



Social Network Analysis untuk Identifikasi Pengguna Twitter Berpengaruh pada Topik Bencana Gempa dan Tsunami di Indonesia

Ibnu Santoso^{#1}, Siskarossa Ika Oktora^{*2}, Siti Muchlisoh^{**3}, Ernawati Pasaribu^{*4}

[#]D-IV Komputasi Statistik, Politeknik Statistika STIS,
Jln. Otto Iskandardinata No. 64C, Jakarta Timur, 13330
¹ibnu@stis.ac.id

^{*}D-IV Statistika, Politeknik Statistika STIS,
Jln. Otto Iskandardinata No. 64C, Jakarta Timur, 13330
²siskarossa@stis.ac.id

⁴ernapasaribu@stis.ac.id
^{**}D-III Statistika, Politeknik Statistika STIS,
Jln. Otto Iskandardinata No. 64C, Jakarta Timur, 13330
³sitim@stis.ac.id

Abstrak—Indonesia merupakan negara yang rawan terjadi bencana alam seperti gempa dan tsunami. Seiring dengan perkembangan teknologi, arus informasi mengenai kebencanaan juga mengalir di media sosial seperti Twitter. Penggunaan Twitter dalam kaitannya dengan kebencanaan telah banyak diteliti antara lain untuk penyebarluasan informasi, alat manajemen dan pengurangan resiko, pemantauan aktivitas tanggap darurat, dan lain-lain. Penelitian ini bertujuan untuk mengidentifikasi pengguna twitter berpengaruh khusus untuk topik bencana gempa dan tsunami di Indonesia dengan menggunakan *Social Network Analysis* (SNA) dengan dan tanpa mempertimbangkan faktor *frequency* dan *engagement*. Hasil SNA tanpa mempertimbangkan faktor *frequency* dan *engagement* menunjukkan bahwa pengguna Twitter yang dinilai paling berpengaruh pada topik bencana gempa dan tsunami adalah situs berita seperti detikcom dengan *influence score* sebesar 0,77. Sedangkan jika mempertimbangkan faktor *frequency* dan *engagement* menunjukkan bahwa pengguna Twitter yang dinilai paling berpengaruh pada topik bencana gempa dan tsunami adalah akun infoBMKG dengan indeks *influence score* sebesar 0,63. Berdasarkan hasil penelitian ini ditemukan bahwa BMKG telah berperan penting dalam pemberian informasi mengenai bencana gempa bumi dan tsunami di Indonesia dan mendapatkan kepercayaan luas dari masyarakat yang ditunjukkan dengan adanya *engagement* yang lebih tinggi dibandingkan akun lainnya.

Kata kunci— *Influence Score Index*, Gempa, *Social Network Analysis*, Tsunami, Twitter

I. PENDAHULUAN

Indonesia merupakan negara kepulauan yang terletak pada tiga lempeng besar dunia (Indo-Australia, Pasifik

dan Eurasia) yang terus aktif bergerak. Kondisi tersebut menyebabkan Indonesia menjadi negara yang rawan bencana alam seperti gempa bumi dan tsunami. Bencana alam yang terjadi tidak hanya menyebabkan kerugian material seperti kerusakan bangunan fisik, namun juga dapat menyebabkan trauma korban seperti depresi dan kekhawatiran tentang terjadinya bencana alam di masa depan [1]. Selama kurun waktu 2000 hingga 2021 telah terjadi beberapa bencana besar di Indonesia, diantaranya tsunami di Aceh pada 26 Desember 2004 akibat gempa berkekuatan magnitudo 9,3; tsunami di Pangandaran pada 17 Juli 2006 berkekuatan magnitudo 8; serta tsunami di Palu dan Donggala pada 28 September 2018 akibat gempa berkekuatan magnitudo 7,4.

Perkembangan teknologi menyebabkan arus informasi tidak lagi terbatas ruang dan waktu. Salah satu platform yang paling banyak digunakan saat ini adalah media sosial seperti twitter. Kejadian bencana alam yang tidak dapat diprediksi seringkali membuat penanganan pasca bencana sulit dilakukan. Dalam kedaruratan bencana, pemenuhan kebutuhan dasar serta pertolongan medis sangat diperlukan. Dalam situasi tersebut media sosial daring seperti twitter dapat dijadikan solusi untuk membantu proses penanggulangan bencana [2]. Twitter mampu meningkatkan kesadaran masyarakat akan situasi tertentu, memfasilitasi koordinasi antar respon, serta mengurangi jeda waktu antara suatu kejadian (bencana alam) dengan tindakan yang dapat diambil. Twitter juga berfungsi sebagai media pertukaran informasi selama kejadian bencana alam serta memberdayakan aktor-aktor baru dalam proses penanggulangan bencana alam [3]. Selain itu twitter juga dapat digunakan untuk menganalisis masalah

publik karena memungkinkan penggunaanya (masyarakat) untuk secara bebas menyampaikan pendapat serta menyebarkan informasi mengenai peristiwa terkini, termasuk bencana alam [4].

Berbagai penelitian untuk mengetahui respon masyarakat terhadap kejadian bencana alam melalui twitter telah cukup banyak dilakukan, diantaranya penelitian yang dilakukan oleh [5] yang mengkaji penggunaan twitter dalam komunikasi krisis selama terjadinya bencana alam yang membuat orang lain di seluruh dunia mengetahui situasi terkini dan dapat memberikan bantuan kepada korban sesuai dengan kebutuhan. Studi tersebut menggambarkan bagaimana Walikota Houston, Texas, Sylvester Turner, menggunakan Twitter selama Badai Harvey pada bulan Agustus dan September 2017. Penelitian tersebut menunjukkan bahwa penggunaan twitter oleh Walikota Turner merupakan contoh praktik terbaik untuk menggunakan media sosial dalam komunikasi krisis kebencanaan.

