

Analisis Jaringan Sosial: Mengidentifikasi Komunitas di twitter dengan Tagar #FREEPALESTINE Menggunakan Metode Louvain

Muhammad Dzaki Arta- 13522149¹

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia

¹author@itb.ac.id

Abstract—Makalah ini menyajikan data dari hastag Free Palestine dengan menggunakan sebuah metode untuk mendeteksi komunitas dalam Free Palestine. Dengan algoritma deteksi komunitas untuk mengidentifikasi komunitas pengguna yang membahas Free Palestine dibutuhkan hubungan antara seorang pengguna dengan pengguna lain yang direpresentasikan dalam sebuah graf. Metode ini dapat mendeteksi sebuah komunitas dan memberikan sajian data kepada orang-orang untuk melihat banyaknya interaksi terhadap hastag Free Palestine dan data mejadi analisis pertama dalam melihat dukungan dan seberapa banyak pembahasan penggunaan jejaring sesosial twiter membahasnya.

Keywords—Pendeteksian Komunitas, Metode Louvain, Free Palestine, Data Twiter, Teori Graf

I. PENDAHULUAN

Konflik yang Palestine menjadi topik pembahasan yang sedang ramai diperbincangkan. Krisis kemanusiaan yang terjadi disana membuat banyak khalayak rame mulai menyuarakan untuk dihentinya genosida yang terjadi pada Palestine. Media sosial adalah salah satu tempat mengekspresikan dan menyampaikan pendapat kepada ruang publik salah satunya adalah Twiter. Tagar #FREEPALESTINE sampai menjadi tagar yang paling banyak digunakan hingga sempat menjadi tagar trending di twitter.

Interaksi antar pengguna akan menimbulkan komunitas-komunitas, sehingga pada makala ini akan menggunakan sebuah metode untuk mendeteksi komunitas tersebut berdasarkan data yang diperoleh melalui twitter dengan tagar #FREEPALESTINE. Dengan menggunakan teori graf sebagai representasi hubungan interaksi pengguna, kemudian menggunakan sebuah algoritma pendeteksi komunitas.

Malaka ini akan memberikan hasil berupa data komunitas yang membahas tagar tersebut sehingga didapatkannya seberapa banyak dan aktifnya tagar tersebut di gunakan dan bisa menjadi representasi banyaknya dukungan pengguna jejaring sosial terhadap tagar #FREEPALESTINE.

II. DASAR TEORI

A. Pengertian dan Analisis Jaringan Sosial

Situs jejaring sosial merupakan situs yang dapat membantu

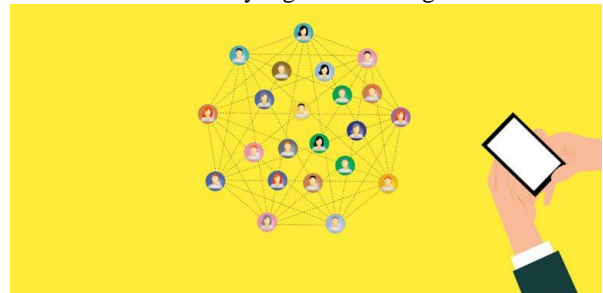
seseorang untuk membuat sebuah profil dan kemudian dapat menghubungkan dengan pengguna lainnya. Situs jejaring sosial adalah aplikasi yang memungkinkan pengguna untuk terhubung menggunakan profil pribadi atau akun pribadinya. Adapun contoh berbagai situs jejaring sosial seperti Friendster, Facebook, Foursquare, Myspace, Twitter, juga termasuk Blackberry Messenger (Juditha, 2011)

Analisis jaringan sosial merupakan bidang studi yang mempelajari koneksi dan hubungan antara individu atau organisasi dalam suatu jaringan sosial. Salah satu jenis analisis ini sering digunakan dalam memahami dinamika dan struktur jaringan sosial, seperti komunitas dan pembentukan kelompok atau penyeberang informasi.

