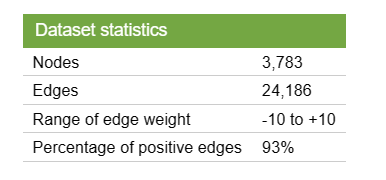
#### **Sous l’Encadrement de : réalisé par :Bougatef Roudayna**

Prof. **feriel ben fraj**

Analyse des réseaux sociaux

**Bitcoin Alpha trust weighted signed network:**

L’ de données représente un réseau de confiance entre les personnes qui effectuent des transactions en utilisant Bitcoin sur une plateforme appelée Bitcoin Alpha. Étant donné que les utilisateurs de Bitcoin sont anonymes, il est nécessaire de conserver un enregistrement de la réputation des utilisateurs pour prévenir les transactions avec des utilisateurs frauduleux et risqués. Les membres de Bitcoin Alpha notent les autres membres sur une échelle de -10 (défiance totale) à +10 (confiance totale) par incréments de 1. Il s'agit du premier réseau dirigé pondéré et signé explicitement disponible pour la recherche.



ensemble de données est la plateforme Bitcoin Alpha, sur laquelle les utilisateurs effectuent des transactions en utilisant Bitcoin. Les données sont recueillies à partir des évaluations que les membres de la plateforme attribuent les uns aux autres, représentant le niveau de confiance entre les utilisateurs.

les entités (nœuds) et les relations entre elles (liens) :Les entités, ou nœuds, dans ce réseau sont les membres de la plateforme Bitcoin Alpha qui effectuent des transactions. Les relations entre les nœuds sont basées sur les évaluations de confiance que les membres attribuent les uns aux autres. Ces évaluations créent des liens dirigés et pondérés entre les nœuds, indiquant le niveau de confiance entre eux.

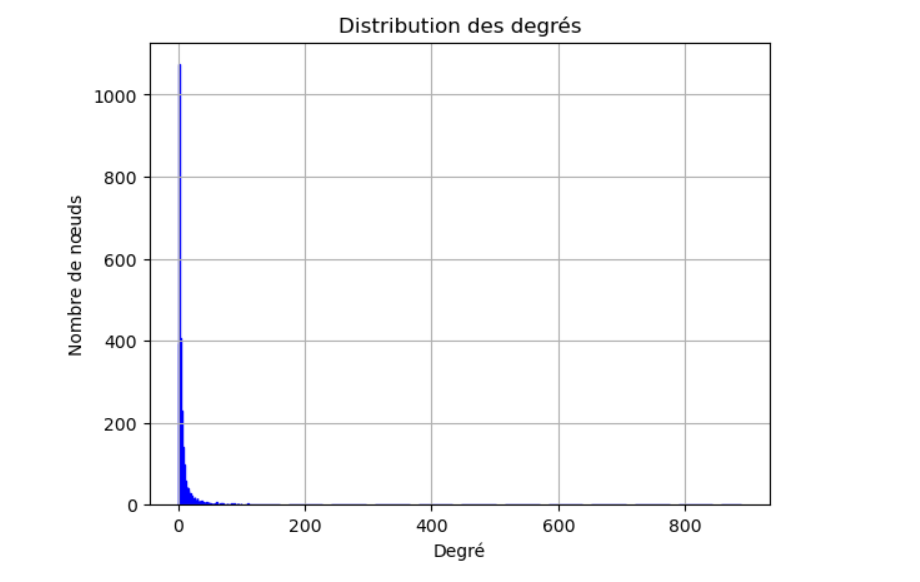
3. Identifier les informations additionnelles valables : Les informations additionnelles dans cet ensemble de données incluent les attributs des nœuds, tels que les identifiants des membres, ainsi que les poids des liens, qui représentent les évaluations de confiance attribuées par les membres. Les poids des liens sont évalués sur une échelle de -10 à +10, où -10 représente une totale défiance et +10 une totale confiance.

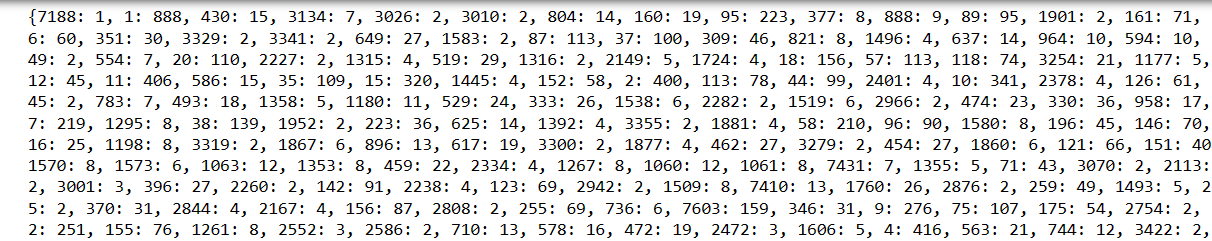
4. Obtenir les données à partir de la source de données sélectionnée : Les données sont obtenues à partir des évaluations des membres de Bitcoin Alpha, qui sont recueillies par la plateforme pour maintenir un enregistrement de la réputation des utilisateurs. Ces données sont ensuite rendues disponibles pour la recherche et l'analyse.