Pada penelitian yang dilakukan oleh [6], optimalisasi komunikasi pada twitter digunakan sebagai alat manajemen dan pengurangan risiko bencana alam. Twitter juga dianggap sebagai sumber data baru yang dapat dimanfaatkan untuk memantau aktivitas tanggap darurat selama fase bencana alam. Dalam penelitiannya, [7] menganalisis penggunaan twitter dengan menganalisis pola sentimen pada tiga fase berbeda yakni kesiapsiagaan, respon, dan pemulihan saat Badai Isaac melanda Louisiana dan Mississippi pada Agustus 2012.

Terkait pemanfaatan twitter pada kejadian bencana alam di Indonesia, [8] menyimpulkan bahwa penggunaan data pada twitter dapat digunakan untuk perencanaan karena mendukung estimasi cakupan kejadian bencana alam, dapat melihat pola spatio-temporal, serta mengidentifikasi opini dari para pemangku kepentingan. Hasil penelitian yang dilakukan oleh [9] dengan cara meneliti "jejak" tweet tentang bencana menyimpulkan bahwa respon masyarakat perkotaan terhadap bencana Siklon Tropis Cempaka yang melanda pantai selatan Pulau Jawa pada bulan November tahun 2017 bersifat positif dengan banyaknya postingan kata-kata doa dan simpati, menyebarkan informasi untuk meningkatkan kesadaran bencana, dan mengorganisir penggalangan dana untuk para korban.

Berbagai penelitian tersebut merupakan hal yang baik, namun adakalanya pengguna twitter yang sangat bervariasi dapat berakibat pada mis-informasi yang diterima oleh masyarakat. Misalnya volatilitas yang tinggi dari respon sebuah *tweet* tidak berasal dari otoritas yang berkepentingan. Padahal sistem manajemen risiko bencana alam yang melibatkan twitter membutuhkan strategi keterlibatan otoritas resmi yang sangat komprehensif seperti BMKG, pimpinan daerah, *channel* berita resmi, *influencer* yang terpercaya, dan sebagainya.

Secara sederhana seorang *influencer* dapat diartikan sebagai *user* yang dapat memengaruhi aksi atau hal yang dilakukan oleh banyak orang di media sosial, atau dapat juga diartikan sebagai orang yang bisa menyebarkan

informasi yang ada di media sosial [10]. Besarnya pengaruh yang dapat diberikan oleh *user* yang ada di media sosial secara sederhana dapat ditentukan dengan cara melihat profil yang dimiliki dari *user* tersebut, seperti berapa banyak *followers*, *following* yang dimiliki oleh *user*, atau dapat juga dengan menghitung rasio dari keduanya. Namun ketika berbicara topik tertentu, bisa jadi *user* tersebut tidak memiliki cukup pengaruh [11].

Pada saat ini, metode sederhana tersebut telah mengalami perkembangan, metode yang paling sering digunakan pada saat ini adalah *Social Network Analysis* (SNA). SNA bertujuan untuk mengidentifikasi *most relevant actors* atau orang yang memiliki pengaruh paling kuat dalam *social network* [12]. Dalam SNA, *social network* dapat direpresentasikan sebagai suatu *graph*, dimana *user* direpresentasikan sebagai *node*, sedangkan *garis* merepresentasikan hubungan yang dimiliki oleh masing-masing *user* seperti melakukan *retweet*, *likes*, atau *replies*.

Untuk beberapa kasus, akan lebih baik mengidentifikasi pengaruh dari seorang *user* berdasarkan topik tertentu. Hal ini dapat dilakukan dengan cara melihat dari topik apa yang sering dibahas oleh *user* tersebut. Menurut referensi [13] metode perhitungan yang hanya melihat pada suatu topik tertentu saja akan lebih efektif daripada melakukan perhitungan secara umum (tanpa topik). Metode perhitungan ini digunakan untuk analisis secara *offline*. Dimana data dari twitter didapatkan melalui *scraping* pada periode tertentu. *Tweet* yang diambil sesuai dengan topik yang ingin dianalisis menggunakan *keyword* atau kata kunci yang bersesuaian dengan topik tersebut.

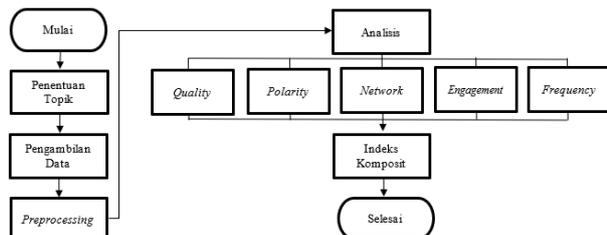
Walaupun demikian perhitungan terkait dengan *influence score* tersebut masih sering mengalami kritik oleh berbagai pihak, khususnya terkait dengan skor yang dihasilkan yang bertujuan untuk menentukan apakah *user* yang bersangkutan memiliki pengaruh atau tidak terkait dengan suatu topik [11]. Oleh karena itu, dalam penelitian ini akan dihitung *influence score* menggunakan metode pada penelitian sebelumnya [12]. Kemudian akan coba dibangun indeks komposit untuk menghitung besarnya *influence* yang dapat diberikan oleh masing-masing *user* terhadap suatu topik dan dievaluasi relevansinya dengan *influence score* yang akan dihitung pada topik bencana gempa dan tsunami di Indonesia.