Dalam analisis jaringan sosial, sebuah jaringan biasanya diwakili sebagai graf yang mana simpul mewakili individu atau organisasi dan tepi mewakili hubungan antara mereka. Berbagai algoritma kemudian diterapkan pada graf tersebut untuk mengidentifikasi pola dan tren pada jaringan tersebut, seperti sentralistis, deteksi komunitas, dan agregasi. (Alisha, 2022)

B. Komunitas dalam media sosial

media sosial merupakan media (alat/wadah) di internet yang memungkinkan individu-individu berkomunikasi, bertukar informasi bahkan bekerja sama dalam ruang publik virtual yang membentuk ikatan sosial secara virtual. Kemampuan media sosial juga terbukti menjangkau berbagai hal termasuk membuka peluang dalam menjalin pertemanan secara virtual. Hal ini memungkinkan munculnya jaringan-jaringan pertemanan hingga membentuk berbagai kelompok komunikasi dalam dunia virtual atau yang disebut dengan virtual community

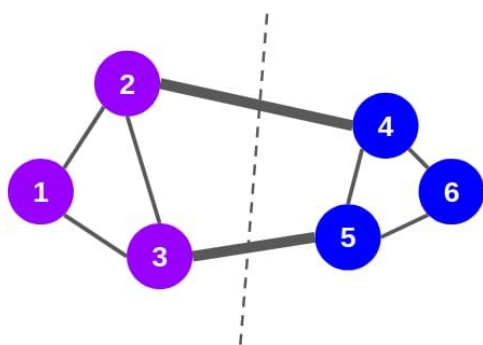


(Gambar 2.1 Komunitas Jejaring Media Sosial)

virtual community merupakan individu-individu yang terlibat

secara bersama, kesamaan pikiran, daya tarik yang membentuk ruang sosial dilingkungan digital sehingga membentuk satu ikatan dan budaya. Dalam media sosial setiap individu selalu terlibat dalam *virtual community*. Ini menggaris bawahi bahwa manusia (agen/individu) adalah makhluk yang memiliki naluri untuk selalu bersama dan menjalin interaksi dengan individu lainnya. Jika individu tidak dapat menggunakan satu bentuk komunikasi untuk terhubung secara efektif, individu akan menemukan yang lain, karena komunikasi dengan orang lain adalah hal mendasar bagi siapa dan bagaimana kita sebagai manusia.

Dengan banyaknya komunitas yang ada dan terjadi dalam media sosial maka komunitas-komunitas ini harus di deteksi untuk melihat seberapa jauh cakupan dan pembahasan yang menjadi bahan interaksi dalam komunitas tersebut. Berbekal sebuah algoritma yang dapat memartisi jaringan tersebut ke dalam beberapa komunitas didapatkannya beberapa metode dalam graf untuk mendeteksi komunitas yaitu: sentralistis, deteksi komunitas, dan agregasi.



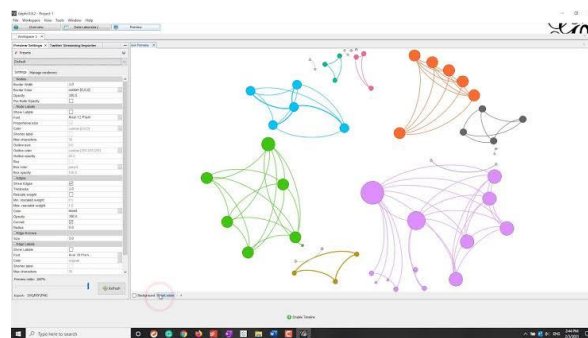
(Gambar 2.2 Komunitas dalam jejaring sosial)

C. Gephi

Gephi merupakan sebuah aplikasi software open source untuk menganalisis jaringan. Gephi dapat membantu menganalisis secara intuitif, Gephi juga menggunakan mesin render 3D untuk menampilkan grafik secara real time. Dapat digunakan untuk menganalisis, memfilter, memanipulasi dan mengeksport semua jenis grafik. Dalam makalah ini gephi digunakan karena Gephi mengimplementasikan metode Louvain dengan antarmuka sehingga dapat memudahkan dalam melakukan analisis.

Salah satu fitur utama yang membuat Gephi menjadi pilihan dalam penelitian ini adalah implementasi metode Louvain yang terintegrasi. Melalui antarmuka yang disediakan oleh Gephi, pengguna dapat dengan mudah mengaplikasikan metode Louvain untuk analisis jaringan. Metode Louvain sendiri merupakan pendekatan yang efektif untuk mendeteksi struktur komunitas dalam jaringan, sehingga Gephi menjadi alat yang sangat berguna untuk mendukung penelitian ini.

