

ANALISIS SENTIMEN TOPIK VIRAL DESA PENARI PADA MEDIA SOSIAL TWITTER DENGAN METODE *LEXICON BASED*

Rifiana Arief¹, Karel Imanuel²
Universitas Gunadarma^{1,2}

Jalan Margonda Raya No 100 Depok Jawa Barat 16424
Sur-el : rifiana@staff.gunadarma.ac.id¹, karel4404@gmail.com²

Abstract : *The horror story of Dancer Village in Indonesia is a viral topic that has become a talk of citizens on Twitter social media. Various responses and public opinions emerged related to the truth of the story of supernatural experiences of students during a Real Work Lecture in an East Java region of Indonesia. This study conducted a sentiment analysis of community comments on Twitter social media on the viral topic using the Lexicon Based method. Sentiment classification is divided into 3 classes namely positive, negative and neutral. The research phase consists of data collection, pre-processing, processing (sentiment analysis) and visualization. Data collection uses Twitter Search API with 1000 Penari Desa keywords in Indonesian. The lexicon assessment results from 1000 tweets data obtained 33 positive, 767 neutral and 200 negative. The percentage of tweets containing positive comments by 3.3%, neutral 76.7% and negative by 20%.*

Keywords: *Dancer Village, Sentiment Analysis, Lexicon Based, Twitter, WorldCloud*

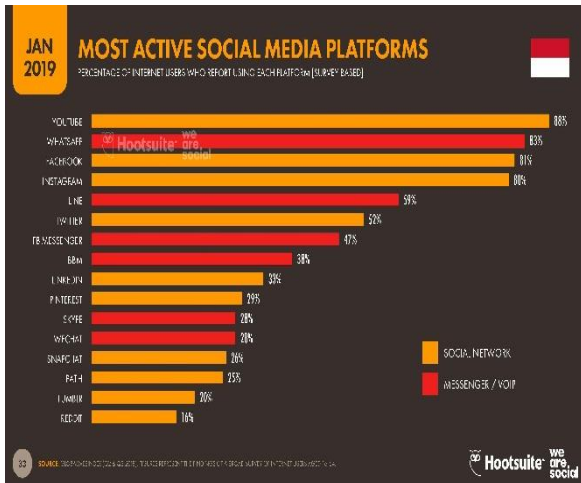
Abstrak : *Kisah horor Desa Penari di Indonesia merupakan topik viral yang menjadi perbincangan warganet pada media sosial twitter. Berbagai tanggapan dan opini masyarakat muncul terkait kebenaran cerita pengalaman supranatural mahasiswa saat Kuliah Kerja Nyata di sebuah wilayah Jawa Timur Indonesia. Penelitian ini melakukan analisis sentimen dari komentar-komentar masyarakat pada media sosial Twitter terhadap topik viral tersebut menggunakan metode Lexicon Based. Klasifikasi sentimen dibagi menjadi 3 kelas yaitu positif, negatif dan netral. Tahap penelitian terdiri dari pengumpulan data, prapengolahan, pengolahan (analisis sentimen) dan visualisasi. Pengumpulan data menggunakan API Search Twitter dengan kata kunci Desa Penari sebanyak 1000 buah komentar (tweet) dalam bahasa Indonesia. Hasil penilaian leksikon dari 1000 data tweet diperoleh 33 tweet bernilai positif, 767 tweet bernilai netral dan 200 tweet negatif. Prosentase tweet berisi komentar positif sebesar 3.3 %, netral 76.7 % dan negatif sebesar 20%.*

Kata Kunci : *Desa Penari, Analisis Sentimen, Lexicon Based, Twitter, WorldCloud*

1. PENDAHULUAN

Perkembangan teknologi informasi dan internet memberi alternatif baru dalam berkomunikasi. Media sosial sebagai kumpulan perangkat lunak memungkinkan individu maupun komunitas untuk berkumpul, berbagi, berkomunikasi, berkolaborasi atau bermain. Ketersediaan berbagai layanan melalui media sosial yang dapat digunakan telah merubah cara berkomunikasi dalam masyarakat [1].

Hoot Suite We Are Social melakukan survei pada Januari 2019 menyatakan bahwa total pengguna aktif media sosial melalui internet di Indonesia mencapai 150 juta orang, artinya mencapai 56% dari total penduduk Indonesia. Penggunaan aplikasi sosial media di Indonesia [2] antara lain *Youtube, Whatsapp, Facebook, Instagram, Line, Twitter, FB Messenger, BBM, Linkedin, Pinterest, Skype, Wechat, Snapchat, Path, Tumblr* dan *Reddit* diilustrasikan pada Gambar 1.