Penelitian ini diawali dengan pendahuluan, diikuti dengan metode penelitian yang mencakup pengumpulan data, *preprocessing*, penghitungan *influence score*. Setelah itu dilakukan analisis berdasarkan hasil pengolahan yang diperoleh dan membentuk kesimpulan.

II. METODE

Penelitian ini diawali dengan pengambilan data pada twitter, *preprocessing* data, analisis data yang terbagi menjadi analisis *quality*, *polarity*, *network*, *engagement*, dan *frequency*, dan pembentukan indeks komposit. Dari indeks komposit yang dihasilkan diharapkan dapat diketahui pengguna yang memiliki pengaruh terbesar pada

topik tertentu. Diagram alir penelitian ditunjukkan pada Gambar 1.



Gambar. 1 Diagram alir penelitian

A. Pengumpulan Data

Data yang digunakan pada penelitian ini merupakan data twitter terkait dengan bencana khususnya gempa dan tsunami di Indonesia. pengambilan data dari twitter tersebut menggunakan *keyword* tertentu. Pemilihan *keyword* sendiri merupakan hal yang penting karena apabila *keyword* yang digunakan terlalu umum, maka *tweet* yang terambil akan terlalu banyak dan tidak sesuai dengan apa yang diinginkan, sedangkan apabila terlalu sempit maka *tweet* yang terambil menjadi sedikit sehingga hasilnya kurang dapat mewakili keadaan sebenarnya. Adapun *keyword* yang digunakan adalah gempa dan tsunami dengan lokasi di Indonesia.

Pengambilan *tweet* yang berkenaan tentang bencana ini dilakukan dengan menggunakan metode *scraping twitter* yang dilakukan dengan menggunakan *library* yang tersedia pada bahasa pemrograman python, yaitu *twint* [14]. *Tweet* yang digunakan adalah *tweet* yang memuat *keyword* pada periode 1 Januari 2010 sampai dengan 28 Maret 2021.

Sebelum dilakukan analisis, perlu dilakukan *preprocessing* terlebih dahulu dengan tujuan membersihkan data. Rangkaian yang dilakukan adalah sebagai berikut:

1. Menghapus duplikasi data, sebuah *tweet* akan memiliki nilai *id_tweet* yang unik, dengan menggunakan *id_tweet* sebagai primary key maka *tweet* yang *redundant* akan dihapus
2. Menghapus karakter yang tidak dibutuhkan. Karakter seperti `~!@#$%^&*()<>{}[]_+|;:'"` tidak diperlukan sehingga dapat dihapus. *Link* URL, angka 0 sampai 9 juga dihapus karena tidak diperlukan dalam analisis
3. *Case Folding*. Merupakan proses untuk merubah semua kata menjadi huruf non kapital. Hal ini dilakukan karena python merupakan bahasa pemrograman yang *case sensitive*, oleh karena itu kata "bencana" dengan kata "Bencana" akan dianggap sebagai kata yang berbeda. Walaupun kata yang dimaksud adalah kata yang sama
4. Normalisasi. Normalisasi teks digunakan untuk mengubah kata singkatan ataupun kata yang salah ketik menjadi bentuk sebenarnya. Misalnya kata "yg" akan diubah menjadi "yang", "tdk" menjadi "tidak" dan sebagainya

5. *Stemming*. Merupakan proses yang dilakukan untuk merubah kata menjadi bentuk bakunya. Misalnya "menjadi" menjadi "jadi", "berkuasa" menjadi "kuasa" dan sebagainya.
6. Menghapus *stopwords*. *Stopwords* atau kata penghubung seperti "yang", "dan" dsb. Kata kata tersebut dihapus karena dapat mengganggu proses analisis karena frekuensinya yang banyak.

B. Menghitung Influence Score

1) *Network Analysis*: Langkah selanjutnya adalah melakukan *network analysis* [15]. *User* yang melakukan *retweet* terhadap *tweet* dari *user* lainnya maka kedua user tersebut akan terhubung. Setelah semua *user* berhasil dihubungkan dan menjadi sebuah graph, *network analysis* yang dilakukan adalah dengan menghitung *Betweenness centrality* (bc), *Eigenvector centrality* (ec), dan *In-degree* (id) dan *Follower-Followee Ratio* (TFF). Untuk menghitung besarnya *network* yang *user* miliki, maka model perhitungan yang dilakukan adalah:

$$U_{\text{network}} = \frac{bc+ec+id+tff}{4} \quad (1)$$

2) *Quality Analysis dan Sentiment Analysis*: Tahapan berikutnya adalah melakukan *quality analysis* [15]. Kualitas yang akan dihitung adalah kualitas dari masing masing *tweet*. Metode perhitungan kualitas yang digunakan adalah *Flesch-Kincaid* [16] dengan menggunakan rumus sebagai berikut:

$$\text{Kualitas} = 0,39 \times \left(\frac{\text{total kata}}{\text{total kalimat}} \right) + 11,8 \times \left(\frac{\text{total suku kata}}{\text{total kata}} \right) - 15,59 \quad (2)$$

Setelah melakukan analisis terhadap kualitas dari sebuah *tweet*, langkah selanjutnya adalah melakukan analisis terkait sentiment dari *tweet* tersebut. *Sentiment analysis* juga pernah dilakukan sebelumnya oleh [17] dan [18]. Menggunakan metode perhitungan analisis sentimen berbasis lexicon (*kamus*). Didalam kamus tersebut terdapat kata dengan nilai polaritas yang dimiliki oleh kata tersebut. Metode *lexicon* ini menggunakan pencocokkan kata yang ada pada *tweet* dengan kata yang ada pada kamus. Apabila skor polaritas pada *tweet* tersebut > 0, maka *tweet* tersebut merupakan *tweet* dengan sentimen positif, begitu juga apabila polaritas < 0, merupakan *tweet* dengan sentimen negatif, dan nilai polaritas = 0 merupakan *tweet* dengan sentimen yang netral.

