Influencer Detection in Social Media Using Degree Centrality

# Group Details

Group Leader: Waleed Haroon-(22k-4596)  
Group Members: Muzamil Soomro-(22k-4598)   
  
Project Title: Influencer Detection in Social Media Using Degree Centrality

# Project Overview

Goal: Identify social media influencers using degree centrality and other network features.  
  
Concepts: Centrality, network analysis.  
  
Tools: Python, NetworkX, Scikit-learn.

# Project Methodology

In this project, the goal was to identify the most influential users on social media platforms, specifically TikTok, by analyzing their engagement metrics and constructing a graph-based social network. The project utilized degree centrality to measure the connectedness of each user within the network, with the assumption that more connected users have a higher degree of influence.  
  
The dataset used for this project contained user metrics such as subscribers, views, likes, comments, and shares. This data was preprocessed and transformed into a graph where each user was represented as a node, and their interactions (followers, engagement metrics) were modeled as edges. Using Python and NetworkX, degree centrality for each user was calculated, and the most connected users were identified as the influencers in the network.  
  
The methodology involved the following steps:  
1. Data Preprocessing: Handling missing data and converting engagement metrics to numeric values.  
2. Graph Construction: Nodes represent users and edges represent interactions.  
3. Degree Centrality Calculation: Using NetworkX to calculate the degree centrality for each node.  
4. Visualization: Plotting the social network to highlight the most influential users.

# Results and Conclusion

**Results:**  
After processing the dataset and calculating degree centrality for each user, we identified the top 10 TikTok users with the highest centrality scores. These users exhibited significant influence within the social media network based on their number of connections and interactions with other users. The degree centrality scores ranged from 0 to 1, with higher values indicating greater influence. The top influencers were visualized in a graph, where nodes with higher degree centrality were emphasized.  
  
**Conclusion:**  
The use of degree centrality in identifying influencers was effective, as it highlighted the most connected users in the TikTok network. This method can be extended to other social media platforms for influencer marketing, trend analysis, and audience targeting. By understanding the network structure, businesses can collaborate with influencers who have the greatest reach and impact on social media.  
  
**Future Work:**  
Future improvements could include the integration of other centrality measures (such as betweenness or closeness centrality) to provide a more comprehensive view of influencer influence. Additionally, the analysis could be extended to real-time data and incorporated into a dynamic influencer recommendation system.