

**Proposition d’un sujet de thèse en cotutelle**

**Université Libanaise**

**Nom : Prénom :**

ABDALLAH

Fahed

**Titre (Prof, HDR, ….) :**

Prof-HDR

**Laboratoire : Adresse Web :**

https://www.ul.edu.lb/faculte/branches.aspx?facultyId=23

EDST

**Etablissement : Adresse Web :**

https://www.ul.edu.lb/faculte/branches.aspx?facultyId=19

UL-Fac Technologie

**Domaines d’expertise :**

Data Science, Machine learning

**Publications importantes en relation avec le sujet proposé :**

El Hajjar, S., Dornaika, F., Abdallah, F-O., & Barrena, N. (2022). Consensus graph and spectral representation for one-step multi-view kernel based clustering. Knowledge-Based Systems, 241, [108250]. <https://doi.org/10.1016/j.knosys.2022.108250>

S. El Hajjar, F. Dornaika, F. Abdallah (2022). One-step multi-view spectral clustering with cluster label correlation graph, Information Sciences, Volume 592, Pages 97-111, ISSN 0020-0255, https://doi.org/10.1016/j.ins.2022.01.017.

**Adresse Web de votre page personnelle :**

https://scholar.google.fr/citations?user=7DMaJnQAAAAJ&hl=fr

Fahed.abdallah76@gmail.com

**Adresse mail :**

**Partenaire à l’étranger :**

**Nom : Prénom :**

Jrad

Nisrine

**Titre (Prof, HDR, …) :**

Maître de Conférences

**Laboratoire : Adresse Web :**

Laboratoire Angevin de Recherche en Ingénierie des Systèmes (LARIS)

<http://laris.univ-angers.fr/fr/index.html>

**Etablissement : Adresse Web :**

Université Catholique de l’Ouest (UCO)

<https://www.uco.fr/fr>

**Domaines d’expertise :**

Ma thématique de recherche porte sur les méthodes d’apprentissage automatique (Machine Learning) et le traitement du signal, en particulier dans le domaine médical.

**Publications importantes en relation avec le sujet proposé :**

**Adresse Web de votre page personnelle :**

<https://recherche.uco.fr/chercheur/nisrine-jrad>

njrad@uco.fr

**Adresse mail :**

**Description du sujet de thèse proposé : Discipline :**

Data Science and machine learning

**Titre et Résumé :**

DC-GNN : Deep clustering with Graph Neural Networks for real world data

**Sujet :**

**Description du sujet (contexte scientifique, description du problème, Objectifs, …..) :**

Clustering is an unsupervised machine learning technique that involves grouping data samples. In the literature, most proposed approaches use shallow models. Recently, some works have used deep architectures such as deep K-means, graph convolutional networks based clustering and online deep clustering [1,2]. In the image domain, several solutions aim to learn deep convolutional neural networks whose loss is given by the reconstruction error and a clustering-based loss (as in Kmeans or spectral clustering). While these methods provided satisfactory results, most proposed deep solutions do not offer clustering as a direct solution. Rather, a data representation is first computed and then a clustering algorithm is applied to this representation. For shallow graph-based models or kernel-based models, it has been shown that direct solutions, also known as one-step clustering solutions, can be superior to the indirect solution. This is because data partitioning is considered in the criterion being optimized.

**Approche méthodologique :**

Learning with graph data, such as social networks and citation networks, has recently gained increasing attention. Graph Neural Network (GNN) has become the most important tool for learning representations on graphs. GNN is an extension of traditional neural networks for processing graph data. Graph Convolutional Networks (GCNs) [3,4] and their variants have been proposed for semi-supervised classification of nodes in graph structured data. The main application of GCNs is semi-supervised classification. The use of GCNs in unsupervised clusters will be explored.

**Résultats attendus :**

The main goal of this research is to propose deep clustering algorithms based on GNN. To this end, the research work consists of: Developing a self-supervised deep graph embedding feature that handles the interaction of high order nodes for clustering, learning such GNN architectures that integrate the attributes and the convolution and aggregation operator for structured information. The output can be the low dimensional embedding vectors which can be seen as the nonlinear data projection. Using of an architecture with two heads. One head provides the nonlinear spectral representation and the other provides the cluster index matrix. Applying proposed algorithms on image clustering.

**Bibliographie :**

[1]- X. Zhanga, H.  Liua, X.  Wuc, X.  Zhang, X. Liu   Spectral Embedding Network for Attributed Graph Clustering.  Neural Networks,  2021.

[2]- X. Zhan, J. Xie, Z.Liu, Y. Ong, C. Loy.  Online deep clustering for unsupervised learning. IEEE CVPR 2020.

[3]- T. N. Kipf and M. Welling. Semi-supervised classification with graph convolutional networks.
International Conference on Learning Representations, 2017.

[4]- M. T. Kejani, F. Dornaika and H. Talebi.   Graph Convolution Networks with Manifold Regularization for Semi-Supervised Learning. Neural Networks,  volume 127, pp. 160-167,  2020.

**Mots clés :**

Deep clustering, graph data, Graph Neural Networks (GNN)

**Possibilité de financement (Justificatif éventuel) :**

Contrat doctoral pour la durée de séjour en France

**Profil Scientifique du candidat :**

Le candidat recherché doit être titulaire d’un Bac+5 (master, diplôme d’ingénieur ou équivalant) et issu des formations en data science, en informatique, mathématiques appliquées ou en ingénierie Mécanique/électrique. Des compétences en programmation seront indispensables.

* À joindre un fichier PDF détaillant le sujet.