Gambar 1. Survey Penggunaan Sosial Media Di Indonesia Awal Tahun 2019 [2]

Masyarakat Indonesia menyukai kisah horor dan mistis baik dalam bentuk film, benda-benda, maupun kisah nyata. Perusahaan media intelligence Isentia menangkap perbincangan terkait keyword 'horor' dan 'mistis' di media sosial terhitung sejak 1 Agustus hingga 3 Oktober 2019 sebanyak 27.399 buzz. Warganet Indonesia memang memiliki kecenderungan tertarik terhadap kisah berbau horor, terbukti dari banyaknya topik viral terkait hal mistis terutama beberapa waktu belakangan ini. Terdapat empat topik utama yang paling viral berdasarkan urutan popularitasnya pada media sosial yaitu Boneka, KKN di Desa Penari, Orderan Mistis Ojol dan Film Danur [3] diilustrasikan pada Gambar 2.



Gambar 2: Trend Topik Cerita Horor [3]

Analisis sentimen atau penambangan opini dapat didefinisikan sebagai tugas mendeteksi, mengekstraksi, dan mengklasifikasikan pendapat tentang sesuatu hal. Tugas ini merupakan jenis pemrosesan bahasa alami (NLP) untuk melacak mood publik untuk hukum, kebijakan, atau pemasaran tertentu, dll. Analisis sentimen melibatkan cara pengembangan untuk pengumpulan dan pemeriksaan komentar dan pendapat tentang undang-undang, kebijakan, dll., yang diposting di media sosial. Selain itu analisis sentimen berurusan dengan penambangan informasi terkait sentimen atau pendapat dari grup untuk topik tertentu. Sentimen analisis juga cukup populer pada berbagai bidang seperti politik yaitu untuk memperkirakan hasil pemilihan dari berbagai forum politik, bisnis yaitu untuk menganalisis sentimen online di media sosial untuk prediksi pasar saham dan pemasaran yaitu untuk memperkirakan penjualan produk spesifik [4].

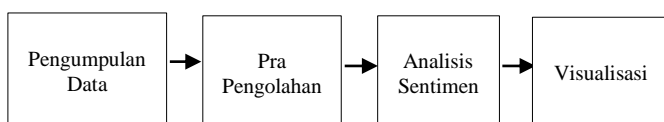
Analisis sentimen adalah sebuah proses untuk menentukan sentimen atau opini dari seseorang yang diwujudkan dalam bentuk teks dan bisa dikategorikan sebagai sentimen positif, negatif atau netral. Pengguna internet banyak menuliskan pengalaman, opini dan segala hal yang menjadi perhatian mereka. Tulisan tentang apa yang mereka rasakan ini berupa perasaan positif, netral maupun negatif yang bisa diungkapkan dengan cara yang cukup kompleks [5] dalam [6]. Ide dasar dari investigasi sentimen adalah untuk mendeteksi polaritas teks dokumen atau kalimat pendek dan mengklasifikasikannya dalam premis ini. Polaritas sentimen dikategorikan sebagai

"positif", "Negatif" atau "tidak memihak" (netral). Teknik yang ada untuk analisis sentimen termasuk pembelajaran mesin (diawasi dan tidak diawasi), dan pendekatan berbasis leksikal [7]. Berbagai penelitian terkait analisis sentimen berbagai topik di Indonesia telah banyak dilakukan antara lain dengan metode *Naïve Bayes* [8], *Support Vector Machine* [9], *DBF* [10].

Penelitian ini akan membahas tentang analisis sentimen dari komentar masyarakat pada media sosial *Twitter* terhadap topik viral desa penari pada media sosial *twitter* dengan metode *lexicon based*. Tujuan penelitian ini adalah untuk mengelompokkan komentar masyarakat ke kategori sentimen positif, negatif dan netral berdasarkan *lexicon*.

2. METODOLOGI PENELITIAN

Pada penelitian ini dilakukan analisis sentimen kisah horor desa penari dengan metode *Lexicon Based*. Sentimen komentar masyarakat akan dikategorikan menjadi positif, negatif dan netral. Tahap penelitian terdiri dari pengumpulan data, pra pengolahan, pengolahan (analisis sentimen) dan visualisasi seperti ditunjukkan pada Gambar 4.