Setelah didapatkan kualitas dan juga sentimen dari masing masing *tweet*, kemudian dihitung kualitas dan sentimen dari masing masing *user* dengan cara menghitung rata rata dari kualitas dan juga sentimen *tweet* yang *user* tersebut

$$u_{\text{quality}} = \sum_{i=1}^{i \leq n_u} \text{kualitas}_i \times \frac{1}{n_u} \quad (3)$$

$$u_{\text{polarity}} = \sum_{i=1}^{i \leq n_u} \text{sentimen}_i \times \frac{1}{n_u} \quad (4)$$

C. Perhitungan Influence Score

Setelah semua variabel didapat, proses berikutnya adalah menghitung *influential score* untuk masing masing *user*, metode perhitungan yang digunakan adalah metode perhitungan Bigonha [15]. *Influential score* dapat dihitung dengan cara:

$$I_{su} = \frac{\alpha \times u_{polarity} + \varphi \times (\beta \times u_{network} + \gamma \times u_{quality})}{\alpha + \beta + \gamma} \quad (5)$$

D. Menyusun Indeks Komposit

Tahapan penyusunan indeks menggunakan metode yang digunakan pada referensi [19] adalah sebagai berikut:

1. Membangun kerangka kerja teoritis dan pemilihan variabel/indikator.

Kerangka kerja harus secara jelas mendefinisikan fenomena yang akan diukur dan sub-komponennya, dan selanjutnya memilih masing-masing indikator dan bobot yang mencerminkan kepentingan relatif mereka dan dimensi komposit keseluruhan.

2. Standardisasi data

Metode normalisasi yang digunakan dalam penelitian ini adalah normalisasi *Min-max*. Metode *Min-Max* mengubah nilai variabel tersebut menjadi nilai antara 0 sampai 1.

3. Analisis faktor

Pada tahapan ini dilakukan beberapa hal seperti menguji kelayakan dari data yang akan digunakan uji kelayakan data, melakukan reduksi atau pengurangan dari variabel/indikator, membentuk faktor, mengestimasi nilai matriks faktor yang terbentuk, melakukan rotasi dari faktor yang terbentuk, dan menghitung nilai skor faktor.

4. Standarisasi data skor faktor

Pada masing masing model faktor yang terbentuk dilakukan standarisasi. Standarisasi data skor faktor digunakan untuk memperkecil *range* nilai faktor indeks menjadi nilai 0 sampai 1. Dalam penelitian ini metode normalisasi data yang digunakan adalah metode *Min-Max* kemudian dikalikan 100 untuk mempermudah dalam interpretasi.

$$f'_{ij} = \frac{[f_{ij} - \text{Min}(f_i)]}{[\text{Max}(f_i) - \text{Min}(f_i)]} \times 100 \quad (6)$$

keterangan :

f'_{ij} : skor faktor ke-i *user* ke-j setelah standarisasi

f_{ij} : skor faktor ke-i *user* ke-j

$\text{Max}(f_i)$: skor maksimal faktor ke-i

$\text{Min}(f_i)$: skor minimal faktor ke-i

5. Menentukan penimbang yang akan digunakan.

Penimbang yang digunakan dalam penelitian ini adalah *unequal weighting* dan *equal weighting*. *Equal weighting* digunakan pada variabel yang ada di masing masing faktor, sedangkan *unequal weighting* digunakan pada faktor faktor yang terbentuk.

- a. Pembobotan skor faktor

Untuk masing masing variabel yang ada pada sebuah faktor, akan dilakukan metode perhitungan *equal weighting* untuk mendapatkan nilai dari faktor tersebut

$$f'_i = \frac{1}{t} \sum_{j=1}^t Y'_{ij} \quad (7)$$

keterangan :

f'_i : Nilai faktor ke-i

Y'_{ij} : Nilai variabel ke-j faktor ke-i

t : Banyaknya variabel pada faktor ke-i

- b. Pembobotan faktor

Pada pembobotan faktor, metode yang digunakan adalah *unequal weighting*. Penimbang didapatkan dari proporsi *explained variance* oleh setiap faktor dengan total *explained variance*

$$W_i = \frac{\text{Explained variance}_i}{\text{Total explained variance}} \quad (8)$$

$$\sum W_i = 1$$

6. Agregasi.

Menghitung nilai dari *influence score*. Dalam penelitian ini, metode perhitungan yang digunakan adalah agregasi linear (*linear aggregation*). Dimana menjumlahkan semua faktor yang telah dikalikan dengan penimbangnya masing masing

$$\text{Indeks} = \sum_{i=1}^n W_i f'_i \quad (9)$$

keterangan :