5. Construire un réseau : Les données recueillies sont utilisées pour construire un réseau de confiance entre les membres de Bitcoin Alpha. En utilisant les nœuds comme représentation des membres et les liens pondérés comme indication de la confiance entre eux, un réseau est créé pour analyser les relations et la dynamique de confiance au sein de la plateforme.

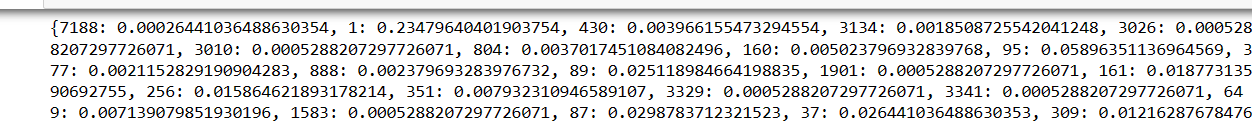
**Partie 2 : Analyse du réseau**

1. L’analyse de la distribution des degrés :

****

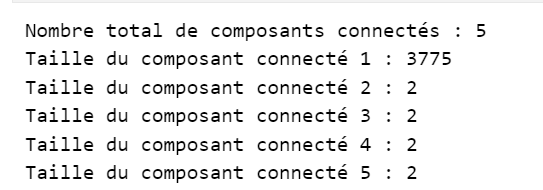
****

un dictionnaire contenant les degrés de chaque nœud dans le graphe.

****

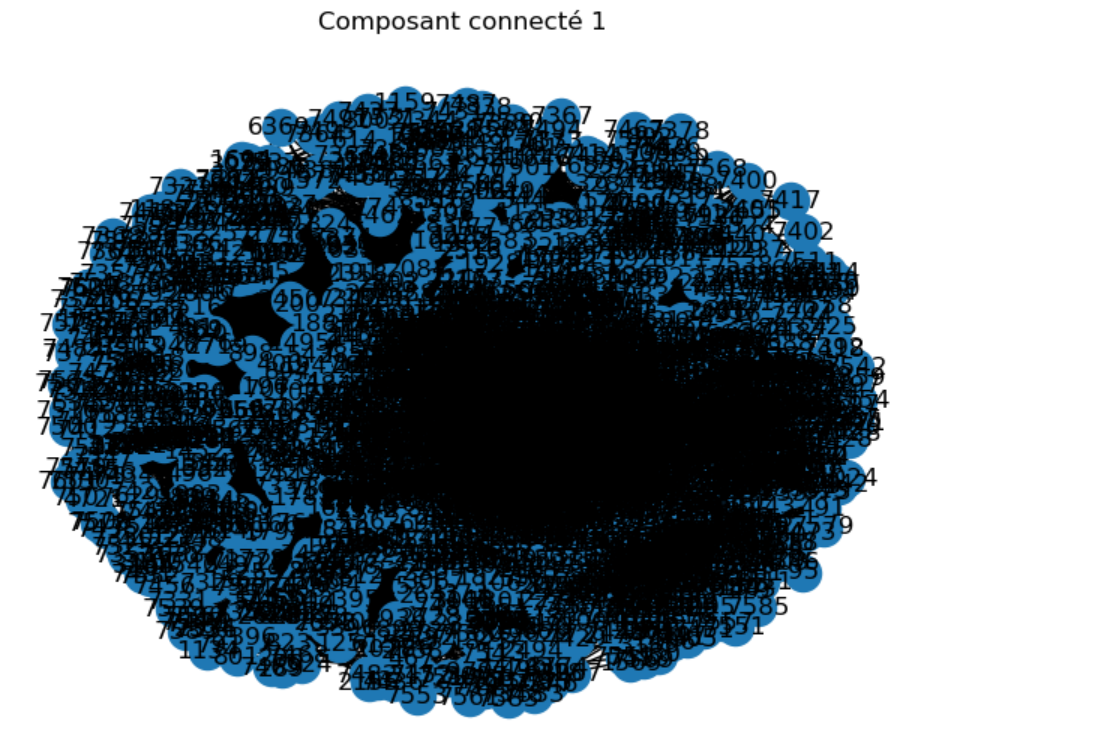
la centralité de degré pour chaque nœud dans le graphe.

