation  $x_i$  à un cluster  $c_j$
- $c_j$  est le centre de dimension de cluster  $j$
- $m$  est le facteur de flou qui doit être supérieur à 1

On peut voir que le FCM diffère du k-means en utilisant les valeurs d'appartenance  $u_{ij}$  et le facteur de flou  $m$ .

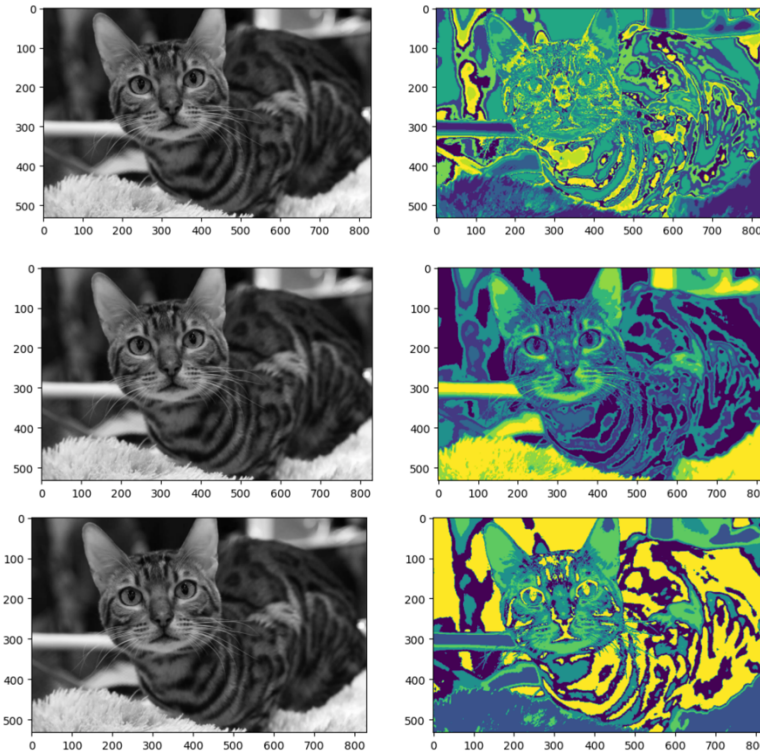
La variables  $u_{ij}$  et  $c_j$  sont définies comme suit :

- Le degré d'appartenance,  $u_{ij}$ , est lié inversement à la distance de  $x$  au centre du cluster.

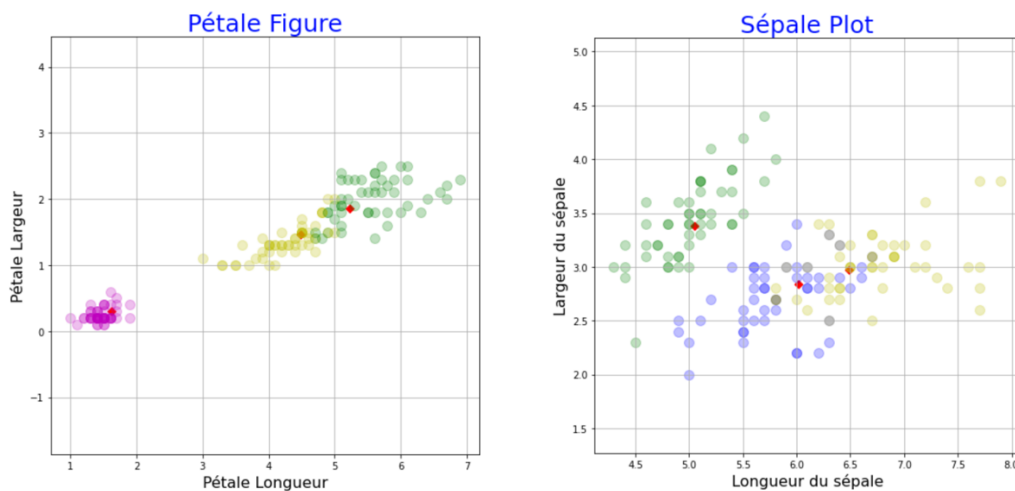
$$u_{ij} = \frac{1}{\sum_{k=1}^C \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}, \quad c_j = \frac{\sum_{i=1}^N u_{ij}^m \cdot x_i}{\sum_{i=1}^N u_{ij}^m}$$

## Résultats

Segmentation d'image (clusters 9, 7, 5)



## Clustering Iris dataset



## Conclusion

Durant ce travail on a implémenté les différentes étapes de l'algorithme fuzzy c-means, et on a appliqué l'algorithme sur deux types de données, une image et le dataset iris. L'algorithme a prouvé sa performance pour la segmentation d'image on peut voir sur les images avec les différents nombres de clusters que l'algorithme a réussi à mettre en valeur et faire apparaître les caractéristiques du chat, ce qui facilite la tâche pour la classification de l'image. L'application de l'algorithme fuzzy c-means sur la dataset Iris nous a permis de faire la segmentation des individus de la datasets en 3 différents classes avec la possibilité d'appartenir à plusieurs classes au même temps, on voit bien sur les graphe la distrubution ainsi que la classification de tous les individus et aussi les centres de chaque classe, on peut dire que l'algorithme est efficace avec ce types de données. Comme pour les réseaux de neurones, l'optimisation des hyperparamètres est aussi importante pour ce type de problème. Il pourrait être intéressant d'utiliser ce réseau sur un ensemble massif de données.