Gambar 4. Tahap Penelitian

2.1 Pengumpulan Data

Data yang digunakan adalah komentar-komentar masyarakat dari *Twitter* dengan kata

kunci Desa Penari dan komentar dalam bahasa Indonesia sebanyak 1000 buah pada tanggal 13 November 2019. Proses diawali dengan membuat aplikasi *Twitter*, mengumpulkan data *tweets* dengan menggunakan akses pada *Twitter Search API*, mengubah kumpulan data (*list tweets*) menjadi *data frame* untuk membedakan konten dari atribut-atribut data yang diperoleh dalam bentuk tabel dan untuk proses selanjutnya hanya mengambil dan menggunakan isi konten dari atribut teks.

2.2 Pra Pengolahan

Pra pengolahan bertujuan untuk pengeksktraksian dan pembersihan *tweet* yang akan digunakan pada tahap pengolahan (analisis sentimen /pengklasifikasian *tweet*). Proses terdiri antara lain: mengubah variabel teks menjadi karakter, mengubah huruf besar menjadi huruf kecil (*case folding*), menghapus tautan, menghapus *mention* (@), menghapus tanda baca, menghapus angka, menghapus *stopword* (kata "rt" dan "gt") dan menghapus *whitespace* (spasi yang berlebihan).

2.3 Pengolahan (Analisis Sentimen)

Setelah selesai melakukan tahap pra-pengolahan, dilanjutkan ke tahap pengolahan yaitu analisis sentimen untuk pengklasifikasian *tweet*. Metode *Lexical based approach* menggunakan sebuah kamus sebagai sumber bahasa atau leksikal. Dalam penelitian ini adalah *opinion lexicon* milik Liu [11] yang telah dimodifikasi dan ditranslate menjadi bahasa Indonesia oleh Devid Haryalesmana [12].

Tahap pengolahan terdiri dari 3 proses, yaitu mengimport kamus kata leksikon positif dan kamus kata leksikon negatif, memecah teks tweets menjadi *bag of words* (penggalan kata) yang bertujuan memudahkan pada saat penentuan skor positif, negatif dan netral serta menghitung skor sentimen positif, negatif dan netral dengan menggunakan metode Lexicon-based.

Skor *dictionary based / lexicon based* dihitung menggunakan rumus (1), (2) dan (3) :

$$\text{if } \sum_k \text{Score}(k) > 0 \text{ then positive} \quad \dots\dots (1)$$

$$\text{if } \sum_k \text{score}(k) < 0 \text{ then negative} \quad \dots\dots (2)$$

$$\text{if } \sum_k \text{score}(k) = 0 \text{ then Neutral} \quad \dots\dots (3)$$

2.4 Visualisasi

Tahap akhir memvisualisasi hasil analisis sentimen dalam bentuk diagram histogram maupun wordcloud. Dengan menggunakan word cloud, gambaran frekuensi kata-kata dapat ditampilkan dalam bentuk yang menarik namun tetap informatif. Semakin sering satu kata digunakan, maka semakin besar pula ukuran kata tersebut ditampilkan dalam *word cloud*.

3. HASIL DAN PEMBAHASAN

3.1. Hasil Pengumpulan Data

Pengambilan data dengan *API Search Twitter* pada tanggal 13 November 2019 dengan kata kunci Desa Penari telah mengumpulkan sebanyak 1000 buah komentar (*tweet*) dalam

