

GraphSAGE : Résumé, Implémentation, Expérimentations

Samar Krimi

samar.krimi@dauphine.tn

2024-2025



Table des matières

1	Partie 1 : Résumé et Implémentation Complète de GraphSAGE	3
1.1	Résumé	3
1.2	Méthodologie : Le Cœur de l’Algorithme GraphSAGE	3
1.2.1	Apprentissage Inductif par Echantillonnage et Agrégation des Voisins : . .	3
1.2.2	Gestion des Graphes de Grande Taille	4
1.3	Stratégies d’Agrégation	5
1.4	Entraînement du Modèle en Apprentissage Supervisé	5
1.5	Entraînement du Modèle en Apprentissage Non Supervisé	6
1.6	Comparaison entre GraphSAGE, GCN et GAT	6
1.6.1	Objectif et Cadre d’Application	6
1.6.2	Mécanismes d’Agrégation	7
1.6.3	Scalabilité et Capacité Inductive	7
1.6.4	Performances et Complexité	7
1.7	Expérimentations et Comparaisons	7
1.8	Conclusion des Expérimentations	8

2	Partie 2 : Implémentation avec Citeseer, OpenAlex et PPI	9
2.1	Introduction	9
2.2	OpenAlex : Graphe de Connaissances Scientifiques en Open-Source	9
2.3	Étude des Datasets	10
2.3.1	Citation Networks (Citeseer)	10
2.3.2	Dataset PPI (Protein-Protein Interaction)	10
2.3.3	Dataset OpenAlex	10
2.4	Acquisition, Prétraitement et Construction du Graphe	10
2.5	Adaptation de l'Implémentation de GraphSAGE	11
2.6	Expérimentation et Évaluation	11
2.6.1	Comparaison Supervisé vs. Non Supervisé	12
2.6.2	Performance par jeu de données	13
2.6.3	Comparaison avec DeepWalk	13
2.6.4	Synthèse	13
	Annexe : Plan Général du Code	16