Indeks : Indeks *influence* pada seorang *user*

W_i : bobot faktor ke-i

f'_i : skor faktor ke-i

7. Melihat hubungan indeks komposit dengan indeks lain.

Tahapan ini ditujukan untuk menilai seberapa baik kemampuan indeks komposit dalam menjelaskan fenomena yang terjadi. Dalam penelitian ini, akan dilihat hubungan antara nilai *influence score* yang dihasilkan dari indeks komposit dengan *influence score* dari metode referensi [15]. Melihat hubungan serta besarnya hubungan dapat dilakukan dengan cara melakukan analisis uji korelasi *rank spearman*

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n-1)} \quad (10)$$

Uji signifikansi korelasi *Rank Spearman* :

H_0 : Korelasi pada populasi adalah nol

H_1 : Korelasi pada populasi lebih besar dari nol

Statistik hitung untuk data lebih dari 30 :

$$z = r_s \sqrt{n - 1} \tag{11}$$

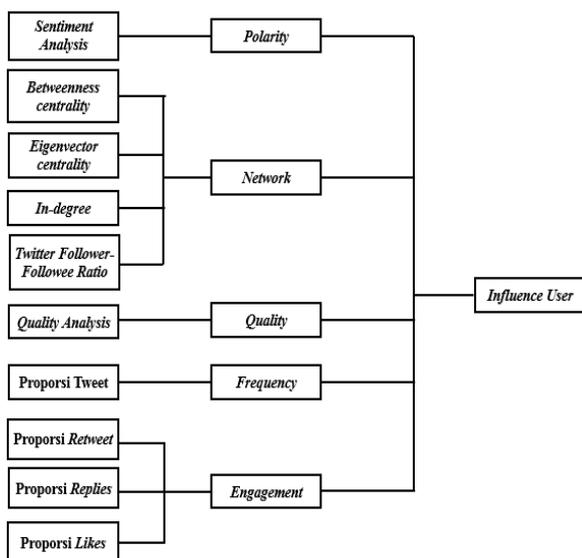
keterangan :

z : Nilai z hitung

r_s : Koefisien korelasi spearman

n : Jumlah sampel penelitian

Berdasarkan [10]–[13], terdapat informasi bahwa besarnya pengaruh seseorang di suatu media sosial juga dapat dipengaruhi oleh beberapa variabel, diantaranya adalah Proporsi *Tweet*, Proporsi *Retweet*, Proporsi *Replies*, Proporsi *Likes*, *Betweenness Centrality*, *Eigenvector Centrality*, *In-Degree Centrality*, *Twitter Follower-Followee Ratio*, *Polaritas Tweet*, dan *Kualitas Tweet* sebagaimana ditunjukkan pada Gambar 2.



Gambar. 2 Komponen Pembentuk Indeks *Influence User*

Variabel *Betweenness Centrality*, *Eigenvector Centrality*, *In-Degree Centrality*, *Twitter Follower-Following Ratio*, *Polaritas Tweet*, dan *Kualitas Tweet* didapatkan dari hasil perhitungan pada metode sebelumnya [15]. Sedangkan pada variabel Proporsi *Tweet*, Proporsi *Retweet*, Proporsi *Replies* dan Proporsi *Likes* didapatkan dengan cara menghitung proporsi banyaknya *tweet*, *retweet*, *replies*, *likes* tentang *tweet* yang berkaitan dengan bencana yang terjadi terhadap seluruh *tweet* yang dikeluarkan oleh *user* dalam rentang waktu penelitian.

$$\text{Proporsi } X = \frac{\text{Banyaknya } X \text{ sesuai topik}}{\text{Total } X \text{ pada periode penelitian}} \tag{12}$$

keterangan

X : *tweet*, atau *retweet*, atau *likes*, atau *replies*

E. Pemilihan Variabel

Pemilihan variabel bertujuan untuk menyeleksi variabel yang akan digunakan untuk membentuk dimensi indeks komposit *influence score* terkait dengan bencana yang terjadi. Sepuluh variabel pembentuk indeks komposit yang akan digunakan perlu diseragamkan terlebih dahulu arahnya untuk mempermudah dalam proses interpretasi. Dalam hal ini semua variabel diubah arahnya menjadi positif.

Setelah dilakukan penyeragaman arah, langkah yang dilakukan selanjutnya adalah melakukan normalisasi. Normalisasi dilakukan karena setiap variabel memiliki satuan yang berbeda-beda dan memiliki nilai *range* atau rentang yang berbeda-beda pula, sehingga dilakukannya normalisasi bertujuan untuk mengurangi *bias* yang terjadi. Indikator yang telah diseragamkan dan dinormalisasi selanjutnya dianalisis menggunakan analisis faktor. Analisis faktor yang dilakukan bertujuan untuk mereduksi dimensi variabel menjadi lebih ringkas dan membentuk faktor. Proses reduksi variabel akan memperhatikan nilai KMO, MSA, komunalitas serta *Bartlett's Test of Sphericity*.

Nilai KMO yang kurang dari 0,5 menandakan bahwa data tersebut belum layak untuk dianalisis. Dalam proses mereduksi variabel, peneliti memperhatikan nilai MSA dan nilai komunalitas. Variabel yang memiliki nilai MSA dan komunalitas kurang dari 0,5 menandakan hubungan yang kurang kuat antara variabel tersebut dengan variabel lainnya, sehingga variabel tersebut dapat dikatakan tidak layak untuk dimasukkan ke dalam model dan dianalisis lebih lanjut, sehingga harus direduksi atau dikeluarkan. Langkah berikutnya adalah pembentukan faktor, perhitungan indeks komposit *influence score*, dan uji korelasi dengan *influence score* yang dihasilkan dengan metode sebelumnya.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan Data dan Preprocessing

Data *tweet* yang diperoleh dari hasil *scraping* sebanyak 457.389 *tweet*. Setelah *preprocessing*, data tereduksi menjadi 262.367 *tweet*. Data pada Tabel 1 merupakan contoh sampel data *tweet* sebelum dilakukan *preprocessing* dan Tabel 2 merupakan contoh sampel data *tweet* setelah dilakukan *preprocessing*.