bahasa Indonesia. Hasil pengambilan data yang diperoleh secara default berupa *tweets* berbentuk *list* ditunjukkan pada Gambar 5. Selanjutnya list tersebut perlu diubah menjadi data frame agar dapat diproses pada tahap analisis selanjutnya. *Data frame* berisi 16 variabel (atau bisa disebut sebagai *field*). Variabel-variabel tersebut adalah: 1). *text*, berisi teks *tweets*. 2). *favorited*, berisi nilai Boolean mengenai apakah *tweets* difavoritekan pengguna/pengguna lain. 3). *favoriteCount*, berisi jumlah kali *tweets* difavoritekan. 4). *replyToSN*, berisi nama akun ke mana *tweets* dibalaskan. 5). *created*, berisi tanggal dan jam *tweets* dibuat. 6). *truncated*, berisi nilai Boolean apakah *tweets* terpotong. 7). *replyToSID*, berisi ID *tweets* balasan. 8). *id*, berisi ID *tweets*. 9). *replyToUID*, berisi ID pengguna ke mana *tweets* dibalaskan. 10). *statusSource*, berisi sumber *user agent* dari *tweets*. 11). *screenName*, berisi nama akun pembuat *tweets*. 12). *retweetCount*, berisi jumlah kali *tweets* di-*retweet*. 13). *isRetweet*, berisi nilai Boolean apakah suatu *tweets* merupakan sebuah *retweet*. 14). *retweeted*, berisi nilai Boolean apakah suatu *tweets* di-*retweet*. 15). *longitude*, berisi ordinat garis bujur tempat *tweets* dibuat. 16). *latitude*, berisi ordinat garis lintang tempat *tweets* dibuat. Hasil perubahan *list data tweet* yang disimpan menjadi data frame ditunjukkan pada Gambar 6. Kemudian melakukan pengambilan hanya konten teksnya saja yaitu pada atribut *text*. Hasil pengumpulan hanya konten teks ditunjukkan pada Gambar 7.

```

analysis
├── [[1]]
│   ├── created
│   │   ├── type: double (S3: POSIXct, POSIXt)
│   │   ├── value: 2019-11-16 02:44:37
│   ├── favoriteCount
│   │   ├── type: double [1]
│   │   ├── value: 0
│   ├── favorited
│   │   ├── type: logical
│   │   ├── value: FALSE
│   ├── id
│   │   ├── type: character [1]
│   │   ├── value: '1195533100337381376'
│   ├── isRetweet
│   │   ├── type: logical
│   │   ├── value: TRUE
│   ├── latitude
│   │   ├── type: character [0]
│   │   ├── value: 
│   ├── longitude
│   │   ├── type: character [0]
│   │   ├── value: 
│   ├── replyToSID
│   │   ├── type: character [0]
│   │   ├── value: 
│   ├── replyToSN
│   │   ├── type: character [0]
│   │   ├── value: 
│   ├── replyToUID
│   │   ├── type: character [0]
│   │   ├── value: 
│   ├── retweetCount
│   │   ├── type: double [1]
│   │   ├── value: 3550
│   ├── retweeted
│   │   ├── type: logical
│   │   ├── value: FALSE
│   ├── screenName
│   │   ├── type: character [1]
│   │   ├── value: 'FebrianaYH'
│   ├── statusSource
│   │   ├── type: character [1]
│   │   ├── value: '<a href="http://twitter.com/download/iphone" rel="nofollow">Twitter for iPhone< ...
│   ├── text
│   │   ├── type: character [1]
│   │   ├── value: 'RT @Evan_sadis: Ini kisah KKN Desa Penari akan cek di youtube Sisiv\&n\#flayan ...
│   ├── truncated
│   │   ├── type: logical
│   │   ├── value: FALSE
│   ├── urls
│   │   ├── type: list [0 x 4] (S3: data.frame)
│   │   ├── value: A data.frame with 0 rows and 4 columns
│   ├── refClassDef
│   │   ├── type: S4 (methods::refClassRepresentation)
│   │   ├── value: S4 object of class refClassRepresentation
│   ├── .self
│   │   ├── type: S4 [1] (twitterR::status)
│   │   ├── value: 
│   ├── created
│   │   ├── type: double (S3: POSIXct, POSIXt)
│   │   ├── value: 2019-11-16 02:44:37
└── analysis[[1]]
    
```