Dengan Gephi, peneliti dapat melakukan analisis mendalam, melakukan filter data, melakukan manipulasi grafik, dan mengeksport hasil analisis dengan lebih efisien. Dengan kombinasi antarmuka yang ramah pengguna dan kemampuan analisis yang canggih, Gephi memberikan kontribusi positif dalam mempermudah pemahaman dan interpretasi data jaringan dalam penelitian ini.



(Gambar 2.3 Antarmuka Gephi)

D. Data Scraping

Dalam era internet modern, keberadaan berbagai situs web dan aplikasi membuat data dan informasi berputar secara cepat. Informasi ini memiliki nilai yang sangat tinggi, terutama dalam konteks bisnis, di mana dapat memberikan keuntungan strategis dan mengungguli pesaing. Meskipun metode manual masih ada, namun menjadi kurang praktis dan lambat ketika berurusan dengan jumlah data yang besar pada sebuah situs web.

Oleh karena itu, teknik yang cepat, efisien, praktis, dan akurat diperlukan, dan inilah yang disebut sebagai data scraping. Dalam dunia bisnis yang kompetitif saat ini, data scraping menjadi opsi yang meningkatkan produktivitas, mendorong perkembangan, dan meraih kesuksesan perusahaan serta bisnis digital. Dukungan dari berbagai alat data scraping juga membuat pengumpulan informasi dari sebuah website menjadi lebih efisien.

Teknik data scraping dianggap mudah dilakukan karena dapat memberikan informasi yang dapat dibaca dan nyata. Meskipun ada risiko keamanan, tetapi data scraping memberikan kemampuan untuk mengumpulkan dan memanipulasi informasi dengan cepat. Jika Anda mencari cara untuk mengumpulkan data dan informasi, berikut adalah penjelasan singkat tentang data scraping.

Data scraping, atau ekstraksi data, adalah metode otomatisasi yang memungkinkan pengguna mengekstrak data dari website, database, aplikasi enterprise, atau sistem legacy. Data yang diekstrak dapat disimpan dalam format tabular atau spreadsheet. Proses otomatisasi ini menggantikan metode manual copy-paste yang memakan waktu berjam-jam atau bahkan berhari-hari.

Data scraping digunakan dalam berbagai konteks, seperti penelitian konten website, perbandingan harga dalam bisnis, atau riset pasar pada sumber data publik. Data pada website umumnya bersifat tidak terstruktur dalam format HTML, namun melalui data scraping, dapat diubah menjadi format terstruktur dalam spreadsheet atau database, memungkinkan manipulasi lebih lanjut. Ada berbagai cara untuk melakukan data scraping, termasuk penggunaan layanan online, API tertentu, atau pengembangan kode khusus oleh perusahaan.

Salah satu cara efektif adalah dengan memanfaatkan API dari situs web besar seperti Google, Twitter, atau Facebook, yang memungkinkan akses data dalam format terstruktur. Namun, metode ini tidak berlaku untuk semua website, terutama yang tidak memiliki API atau tidak mengizinkan akses ke data dalam bentuk terstruktur.