B. Perhitungan Influence Score

Setelah data *tweet* selesai di-*preprocessing*. Selanjutnya adalah melakukan analisis untuk menghitung nilai dari *influence Score*. Untuk menghitung nilai tersebut diperlukan tiga variabel, yaitu *network*, *polarity*, dan *quality* dilakukan perhitungan pada masing masing *tweet* terlebih dahulu. Tabel 3 merupakan contoh sampel data sentimen dan kualitas dari *tweet*. Setelah perhitungan terkait sentimen dan kualitas *tweet* dilakukan, kemudian

nilai tersebut dikelompokkan berdasarkan *user*, sehingga akan didapat nilai polaritas dan kualitas *tweet* dari masing masing *user*. Perhitungan nilai *influence score* dilakukan setelah variabel *sentiment*, *polarity*, dan *network* dari *user* didapatkan. Tabel 4 merupakan contoh sampel data *network*, *quality*, *polarity* dan *influence square* dari *user*. *Influence score* ini belum mempertimbangkan frequency dan engagement.

TABEL I
CONTOH SAMPEL DATA TWEET SEBELUM PREPROCESSING

Id Tweet	Tweet
1369854870405095428	*[Bencana Alam dan Mitigasinya]* (liat pict biar hemat capt) Sesuatu bisa dikatakan sebagai bencana ketika menimbulkan kerugian. Kenapa bencana alam banyak terjadi di Indonesia? 1. Potensi gempa bumi : Indonesia berada di pertemuan 3 lempeng (Pasifik, Eurasia, Indo-Australia)
1367748495843622913	Gempa Bumi Terkini (USGS) Waktu : 05-03-2021 14:32:54 WIB Lokasi : 157 km SW of Padang, Indonesia Magnitudo : 5.6 M Kedalaman : 17.97 Km Potensi : tidak berpotensi TSUNAMI Koordinat : 1.7901° LS - 99.2082° BT Info : https://t.co/KFGbBfiYqh Sumber : https://t.co/2EZCipYdc5

TABEL II
CONTOH SAMPEL DATA TWEET SETELAH PREPROCESSING

Id Tweet	Tweet
1369854870405095428	bencana alam mitigasinya liat pict biar hemat capt sesuatu kata bencana timbul rugi kenapa bencana alam banyak jadi indonesia potensi gempa bumi indonesia ada temu lempeng pasifik eurasia indo australia
1367748495843622913	gempa bumi kini usgs waktu wib lokasi km sw of padang indonesia magnitudo m dalam km potensi potensi tsunami koordinat ls bt info sumber

TABEL III
CONTOH SAMPEL DATA SENTIMEN DAN KUALITAS TWEET

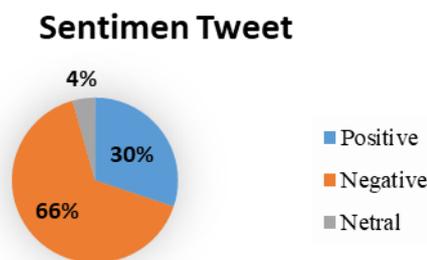
Id Tweet	Sentimen	Kualitas
1373212141059072000	-0.06	7.2
1369854870405095428	-0.22	17.4
1367748495843622913	-0.06	15

TABEL IV
CONTOH SAMPEL DATA NETWORK, QUALITY, POLARITY, DAN INFLUENCE SCORE DARI USER

Username	u_network	u_qualit	u_polarit	influence_score
detikcom	1	0,39394	-0,0841	0,770731364
kompasco	0,352226	0,34988	-0,00086	0,288050458
infoBMK	0,341899	0,38553	0,035221	0,286904151
VIVAcoid	0,298752	0,32320	-0,04216	0,250489163
aagym	0,201618	0,36224	-0,12	0,190471433

C. Analisis Sentimen Terkait Bencana

Untuk melihat gambaran umum terkait dengan pendapat masyarakat terhadap bencana yang terjadi maka dilakukan analisis sentimen. Pada analisis ini data dibagi menjadi tiga berdasarkan sentimen yang dimiliki yaitu positif, negatif, dan netral. Berikut gambaran pendapat masyarakat terkait dengan bencana yang terjadi. Gambar 3 adalah proporsi sentimen yang diolah dari tweet yang didapatkan, dimana 66% adalah sentiment negative, 30% adalah sentiment positif, dan 4% adalah sentiment netral.



Gambar. 3 Persentase Sentimen terhadap Bencana

Secara umum pendapat masyarakat terkait dengan bencana yang terjadi bermuatan negatif, hal ini dapat dibuktikan dengan melihat banyaknya *tweet* yang memiliki polaritas negatif dibandingkan dengan *tweet* yang memiliki polaritas positif maupun netral.