Gambar 5. Hasil Pengambilan Data Dari API Search Tweeter

text	favorited	favoriteCount	replyToSN	created	truncated	replyToSID	id	replyTo
1 RT @Evan_sadis: Ini kisah KKN Desa Penari akan cek di y...	FALSE	0	NA	2019-11-16 02:44:37	FALSE	NA	1195533100337381376	NA
2 RT @levannaayu: Postingan ini hilang , sama kayak thre...	FALSE	0	NA	2019-11-16 02:42:31	FALSE	NA	1195532569632108544	NA
3 KKN desa penari <U+0001F9DA> <U+200D> <U+2640> >...	FALSE	0	NA	2019-11-16 02:36:07	FALSE	NA	1195530962349838336	NA
4 https://t.co/dVQjqI9Ftt Gunung Lawu	FALSE	0	NA	2019-11-16 02:27:31	FALSE	NA	1195528797405081601	NA
5 @iasssssssss @vulpeculai Ini ayamnya macam bima kete...	FALSE	0	iasssssssss	2019-11-16 02:00:54	FALSE	119523766083656801	1195522099458134016	1047
6 RT @nurulthfh: MET PAGI CANTIQUÉ @malikatazkia YAN...	FALSE	0	NA	2019-11-16 01:18:08	FALSE	NA	1195511335926161411	NA
7 MET PAGI CANTIQUÉ @malikatazkia YANG KUAT DAN MEN...	FALSE	1	NA	2019-11-16 01:16:20	FALSE	NA	1195510882366672901	NA
8 RT @fianahsrnd: asyikkk bener deh jogetnya <U+0001F...	FALSE	0	NA	2019-11-16 01:10:45	FALSE	NA	1195509477362896897	NA
9 @SimpleM81378523 @auraaapi @bacahorror Kkn desa ...	FALSE	0	SimpleM81378523	2019-11-16 01:10:28	FALSE	1180793234793959425	1195509405065740288	1095
10 @juhyuqn kerasukan desa penari	FALSE	0	juhyuqn	2019-11-16 00:50:26	FALSE	1195502572905189376	1195504363612987392	1814
11 RT @anakOrang: @Askmenfess Jan: 80jt Feb: unboxing s...	FALSE	0	NA	2019-11-16 00:43:10	FALSE	NA	1195502533692669953	NA
12 RT @Juyeonescent: Soonyoung sepulangny dari KKN di...	FALSE	0	NA	2019-11-16 00:31:48	FALSE	NA	1195499674733436928	NA
13 https://t.co/dVQjqI9Ftt Kokos Jiang: Kokos Leo Lim	FALSE	0	NA	2019-11-15 20:28:01	FALSE	NA	1195438325739294720	NA
14 @kalonggg Pless deh ea, kamutu jan riyak taw gak, me...	FALSE	0	kalonggg	2019-11-15 20:19:53	TRUE	1195434802280747008	1195436276645687296	1123
15 RT @korantempo: Wisata kuliner kelapa kopyor ini seda...	FALSE	0	NA	2019-11-15 20:07:00	FALSE	NA	1195433037829492737	NA
16 RT @korantempo: Nama Desa Penari tiba-tiba menjadi p...	FALSE	0	NA	2019-11-15 19:22:00	FALSE	NA	1195421712550187008	NA
17 @kdryvk Dia ini sto yg kkn desa penari :(FALSE	0	kdryvk	2019-11-15 18:55:20	FALSE	1195414882742489088	1195415000770203648	1125
18 @Askmenfess Gambaran widya lagi nari di kisah kkn de...	FALSE	0	Askmenfess	2019-11-15 18:51:58	FALSE	1195413892547039232	1195414153831206912	8104
19 @Askmenfess penari desa kkn	FALSE	0	Askmenfess	2019-11-15 18:51:44	FALSE	1195413892547039232	1195414092443344896	8104
20 RT @debrisumule: Anna juga gamau kalah donggg.. nar...	FALSE	0	NA	2019-11-15 18:17:49	FALSE	NA	1195405557517041664	NA

Gambar 6. Hasil List Tweet Yang Telah Diubah Menjadi Data Frame

1	RT @Evan_sadis: Ini kisah KKN Desa Penari akan cek di y...
2	RT @levannaayu: Postingan ini hilang , sama kayak thre...
3	KKN desa penari <U+0001F9DA> <U+200D> <U+2640> >...
4	https://t.co/dVQjqI9Ftt Gunung Lawu
5	@iasssssssss @vulpeculai Ini ayamnya macam bima kete...
6	RT @nurulthfh: MET PAGI CANTIQUÉ @malikatazkia YAN...
7	MET PAGI CANTIQUÉ @malikatazkia YANG KUAT DAN MEN...
8	RT @fianahsrnd: asyikkk bener deh jogetnya <U+0001F...
9	@SimpleM81378523 @auraaapi @bacahorror Kkn desa ...
10	@juhyuqn kerasukan desa penari
11	RT @anakOrang: @Askmenfess Jan: 80jt Feb: unboxing s...
12	RT @Juyeonescent: Soonyoung sepulangny dari KKN di...
13	https://t.co/dVQjqI9Ftt Kokos Jiang: Kokos Leo Lim
14	@kalonggg Pless deh ea, kamutu jan riyak taw gak, me...
15	RT @korantempo: Wisata kuliner kelapa kopyor ini seda...
16	RT @korantempo: Nama Desa Penari tiba-tiba menjadi p...
17	@kdryvk Dia ini sto yg kkn desa penari :(
18	@Askmenfess Gambaran widya lagi nari di kisah kkn de...
19	@Askmenfess penari desa kkn
20	RT @debrisumule: Anna juga gamau kalah donggg.. nar...
21	RT @debrisumule: Anna juga gamau kalah donggg.. nar...
22	@debrisumule: Anna juga gamau kalah donggg.. nar...