III. METODOLOGI PENELITIAN

A. Figures and Tables

Pada makalah ini, dilakukan pengrayapan data pada twitter menggunakan API Twitter. Data yang dikumpulkan adalah informasi terkait pengguna yang paling berpengaruh di twitter terkait topik piala dunia dengan tagar #FIFAWorldCup. Alat yang digunakan untuk memperoleh data tersebut adalah tweepy. Berikut merupakan potongan kode yang digunakan untuk melakukan pengrayapan data menggunakan tweepy.

```
qry=#FREEPALESTINE # Query yang dicari
maxTweets = 1000 # jumlah max tweet yg dicari
tweetsPerQry = 100 # <101, Aturan Twitter
fName=#FREEPALESTINE.json # Nama File hasil Crawling
max_id, tweetCount = -1, 0
```

```
print("Mengunduh maksimum {0} tweets".format(maxTweets))
with open(fName, 'w') as f:
    while tweetCount < maxTweets:
        if max_id <= 0:
            new_tweets = api.search_tweets(q=qry,
            count=tweetsPerQry)
        else:
            new_tweets = api.search_tweets(q=qry,
            count=tweetsPerQry, max_id=str(max_id - 1))

        if not new_tweets:
            print("\nTidak ditemukan lagi tweet dengan
            Query=" + qry).format(qry)
            break

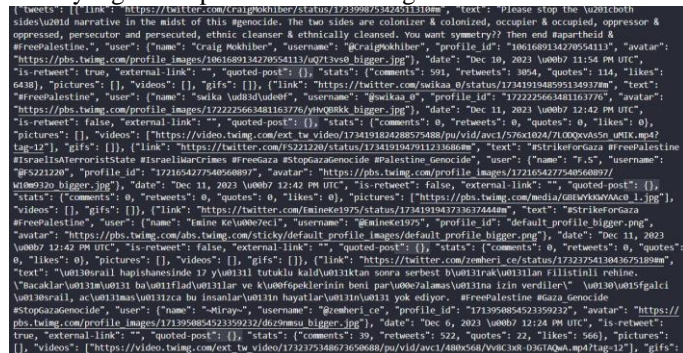
        for tweet in new_tweets:
            f.write(jsonpickle.encode(tweet._json,
            unpicklable=False) + '\n')

            tweetCount += len(new_tweets)
            sys.stdout.write("\r")
            sys.stdout.write("Tweets tersimpan: %.0f" % tweetCount)
            sys.stdout.flush()

            max_id = new_tweets[-1].id

    print("\nSelesai! {0} tweets tersimpan di
    "{1}").format(tweetCount, fName))
```

Hasil yang dikumpulkan adalah sebagai berikut



(Gambar 3.1 Data yang berhasil dikumpulkan menggunakan tweepy)

Dataset tersebut memiliki atribut seperti tanggal dibuat, ID,

status, reply, dan lainnya

B. Pemrosesan Data

Data yang sudah didapat akan diproses sehingga akan mendapatkan graf dari file json yang sudah didapat Representasi graf pada makalah ini menggunakan module python NetworkX.