D. Membangun Indeks Komposit

Berdasarkan proses reduksi variabel yang dilakukan, didapatkan 3 variabel yang tereduksi yaitu variabel Twitter *Follower-Following Ratio*, Polaritas *Tweet*, dan Kualitas *Tweet*. Melalui analisis multivariat analisis faktor didapatkan juga terdapat 2 dimensi atau faktor yang terbentuk. Penentuan faktor ini didasarkan pada nilai *eigen value* yang lebih kecil atau sama dengan 1, maka faktor tersebut dianggap dominan dan signifikan. *Weighting* pada masing masing faktor didapatkan melalui perbandingan antara besarnya keragaman yang dapat dijelaskan (*percentage of variance explained*) oleh suatu faktor dan besarnya keragaman total yang dapat dijelaskan oleh model. Ringkasan perhitungan dapat dilihat pada Tabel 5.

TABEL V
RINGKASAN PEMBENTUKAN DAN PENIMBANGAN FAKTOR

No	Nama Faktor	Variabel	% of variance	Penimbang
1	Engagement Rate	Proporsi Retweet	46.6	0.515
		Proporsi Replies		
		Proporsi Likes		
2	Frequency and Network Analysis	Betweenness Centrality	43.9	0.485
		Eigen Centrality		
		In-Degree Centrality		
		Proporsi Tweet		

E. Perhitungan Indeks Influence Score

Dalam memperoleh nilai indeks komposit dilakukanlah penjumlahan hasil perkalian antara penimbang dan skor faktor untuk tiap faktor yang sudah dinormalisasi.

$$IIS = 0.515 * FAKTOR_1 + 0.485 * FAKTOR_2$$

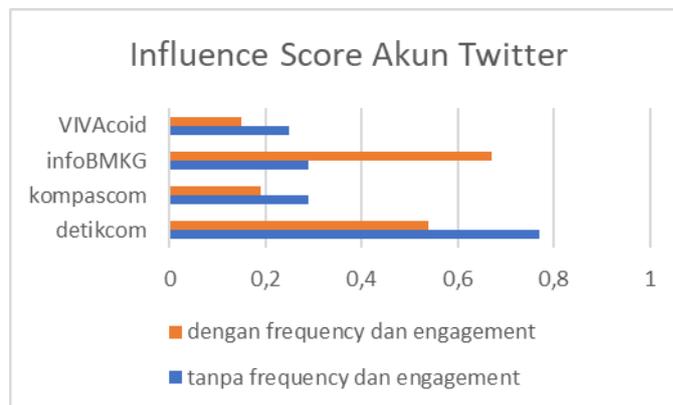
Karena sebelumnya telah dilakukan penyeragaman arah variabel menjadi positif, maka interpretasi dari indeks yang dihasilkan adalah semakin tinggi nilai indeks *influence score* terkait dengan bencana maka semakin besar pula pengaruh yang dapat diberikan oleh seorang *user* terkait dengan bencana. Pada Tabel 6 terlihat bahwa *user* yang memiliki indeks *influence score* pada topik bencana dan tsunami di Indonesia adalah infoBMKG, diikuti dengan *user* media seperti detikcom, kompascom, vivacoid, dan cnnindonesia. Semakin besar nilai indeks *influence score*-nya maka makin berpengaruh *user* tersebut pada topik tertentu.

TABEL VI
SAMPEL INFLUENCE SCORE YANG DIDAPAT DARI INDEKS KOMPOSIT

No	User	Indeks Influence Score
1	infoBMKG	0,63698
2	detikcom	0,54118
3	kompascom	0,18755
4	VIVAcoid	0,15767
5	CNNIndonesia	0,10387

Untuk melihat seberapa besar korelasi dari nilai *influence score* yang dihasilkan dari indeks komposit dengan metode pada [15] dilakukan uji korelasi *rank spearman* untuk membandingkan *ranking* yang dihasilkan pada masing masing metode. Dari perhitungan didapatkan besarnya korelasi *rank spearman* antara kedua indeks tersebut adalah sebesar 0,207 (jika nilai korelasi antara 0 – 0,25 maka termasuk hubungan yang lemah). Dilanjutkan melakukan uji signifikansi dan menghasilkan keputusan tolak H₀. Sehingga kesimpulan yang dapat diambil adalah terdapat cukup bukti untuk menyatakan bahwa kedua

indeks tersebut memiliki hubungan walaupun korelasi yang terbentuk termasuk korelasi kecil/lemah.



Gambar. 4 Influence Score Akun Twitter

Gambar 4 menunjukkan *influence score* akun twitter tanpa dan dengan mempertimbangkan faktor *engagement rate (engagement)* dan faktor *frequency and network analysis (frequency)*. Terlihat bahwa tanpa mempertimbangkan kedua faktor tersebut maka akun detikcom menjadi akun dengan *influence score* tertinggi sebesar 0,77. Sedangkan jika dengan mempertimbangan kedua faktor tersebut ke dalam analisis, maka akun infoBMKG menjadi akun yang memiliki *influence score* tertinggi sebesar 0,64.

IV. KESIMPULAN

Hasil *Social Network Analysis* tanpa mempertimbangkan faktor *frequency* dan *engagement* menunjukkan bahwa pengguna Twitter yang dinilai paling berpengaruh pada topik bencana gempa dan tsunami adalah situs berita seperti detikcom, sedangkan jika mempertimbangkan faktor *frequency* dan *engagement* menunjukkan bahwa pengguna Twitter yang dinilai paling berpengaruh pada topik bencana gempa dan tsunami adalah akun infoBMKG. Untuk penelitian selanjutnya dapat dianalisis mengenai akun Twitter berpengaruh yang berfokus pada penyaluran bantuan ketika terjadi bencana alam dan konektivitasnya dengan akun penyedia informasi bencana.

