PRACTICAL NO: 7

AIM: Social Network Analysis for influencer identification.

---CODE---

```
# !kaggle datasets download -d goyaladi/twitter-dataset
# !unzip /content/twitter-dataset.zip

import pandas as pd
import networkx as nx

data = pd.read_csv('/content/twitter_dataset.csv')

G = nx.DiGraph()

for _, row in data.iterrows():
    G.add_node(row['Username'])
    G.add_edge(row['Username'], row['Username'], weight=row['Likes'])

    if row['Retweets'] > 0:
        for _ in range(row['Retweets']):
            G.add_edge(row['Username'], row['Username'], weight=1)

pagerank = nx.pagerank(G)

influencers = pd.DataFrame({
    'Username': list(G.nodes),
    'pagerank': pagerank.values()
})

top_influencers = influencers.nlargest(10, 'pagerank')
print(top_influencers)
```

---OUTPUT---

```
         Username  pagerank
0         julie81  0.000106
1   richardhester  0.000106
2  williamsjoseph  0.000106
3     danielsmary  0.000106
4      carlwarren  0.000106
5  ramirezmikayla  0.000106
6     fieldsbrian  0.000106
7           jgood  0.000106
8     turneredgar  0.000106
9    audreymooney  0.000106
```