Berikut adalah kode untuk pemrosesan data pada python ini

```
import json
filename = '#FREEPALESTINE.json'
tweet_list = []
with open(filename, 'r') as fopen:
    for line in fopen:
        if line != '\n':
            tweet_list.append(json.loads(line))

def getTweetID(tweet):
    return tweet.get('id')

def getUserIDandScreenName(tweet):
    user = tweet.get('user')
    if user is not None:
        uid = user.get('profile_id')
        screen_name = user.get('username')
        return uid, screen_name
    else:
        return (None, None)

def getRetweetedUserIDandScreenName(tweet):
    retweet = tweet.get('is-retweet')
    if retweet is not None:
        return getUserIDandScreenName(retweet)
    else:
        return (None, None)

def getRepliedUserIDandScreenName(tweet):
    reply_id = tweet.get('in_reply_to_user_id')
    reply_screenname = tweet.get('in_reply_to_screen_name')
    return reply_id, reply_screenname

def getUserMentionsIDandScreenName(tweet):
    mentions = []
    entities = tweet.get('entities')
    if entities is not None:
        user_mentions = entities.get('user_mentions')
        for mention in user_mentions:
            mention_id = mention.get('id')
            screen_name = mention.get('screen_name')
            mentions.append((mention_id, screen_name))
    return mentions

def getQuotedUserIDandScreenName(tweet):
    quoted_status = tweet.get('quoted_status')
    if quoted_status is not None:
        return getUserIDandScreenName(quoted_status)
    else:
        return (None, None)

def getAllInteractions(tweet):
    # Mendapatkan user id dan screen name pengguna
    tweeter = getUserIDandScreenName(tweet)
    # Tidak melakukan apa-apa bila tweeter tidak ada
    if tweeter[0] is None:
        return (None, None), []
    # Digunakan struktur data set untuk menghindari duplikat
    interacting_users = set()
    # Menambahkan user yang di reply
    interacting_users.add(getRepliedUserIDandScreenName(tweet))
    # Menambahkan user yang di retweeted
    interacting_users.add(getRetweetedUserIDandScreenName(tweet))
    # Menambahkan user yang di quoted
```

```

interacting_users.add(getQuotedUserIDandScreenName(tweet))
# Menambahkan user yang di mention
interacting_users.update(getUserMentionsIDandScreenName(tweet))
# Mengapus tweeter jika terdapat didalam set
interacting_users.discard(tweeter)
# Menghapus none jika ada
interacting_users.discard((None, None))
# Mengembalikan tweeter dan list interaksi
return tweeter, list(interacting_users)

import networkx as nx

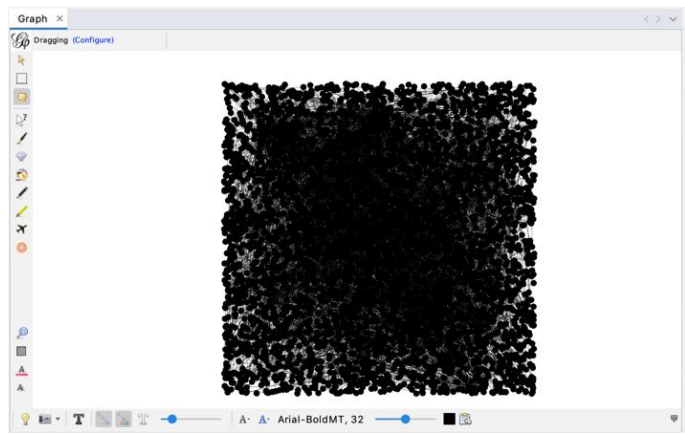
# Menggunakan graf berarah karena sisi mengarah dari user ke
# tweet yang dituju (reply, retweet, quote, mention)
G = nx.Graph()

# Perulangan untuk semua tweet dan menambahkan sisi
for tweet in tweet_list:
    # Menemukan influencer pada tweet
    tweeter, interactions = getAllInteractions(tweet)
    tweeter_id, tweeter_name = tweeter
    tweet_id = getTweetID(tweet)

    # Perulangan untuk menambahkan sisi ke graf
    for interaction in interactions:
        interact_id, interact_name = interaction
        G.add_nodes_from(G.nodes(data=True)) # Menambahkan node
        interact ke graf (jika belum ada)
        G.add_edges_from(G.edges(data=True)) # Menambahkan sisi ke
        graf
        G.add_edge(tweeter_id, interact_id, tweet_id=tweet_id)
        G.nodes[tweeter_id]['name'] = tweeter_name
        G.nodes[interact_id]['name'] = interact_name

# Dapat menggunakan G.nodes(data=True) dan G.edges(data=True) untuk
mendapatkan informasi node dan edge

```



(Gambar 4.2 Visualisasi graf menggunakan depth)

Bisa kita lihat ada banyak node dan sisi yang ada pada gambar tersebut. Setiap node merepresentasikan sebuah tweet dan setiap sisi merepresentasikan reply tweet yang ada. Pada gambar 4.2 bisa dilihat dari data set banyak tweet yang merepresentasikan sebagai simpul yang berwarna hitam.

B. Analisis Derajat

Berikut merupakan hasil simpul dengan derajat tertinggi. Simpul dengan derajat tertinggi dapat diidentifikasi sebagai pengguna dengan pengaruh atau influence terbesar dalam komunitasnya.

(Tabel 4.1 Tabel dengan ID dan derajat terbesar)

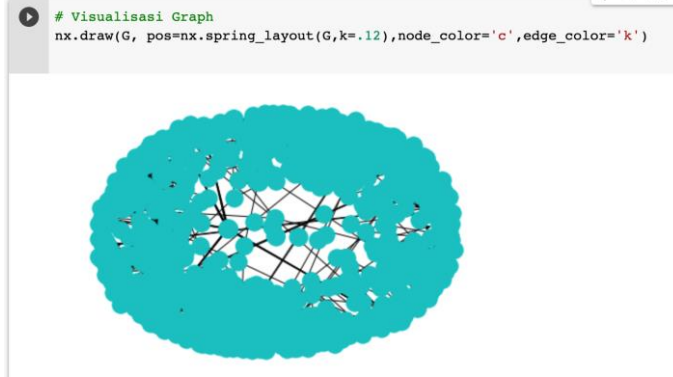
ID	Label Username	Derajat
1486702112910229508	@AvivaKlompas	6449
1061689134270554113	@CraigMokhiber	591
1579442135207710721	@DiazCanelB	591
886593368054222849	@drhossamsamy65	511
1377673590322118662	@uzunabdurrahman	451
1437085900366811145	@yeftale	333
1581282051508080641	@subjectiveviews	293
1591034558920724480	@Naila_Ayad	277
1206744523226914816	@DerbyChrisW	202
1698207856598691840	@Goaldorak2	191

Graf yang dibuat pada makalah ini merupakan graf berarah dimana setiap simpul merupakan pengguna pada media sosial twitter dan sisi merupakan interaksi atau hubungan diantara kedua pengguna.