UCAPAN TERIMA KASIH / ACKNOWLEDGMENT

Terima kasih kepada PPPM Politeknik Statistika STIS yang telah membiayai penelitian ini. Penelitian ini adalah bagian dari Penelitian Tim Riset Kebencanaan Politeknik Statistika STIS yang diketuai oleh Ibu Dr. Tiodora Hadumaon Siagian M. Pop. Hum. Res.

REFERENSI

[1] D. N. Sattler, M. Claramita, and B. Muskavage, "Natural Disasters in Indonesia: Relationships Among Posttraumatic Stress, Resource Loss, Depression, Social Support, and Posttraumatic Growth," *J Loss Trauma*, vol. 23, no. 5, pp. 351–365, Jul. 2018, doi: 10.1080/15325024.2017.1415740.

- [2] S. Behl, A. Rao, S. Aggarwal, S. Chadha, and H. S. Pannu, "Twitter for disaster relief through sentiment analysis for COVID-19 and natural hazard crises," *International Journal of Disaster Risk Reduction*, vol. 55, Mar. 2021, doi: 10.1016/j.ijdr.2021.102101.
- [3] P. Kumar, "Twitter, disasters and cultural heritage: A case study of the 2015 Nepal earthquake," *Journal of Contingencies and Crisis Management*, vol. 28, no. 4, pp. 453–465, Dec. 2020, doi: 10.1111/1468-5973.12333.
- [4] A. Karami, V. Shah, R. Vaezi, and A. Bansal, "Twitter Speaks: A Case of National Disaster Situational Awareness," Mar. 2019, [Online]. Available: <http://arxiv.org/abs/1903.02706>
- [5] C. M. Vera-Burgos and D. R. Griffin Padgett, "Using Twitter for crisis communications in a natural disaster: Hurricane Harvey," *Heliyon*, vol. 6, no. 9, Sep. 2020, doi: 10.1016/j.heliyon.2020.e04804.
- [6] G. P. Cooper, V. Yeager, F. M. Burkle, and I. Subbarao, "Twitter as a potential disaster risk reduction tool. part i: Introduction, terminology, research and operational applications," *PLoS Curr*, vol. 7, no. DISASTERS, Jun. 2015, doi: 10.1371/currents.dis.a7657429d6f25f02bb5253e551015f0f.
- [7] K. Wang, N. S. N. Lam, L. Zou, and V. Mihunov, "Twitter use in hurricane isaac and its implications for disaster resilience," *ISPRS Int J Geoinf*, vol. 10, no. 3, Mar. 2021, doi: 10.3390/ijgi10030116.
- [8] K. M. Carley, M. Malik, P. M. Landwehr, J. Pfeffer, and M. Kowalchuk, "Crowd Sourcing Disaster Management: The Complex Nature of Twitter Usage in Indonesia," 2019. [Online]. Available: <https://www.elsevier.com/open-access/userlicense/1.0/>
- [9] U. Widyanarko and D. R. Hizbaron, "Urban response towards tropical cyclone using twitter in Indonesia," in *IOP Conference Series: Earth and Environmental Science*, Mar. 2020, vol. 451, no. 1. doi: 10.1088/1755-1315/451/1/012098.
- [10] F. Morone and H. A. Makse, "Influence maximization in complex networks through optimal percolation," Jun. 2015, doi: 10.1038/nature14604.
- [11] R. Herrera, J. P. Cárdenas, and R. Alfaro, "User Influence in Online Social Networks," 2014.
- [12] F. Riquelme and P. González-Cantergiani, "Measuring user influence on Twitter: A survey," *Inf Process Manag*, vol. 52, no. 5, pp. 949–975, Sep. 2016, doi: 10.1016/j.ipm.2016.04.003.
- [13] M. Kardara, G. Papadakis, A. Papaioikonomou, K. Tserpes, and T. Varvarigou, "Large-scale evaluation framework for local influence theories in Twitter," *Inf Process Manag*, vol. 51, no. 1, pp. 226–252, 2015, doi: 10.1016/j.ipm.2014.06.002.
- [14] "https://pypi.org/project/twint/."
- [15] C. Bigonha, T. N. C. Cardoso, M. M. Moro, M. A. Gonçalves, and V. A. F. Almeida, "Sentiment-based influence detection on Twitter," *Journal of the Brazilian Computer Society*, vol. 18, no. 3, pp. 169–183, Sep. 2012, doi: 10.1007/s13173-011-0051-5.
- [16] Sandy Ressler, *Perspectives on electronic publishing: standards, solutions, and more*. Prentice-Hall, Inc., 1993.
- [17] F. D. Ananda and Y. Pristyanto, "Analisis Sentimen Pengguna Twitter Terhadap Layanan Internet Provider Menggunakan Algoritma Support Vector Machine," *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 20, no. 2, pp. 407–416, 2021.
- [18] A. Faesal, A. Muslim, A. H. Ruger, and K. Kusriani, "Sentimen analisis pada data tweet pengguna twitter terhadap produk penjualan toko online menggunakan metode k-means," *MATRIK: Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 19, no. 2, pp. 207–213, 2020.
- [19] European Commission. Joint Research Centre. and Organisation for Economic Co-operation and Development., *Handbook on constructing composite indicators: methodology and user guide*. OECD, 2008.