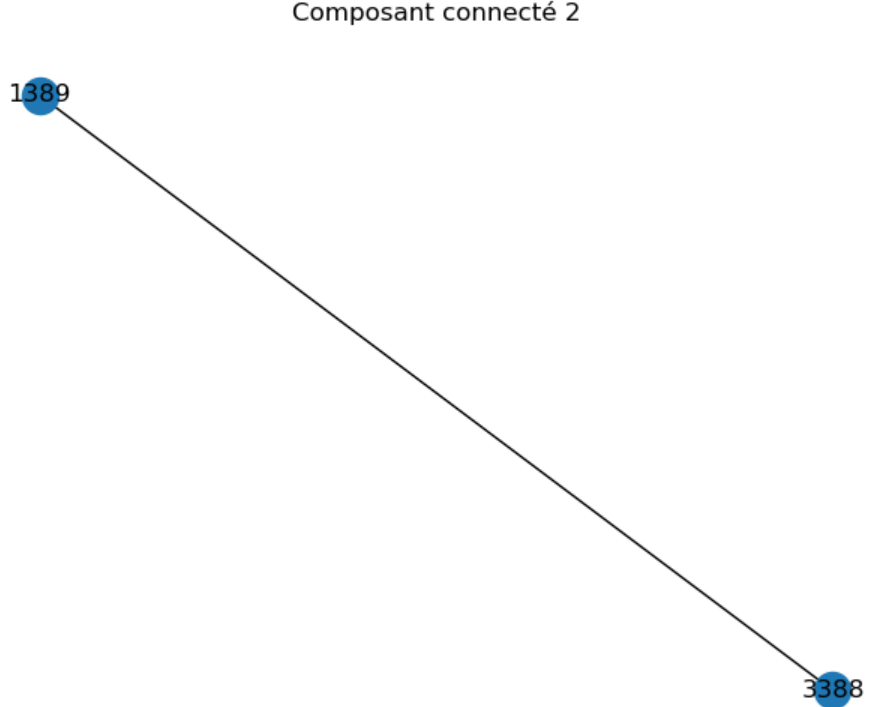
1. L’analyse des composants connectés :

****

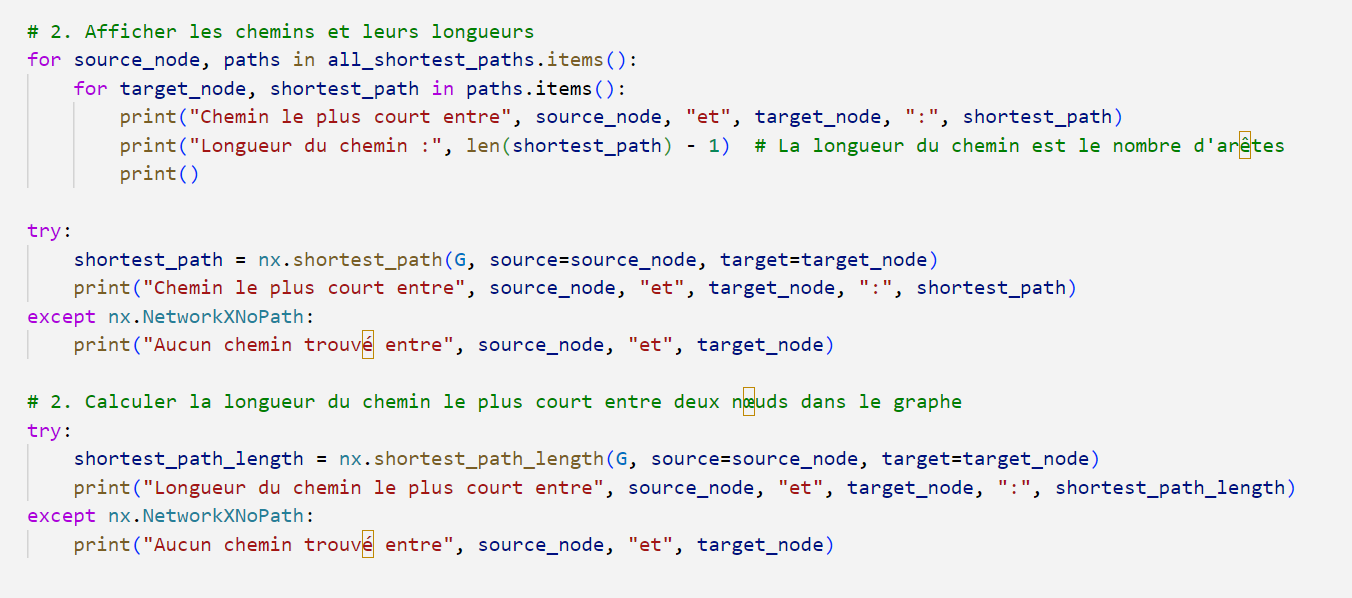
Ces résultats indiquent que le graphe étudié se compose de 5 composants connectés distincts. Le plus grand de ces composants comprend 3775 nœuds, ce qui représente une partie significative du graphe.