Showing 1 to 22 of 1,000 entries

Gambar 7. Hasil Pengambilan Konten Teks Dari Atribut Text Pada Data Frame

3.2. Hasil Pra Pengolahan

Pra pengolahan dilakukan terhadap konten teks dari atribut teks pada data frame yaitu mengubah variabel teks menjadi karakter, mengubah huruf besar menjadi huruf kecil (*case folding*), menghapus tautan, menghapus *mention* (@), menghapus tanda baca, menghapus angka, menghapus *stopword* (kata “rt” dan “gt”) dan menghapus *whitespace* (spasi yang berlebihan). Perubahan isi konten teks melalui proses pra pengolahan secara berkelanjutan ditunjukkan pada Tabel 1 sampai Tabel 4. Hasil akhir setelah prapengolahan ditunjukkan pada Gambar 8.

Tabel 1. Proses case folding

No	Sebelum	Sesudah
1	RT @Evan_sadis: Ini rt kisah KKN Desa Penari akan cek di youtube akan cek di youtube Sisitiv. #layangputus Inalilahiwainailaihirojiun #kisahanakkampung un https://...	@evan_sadis: ini kisah kkn desa penari akan cek di youtube akan cek di youtube sisitiv. #layangputus inilalahiwainailaihiroji n #kisahanakkampung un https://...
2	RT @levannaayu: rt Postingan ini hilang, sama kayak thread kkn desa penari waktu itu. #bacahoror @bacahorror https://t.co/LyladcDBKy	@levannaayu: rt Postingan ini hilang, sama kayak thread kkn desa penari waktu itu. #bacahoror @bacahorror https://t.co/lyladcdbky
3	@SimpleM81378523 @auraapi @bacahorror Kkn desa penari saja ku belum baca :(@simplem81378523 @auraapi @bacahorror kkn desa penari saja ku belum baca :(
4	RT @korantempo: Nama Desa Penari tiba-tiba menjadi perbincangan masyarakat Indonesia. Bermula dari cerita horor berjudul KKN Desa Pen...	@korantempo: nama desa penari tiba-tiba menjadi perbincangan ramai masyarakat indonesia. bermula dari cerita horor berjudul kkn desa pen...

Tabel 2. Proses Penghapusan Tautan

No	Sebelum	Sesudah
1	rt @evan_sadis: ini kisah kkn desa penari akan cek di youtube akan cek di youtube #layangputus inalilahiwainailaihirojiun #kisahanakkampung https://...	@evan_sadis: ini kisah kkn desa penari akan cek di youtube akan cek di youtube sisitiv. #layangputus inilalahiwainailaihir ojiun #kisahanakkampung https://...
2	rt @levannaayu: rt Postingan ini hilang, sama kayak thread kkn desa penari waktu itu. #bacahoror @bacahorror https://t.co/lyladcdbky	@levannaayu: rt Postingan ini hilang, sama kayak thread kkn desa penari waktu itu. #bacahorror @bacahorror https://t.co/lyladcdbky
3	@simplem81378523 @auraapi @bacahorror kkn desa penari saja ku belum baca :(@simplem81378523 @auraapi @bacahorror kkn desa penari saja ku belum baca :(
4	rt @korantempo: nama desa penari menjadi perbincangan indonesia. bermula dari cerita horor berjudul kkn desa pen...	@korantempo: nama desa penari menjadi perbincangan ramai masyarakat indonesia. bermula dari cerita horor berjudul kkn desa pen...

Tabel 3. Proses Penghapusan Mention

No	Sebelum	Sesudah
1	rt @evan_sadis: ini kisah kkn desa kkn desa penari akan cek di youtube akan cek di youtube sisitiv. #layangputus inalilahiwainailaihiroji un #kisahanakkampung n	@evan_sadis: ini kisah kkn desa kkn desa penari akan cek di youtube akan cek di youtube sisitiv. #layangputus inalilahiwainailaihiroji un #kisahanakkampung n
2	rt @levannaayu: rt: postingan ini hilang, sama kayak thread kkn desa penari waktu itu. #bacahoror @bacahorror	@levannaayu: rt: postingan ini hilang, sama kayak thread kkn desa penari waktu itu. #bacahorror @bacahorror
3	@simplem81378523 @auraapi @bacahorror kkn desa penari saja ku belum baca :(@simplem81378523 @auraapi @bacahorror kkn desa penari saja ku belum baca :(
4	rt @korantempo: nama desa penari tiba-tiba menjadi perbincangan ramai masyarakat indonesia. bermula dari cerita horor berjudul kkn desa pen...	@korantempo: nama desa penari tiba-tiba menjadi perbincangan ramai masyarakat indonesia. bermula dari cerita horor berjudul kkn desa pen...

Proses memecah teks *tweets* menjadi *bags of words* agar *tweets* menjadi penggalan kata ditunjukkan pada Gambar 11.