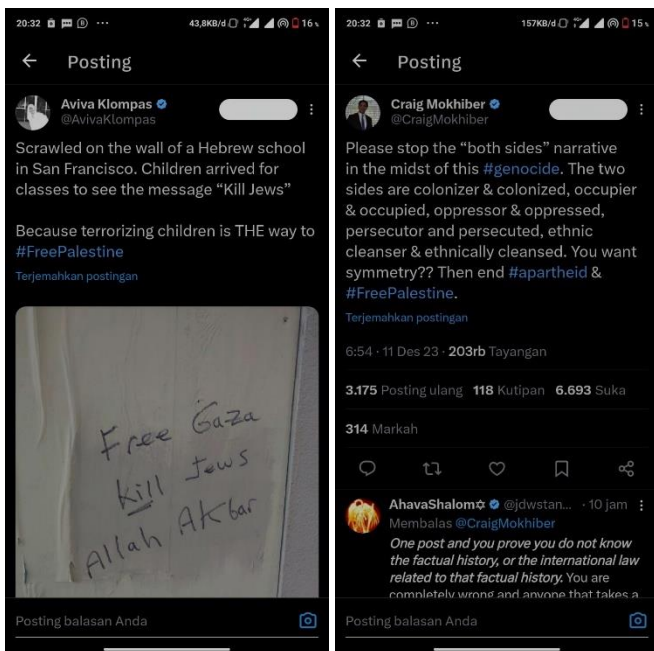
IV. ANALISIS DAN PEMBAHASAN

A. Analisis Graf

Berikut adalah visualisasi graf yang telah didapat.



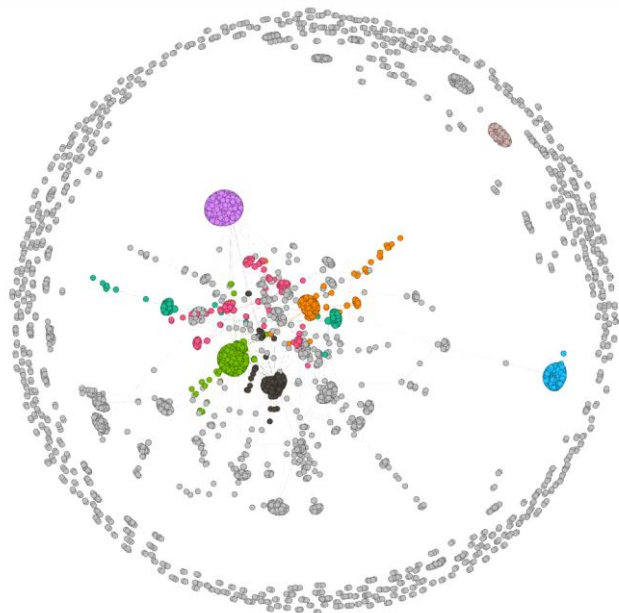
(Gambar 4.1 Visualisasi Graf)



(Gambar 4.3 gambar 2 buah tweet yang memiliki derajat paling banyak)

Dari data tersebut kita bisa lihat pengguna lihat bahwa kedua akun tersebut merupakan public figure yang secara otomatis akan memiliki reply terbanyak dan otomatis akan memiliki derajat paling tinggi

C. Visualisasi dengan Gephi



(Gambar 4.4 Visualisasi Komunitas #FREEPALESTINE menggunakan gephi)

Berikut merupakan hasil pendeteksian komunitas yang didapatkan. Pada graf tersebut, terdapat 8 komunitas utama yang diwakili oleh warna-warna yang berbeda. Perlu diperhatikan bahwa terdapat banyak pengguna yang tidak berada dalam komunitas tersebut. Dalam kasus ini, dapat dianalisis bahwa banyak pengguna yang menggunakan tagar #FREEPALESTINE tetapi tidak berpartisipasi secara aktif dengan pengguna pengguna lainnya.