On peut visiualiser les résultats ci-dessous :

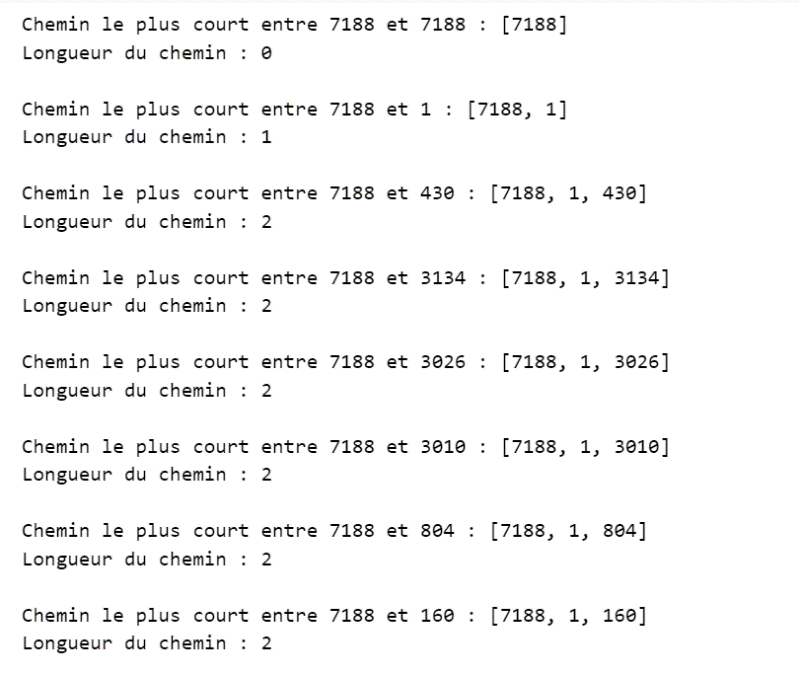




1. L’analyse des chemins :

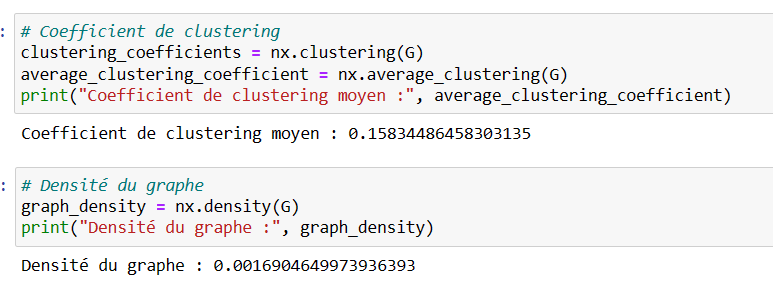
****

Ce code nous donne la résultats ci-dessous :

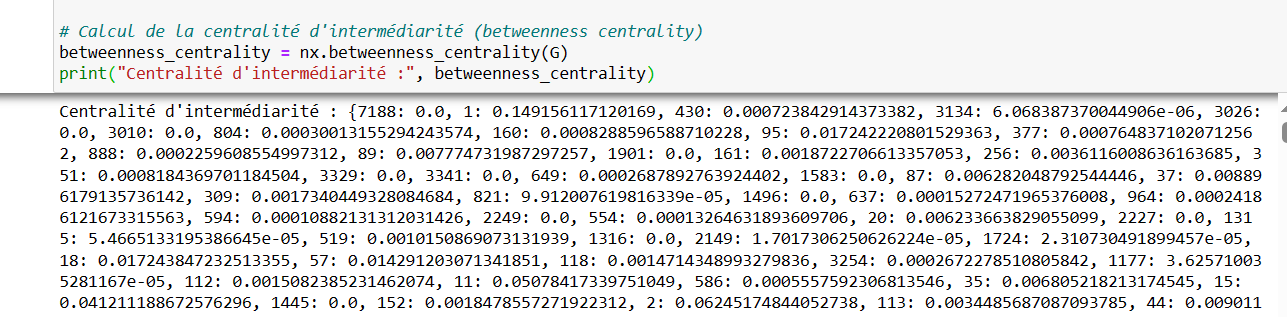


La résultat affiche le chemin le plus court entre deux nœuds et le longueur du ce chemin .