Name	Type	Value
bo_teks_analisis	list [1000]	List of length 1000
[[1]]	character [13]	'rt' kisah 'kkn' 'desa' 'penari' 'cek' ...
[[2]]	character [10]	'rt' 'postingan' 'hilang' 'kayak' 'thread' 'kkn' ...
[[3]]	character [10]	'kkn' 'desa' 'penari' '\U0001f9da2\U0001f9da2' 'thread' 'kkn' ...
[[4]]	character [3]	"" 'gunung' 'lawu'
[[5]]	character [16]	"" 'ayamnya' 'bima' 'ketemu' 'badarawuhi' 'desa' ...
[[6]]	character [20]	'rt' 'met' 'pagi' 'cintique' 'kuat' 'mengalir' ...
[[7]]	character [17]	'met' 'pagi' 'cintique' 'kuat' 'mengalir' 'jemig' ...
[[8]]	character [18]	'rt' 'asyikk' 'bener' 'deh' 'jogetnya' '\U0001f602\U0001f923' ...
[[9]]	character [7]	"" 'kkn' 'desa' 'penari' 'ku' 'baca' ...
[[10]]	character [4]	"" 'kerasukan' 'desa' 'penari'
[[11]]	character [17]	'rt' 'jan' 'j' 'feb' 'unboxing' 'skupi' ...
[[12]]	character [11]	'rt' 'soonyoung' 'sepulangny' 'kkn' 'desa' 'yessoda' ...
[[13]]	character [6]	"" 'kokos' 'jiang' 'kokos' 'leo' 'lim'
[[14]]	character [20]	"" 'ples' 'deh' 'ea' 'kamutu' 'jan' ...
[[15]]	character [13]	'rt' 'wisata' 'kuliner' 'kelapa' 'kopyor' 'heboh' ...
[[16]]	character [14]	'rt' 'nama' 'desa' 'penari' 'perbincangan' 'rama' ...
[[17]]	character [7]	"" 'sto' 'yg' 'kkn' 'desa' 'penari' ...
[[18]]	character [8]	"" 'gambaran' 'widy' 'nari' 'kisah' 'kkn' ...
[[19]]	character [4]	"" 'penari' 'desa' 'kkn'
[[20]]	character [15]	'rt' 'anna' 'namai' 'kalah' 'dennoni' 'nari' ...

Gambar 11. Hasil Pemecahan Teks Menjadi Bag of Word (Penggalan Kata)

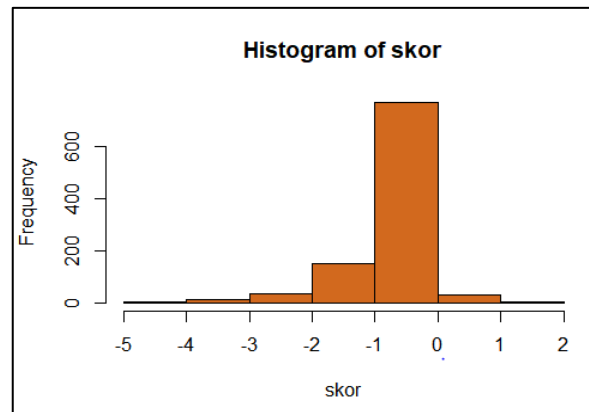
Tahap selanjutnya penghitungan skor untuk menggolongkan *tweets* positif, negatif dan netral. *Tweet* yang memiliki jumlah skor positif lebih besar dari skor negatif maka digolongkan pada sentimen positif, begitu sebaliknya. Jika *tweet* tersebut memiliki skor positif dan negatif yang sama, maka digolongkan dalam sentimen netral. Hasil analisis ditunjukkan pada Gambar 12.

teks_analisis_22	list [1002]	List of length 1002
[[1]]	integer [1]	0
[[2]]	integer [1]	0
[[3]]	integer [1]	0
[[4]]	integer [1]	0
[[5]]	integer [1]	-1
[[6]]	integer [1]	-2
[[7]]	integer [1]	-2
[[8]]	integer [1]	-1
[[9]]	integer [1]	-1
[[10]]	integer [1]	0
[[11]]	integer [1]	-1
[[12]]	integer [1]	0
[[13]]	integer [1]	0
[[14]]	integer [1]	-1
[[15]]	integer [1]	0
[[16]]	integer [1]	-2
[[17]]	integer [1]	0
[[18]]	integer [1]	0
[[19]]	integer [1]	0
[[20]]	integer [1]	0

Gambar 12. Hasil Skor Analisis Sentimen Dengan Lexicon Based

3.4. Hasil Visualisasi

Tahap terakhir dalam penelitian ini adalah memvisualisasi hasil klasifikasi algoritma *Lexicon Based* untuk proporsi bagian – bagian tweet yaitu positif, negative, dan netral. Hasil Visualisasi Dalam Histogram disajikan pada Gambar 13 dan dan WordCloud pada Gambar 14.