Ketika setiap komunitas tersebut diperbesar, terdapat satu simpul utama dengan ukuran simpul terbesar. Besar simpul pada visualisasi ini merepresentasikan banyaknya derajat yang dimiliki. Dapat dianalisis bahwa ID simpul terbesar pada setiap komunitas terdapat pada table 4.1., yang menunjukkan bahwa simpul tersebut memiliki derajat terbesar.

V. KESIMPULAN

Pada makalah ini, dengan memanfaatkan teori graf dalam mendeteksi komunitas-komunitas dalam jejaring sosial khususnya adalah Twitter dengan tagar #FreePalestine. Sehingga didapatkannya seberapa besar tagar tersebut menjadi topik perbincangan. Disini kita dapat mengetahui topik-topik apa saja yang dibahas dalam tagar tersebut, seberapa besar pengaruh dari tagar tersebut, isu-isu yang beredar, informasi terbaru yang diterima dan tidak di terima oleh penggunaan jejaring sosial dikarenakan ada beberapa hoax didalamnya, dan diapkannya individu-individu yang berperan aktif dalam tagar tersebut.

Dengan jejaring sosial Twitter yang marak digunakan pada tahun 2023 dengan tagar yang tengah trending yaitu #FreePalestine dengan mendeteksi komunitas dengan teori graf akan di didapatkannya informasi-informasi yang dapat membantu mengetahui komunitas mana yang berperan aktif dan komunitas mana yang memberi informasi valid atau hoax dengan parameter penentu adalah penggunaan jejaring sosial yang memberi komentar atau balasan yang setuju ataupun tidak setuju dan membuat pemahaman mengenai interaksi sosial manusia di jejaring sosial.

VI. LAMPIRAN

File dataset yang digunakan pada pengerjaan makalah ini : <https://drive.google.com/file/d/1kiqaawAhjKOGcX9yea3OoowSqAwH9zVL/view?usp=sharing>

VII. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada pihak-pihak yang membantu dalam menyelesaikan makalah ini tepat pada waktunya,

terutama kepada:

1. Tuhan Yang Maha Esa, berkat rahmat dan karunia-Nya, makalah ini dapat terselesaikan tepat pada waktunya
2. Orang tua yang telah memberikan semangat serta dukungan
3. Teman-teman seperjuangan yang memberi semangat serta kritik membangun.
4. Dr. Fariska Zakhralatifa Ruskanda, S.T., M.T., selaku dosen mata kuliah IF2120 Matematika Diskrit yang kesempatan dan ilmu dalam menyelesaikan makalah ini.
5. Teman-teman Angkatan yang telah bersedia menemani selama satu semester ini.

Semoga atas semua kebaikan yang diberikan akan dibalas oleh Tuhan yang Maha Esa dan sekitarnya makalah ini dapat bermanfaat baik bagi penulis maupun pembaca.

REFERENSI

- [1] Joshi, P. (2020, April 13). Getting started with community detection in graphs and Networks. Analytics Vidhya. Diakses pada tanggal 10

Desember 2022. from
<https://www.analyticsvidhya.com/blog/2020/04/community-detectiongraphs-networks/>

- [2] Representation of the steps applied by the Louvain method on a ... (n.d.). Diakses pada tanggal 10 Desember 2022., from https://www.researchgate.net/figure/Representation-of-the-steps-appliedby-the-Louvain-method-on-a-synthetic-graph-At-each_fig1_221016951
- [3] Munir, Rinaldi. 2022. Graf (Bag. 1): Bahan Kuliah IF2120 Matematika Diskrit. Merupakan slide bahan ajar perkuliahan. Diakses pada tanggal 9 Desember 2022
- [4] Munir, Rinaldi. 2022. Pohon (Bag. 2): Bahan Kuliah IF2120 Matematika Diskrit. Merupakan slide bahan ajar perkuliahan. Diakses pada tanggal 9 Desember 2022.
- [5] Munir, Rinaldi. 2022. Pohon (Bag. 3): Bahan Kuliah IF2120 Matematika Diskrit. Merupakan slide bahan ajar perkuliahan. Diakses pada tanggal 9 Desember 2022.

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 3 Desember 2023



Nama dan NIM