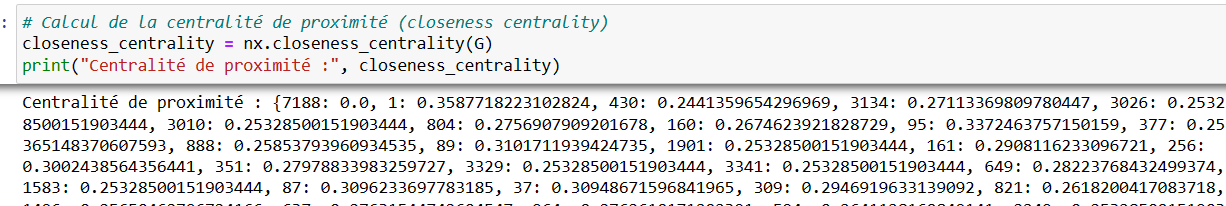
1. Le coefficient de clustering et l’analyse de la densité

****

1. L’analyse de la centralité.

****

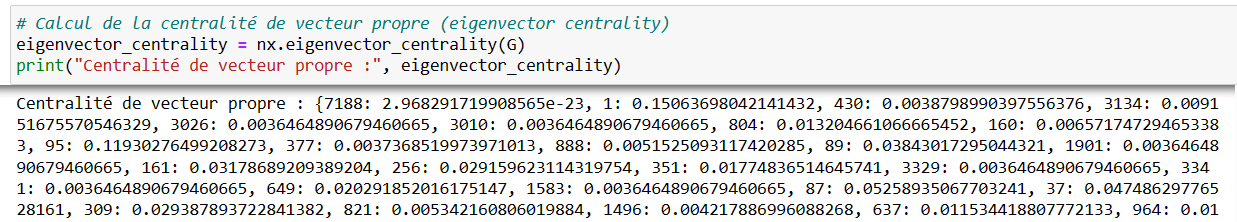
La centralité d'intermédiarité mesure l'importance des nœuds dans un réseau en fonction du nombre de chemins les traversant.

****

Centralité de proximité :

- La centralité de proximité mesure la proximité d'un nœud avec tous les autres nœuds du graphe. Plus la valeur de la centralité de proximité est élevée, plus le nœud est central dans le réseau.

- dans notre cas chaque nœud est associé à une valeur de centralité de proximité, allant de 0 à 1. Une valeur de 1 indiquerait que le nœud est le plus central possible, tandis qu'une valeur proche de 0 indiquerait que le nœud est moins central.

****

Centralité de vecteur propre :

- La centralité de vecteur propre mesure l'importance d'un nœud dans le réseau en tenant compte de l'importance des nœuds avec lesquels il est connecté.

- Dans notre cas chaque nœud est associé à une valeur de centralité de vecteur propre, allant de 0 à 1.

Une valeur de 1 indiquerait que le nœud est le plus important possible, tandis qu'une valeur proche de 0 indiquerait que le nœud est moins important.

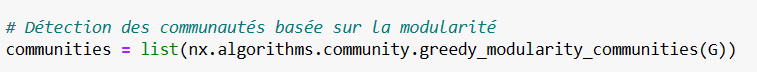
En comparant ces deux mesures de centralité, vous pouvez avoir une idée de la manière dont les nœuds sont positionnés dans le réseau et de leur influence potentielle sur la diffusion