Gambar 13. Hasil Visualisasi Dalam Histogram

Visualisasi dalam *wordcloud* yang menunjukkan kata-kata yang sering muncul dalam keseluruhan *tweet* Desa Penari ditunjukkan pada Gambar 14.



Gambar 14. Visualisasi Dalam WordCloud

Tabel 5. Penilaian Dengan Lexicon Based

Kategori	Jumlah	Prosentase
Positif	33	3.3 %
Netral	767	76.7 %
Negatif	200	20%

Tabel 5 menunjukkan penilaian leksikon dari 1000 data *tweet* diperoleh 33 *tweet* bernilai positif, 767 *tweet* bernilai netral dan 200 *tweet* negatif. Prosentase *tweet* berisi komentar positif sebesar 3.3 %, netral 76.7 % dan negatif sebesar 20%.

4. KESIMPULAN

Paper ini melakukan analisis sentimen topik viral desa penari pada media sosial *twitter* dengan metode *lexicon based*. Berdasar 1000 data *tweet* yang dianalisis diperoleh hasil 33 *tweet* positif (3.3%), 767 *tweet* netral (76.7%) dan 200 *tweet* negatif (20%). Dapat disimpulkan analisis sentimen komentar masyarakat terhadap cerita ini kebanyakan adalah bersifat netral.

DAFTAR PUSTAKA

- [1] A. Clara Sari, R. Hartina, R. Awalia, H. Irianti, N. Ainun, "Komunikasi dan Media Sosial, Research Gate, 2019.
- [2] Websindo, "Indonesia Digital 2019: Media Sosial", *Websindo*. [Online]. Available: <https://websindo.com/indonesia-digital-2019-media-sosial/> [Accessed: 17-Oct-2019).
- [3] DetikInet, "Kisah Horor yang Viral di Medsos selain KKN di Desa Penari", *DetikInet*. [Online]. Available: <https://inet.detik.com/cyberlife/d-4733844/kisah-horor-yang-viral-di-medsos-selain-kkn-di-desa-penari>. [Accessed: 10-Oct-2019).
- [4] B. Saberi and S. Saad, "Sentiment Analysis or Opinion Mining: A Review", *Ijaseit* Vol.7 No. 5 pp 1660-1666, 2017.
- [5] Troussas, C. et al, "Sentiment analysis of Facebook statuses using Naive Bayes classifier for language learning", In IISA 2013. IEEE, pp. 1-6. Available at: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6623713> [Accessed May 7, 2015]
- [6] N.M.S. Hadna, P.I. Santosa, W.W. Winarno, "Studi Literatur Tentang Perbandingan Metode Untuk Proses Analisis Sentimen Di Twitter", Seminar Nasional Teknologi Informasi dan Komunikasi 2016 (SENTIKA 2016), 2016, pp 57-64.
- [7] A. Alsaedi, M.Z. Khan, "A Study on Sentiment Analysis Techniques of Twitter Data", *International Journal of Advanced Computer Science and Applications*, Vol. 10, No. 2, pp 361-374, 2019.
- [8] Billy Gunawan, Helen Sastypratiwi, Enda Esyudha Pratama, "Sistem Analisis Sentimen pada Ulasan Produk Menggunakan Metode Naive Bayes", *Jurnal Edukasi dan Penelitian Informatika*, Vol 4, No 2, pp 113-118, 2018
- [9] W. Athira Luqyana, I. Cholissodin, R.Setya Perdana, "Analisis Sentimen Cyberbullying pada Komentar Instagram dengan Metode Klasifikasi Support Vector Machine", *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer* Vol. 2, No. 11, November 2018, pp. 4704-4713, 2018
- [10] I. Zulfa, E Winarko, "Sentimen Analisis Tweet Berbahasa Indonesia dengan Deep Belief Network, IJCCS, Vol.11, No.2, July 2017, pp. 187-198, 2017
- [11] Liu, Bing, Hu, Mingqing, and Cheng, Junsheng (2005). "Opinion Observer: Analyzing and Comparing Opinions on the Web." *Proceedings of the 14th International World Wide Web Conference (WWW-2005)*, May 10-14, Chiba, Japan.
- [12] Github. Devid Haryalesmana. ID-Opinion Words. 27 Juli 2019 [Online]. Tersedia : <https://github.com/masdevid/ID-OpinionWords>. [Diakses : 27 Juli 2019].