**Partie 3 : Identification des communautés :**

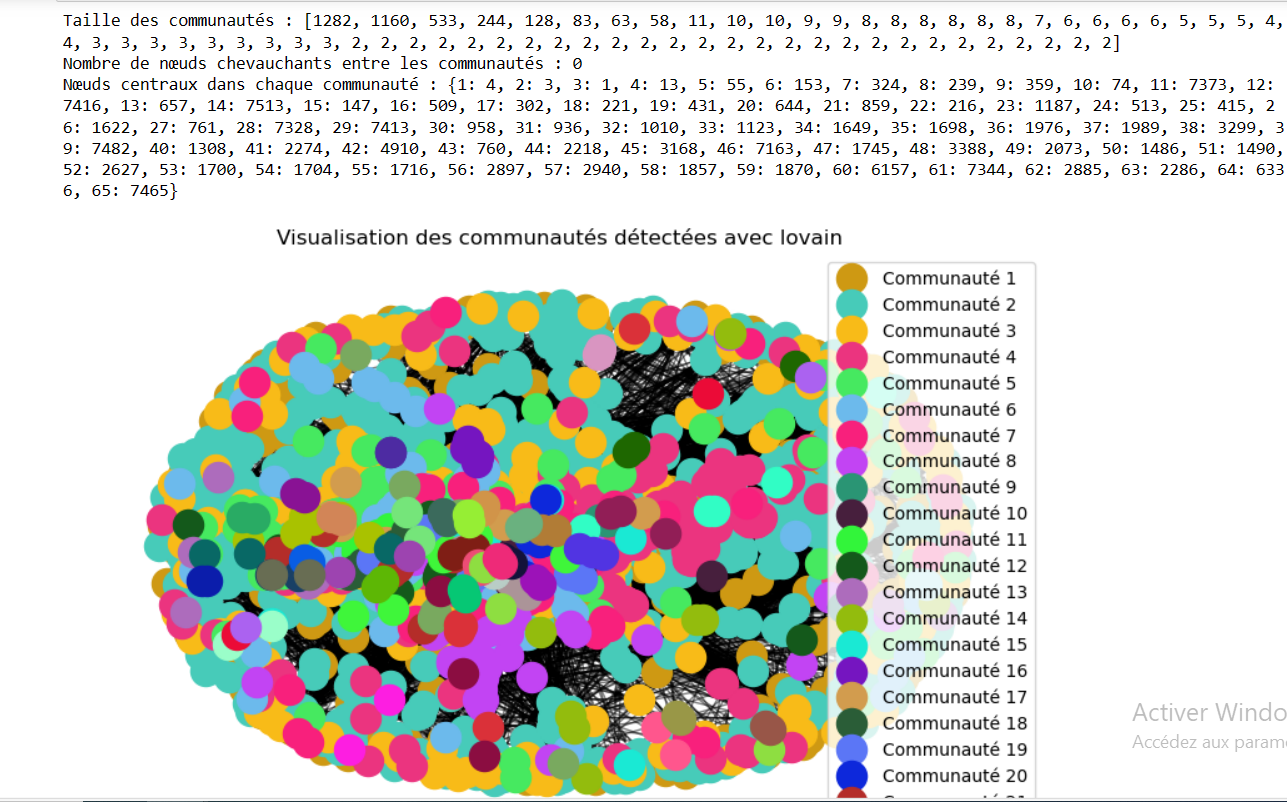
Cette étape affiche une détection des communautés de nœuds qui sont étroitement liés les uns aux autres dans le réseau avec des différents algorithme :

1. Algorithme lovaine :

La fonction de détection des communautés :



Les résultats :



Analyse des résultats :

1. Taille des communautés :

- Les communautés varient en taille, allant de 2 à 1282 nœuds. Cela suggère une certaine hétérogénéité dans la structure du réseau, avec quelques grandes communautés et de nombreuses petites communautés.

2. Nombre de nœuds chevauchants entre les communautés:

- Le nombre de nœuds chevauchants entre les communautés est de 0. Cela signifie qu'il n'y a pas de chevauchement de nœuds entre les différentes communautés détectées.

3. Nœuds centraux dans chaque communauté :

- Pour chaque communauté détectée, un nœud central est identifié. Ces nœuds centraux peuvent être importants pour comprendre la structure et les interactions au sein de chaque communauté. Par exemple, dans la communauté 11, le nœud central est identifié comme ayant l'ID 7373, et dans la communauté 12, le nœud central est identifié comme ayant l'ID 7416.

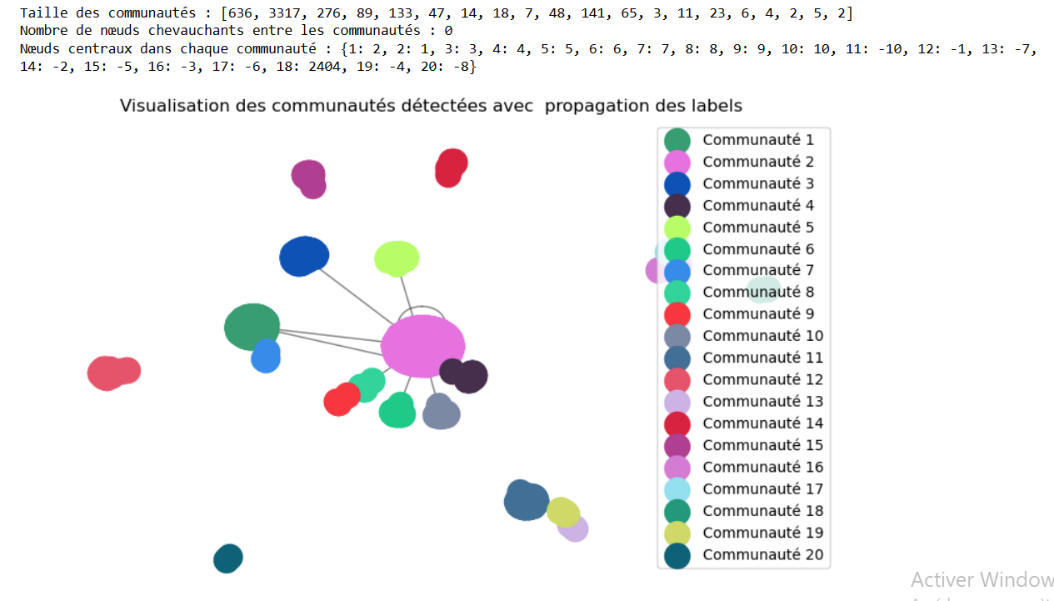
Ces résultats suggèrent une structure communautaire distincte dans le réseau social analysé, avec des communautés de tailles différentes et des nœuds centraux identifiés dans chaque communauté. Cette analyse fournit un aperçu précieux de la manière dont les individus sont regroupés et interagissent dans le réseau.

1. algorithme de propagation des labels :

la fonction du détection du communauté par cet algorithme :



Les résultats :



1. Taille des communautés :

- La taille des communautés varie de 2 à 3317 nœuds.

- La communauté la plus petite contient seulement 2 nœuds, tandis que la plus grande en contient 3317.

2. Nombre de nœuds chevauchants entre les communautés:

- Il n'y a aucun nœud chevauchant entre les communautés, ce qui signifie que chaque nœud n'appartient qu'à une seule communauté.

3. Nœuds centraux dans chaque communauté :

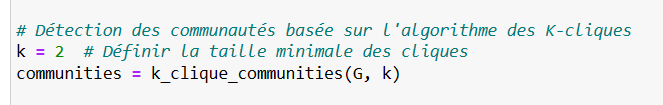
- Chaque communauté semble avoir un nœud central spécifié. Cela peut être important pour comprendre les structures de réseau et les dynamiques communautaires.

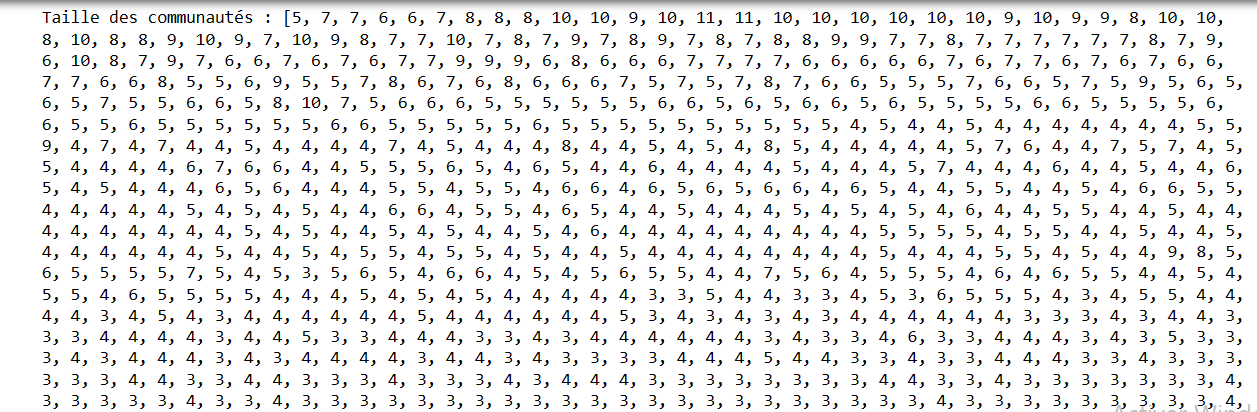
- Les nœuds centraux sont spécifiés par leur identifiant. Par exemple, dans la première communauté, le nœud central est le nœud 2.

Cette analyse fournit une vue d'ensemble des communautés détectées dans le réseau, ainsi que des informations sur leur taille, les chevauchements entre elles et les nœuds centraux de chaque communauté.

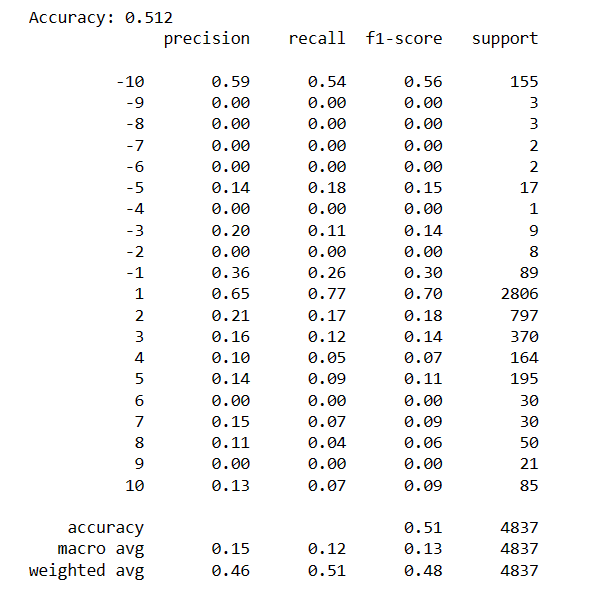
1. L’algorithme de k-clique :

L'algorithme de k-clique est utilisé pour détecter des groupes d'individus dans un réseau qui forment des sous-graphes complets de taille k.





1. Partie 4 : Prédiction des liens :



Ces résultats sont des statistiques de performance pour un modèle de classification basé sur des prédictions des liens. Voici une analyse de ces résultats :

1. Accuracy (précision globale) : L'accuracy est de 0.512, ce qui signifie que le modèle prédit correctement la classe des éléments dans l'ensemble de données dans environ 51,2 % des cas.

2. Précision, rappel et f1-score par classe :

- Pour chaque classe, la précision mesure la proportion d'éléments prédits comme appartenant à cette classe qui sont réellement de cette classe.

- Le rappel mesure la proportion d'éléments appartenant réellement à une classe qui sont correctement prédits par le modèle comme appartenant à cette classe.

- Le f1-score est la moyenne harmonique de la précision et du rappel et est une mesure de la performance du modèle pour une classe particulière.

3. Support : Le support est le nombre d'occurrences de chaque classe dans l'ensemble de données.

4. Analyse par classe :

- Les classes -10, -5, -3, -1, 1 et 2 ont des performances relativement meilleures, avec des précisions et des rappels plus élevés.

- Les classes -9, -8, -7, -6, -4, -2, 3, 4, 5, 6, 7, 8, 9 et 10 ont des performances inférieures, avec des précisions et des rappels plus bas.

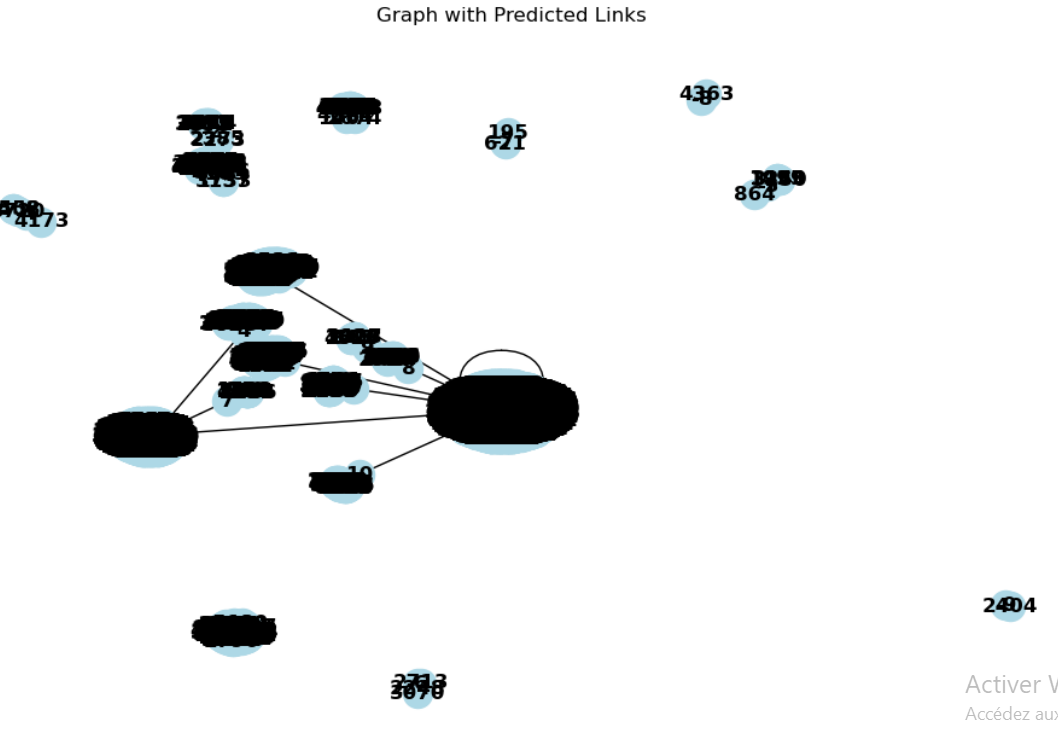
- Certaines classes ont des performances très faibles, avec des précisions et des rappels de 0 pour certaines classes.

5. Analyse globale :

- La macro moyenne des précisions, des rappels et des f1-scores est relativement faible, indiquant une performance globale médiocre pour le modèle.

- La moyenne pondérée des précisions, des rappels et des f1-scores est légèrement meilleure, mais montre toujours une performance moyenne.

En résumé, les résultats indiquent que le modèle a des performances variables selon les classes, avec une précision globale assez faible on a afficher les lien prédites dans la graphe ci-dessous :



**Conclusion :**

Dans notre projet d'analyse des réseaux sociaux, nous avons exploré la détection de communauté, la prédiction des liens et l'analyse des relations. Grâce à ces analyses, nous avons identifié des groupes d'utilisateurs, prédit de nouvelles connexions et examiné les dynamiques d'engagement et d'influence.