

ANALISIS SENTIMEN TERHADAP DATA TWEET UNTUK BADAN PENYELENGGARA JAMINAN SOSIAL (BPJS) MENGGUNAKAN PROGRAM R

SENTIMENT ANALYSIS ON TWEET DATA FOR BADAN PENYELENGGARA JAMINAN SOSIAL (BPJS) USING R PROGRAMMING

Oleh: Fahmi Nurulbaiti ¹⁾ Retno Subekti, M. Sc.²⁾
Program Studi Matematika, Jurusan Pendidikan Matematika, FMIPA UNY
1) f.nurulbaiti@gmail.com 2) retnosubekti@uny.ac.id

Abstrak

Seiring dengan perkembangan teknologi yang terjadi saat ini, pendapat masyarakat terhadap program pemerintah Indonesia dapat dilihat melalui jejaring sosial Twitter. Badan Penyelenggara Jaminan Sosial (BPJS) adalah salah satu penyelenggara program pemerintah Indonesia yang sering mendapat tanggapan baik positif maupun negatif dari masyarakat. Penelitian ini dilakukan dengan tujuan untuk mengetahui langkah-langkah dalam melakukan analisis sentimen masyarakat Indonesia dengan menggunakan Program R, mengetahui topik-topik yang sering diperbincangkan masyarakat Indonesia terkait dengan BPJS serta mengetahui tanggapan masyarakat terhadap keberadaan BPJS. Penelitian ini menggunakan data berbentuk teks yang didapatkan dari tulisan warganet pada jejaring sosial Twitter. Data yang digunakan yaitu sebanyak 1000 *tweets* dan dibatasi hanya tweet berbahasa Indonesia saja. Teknik analisis data pada penelitian ini dilakukan dengan *text mining* dan *scoring sentiment*. Pengambilan data, pengolahan data, serta visualisasi informasi dilakukan dengan bantuan Program R. Hasil penelitian menunjukkan, langkah-langkah untuk melakukan analisis sentimen dengan program R yaitu dengan memiliki API untuk melakukan *crawling data*, kemudian melakukan *transforming*, *filtering*, dan *stemming* pada tahap *preprocessing*, lalu memvisualisasikan data dalam bentuk histogram dan *wordcloud*, terakhir *scoring tweet* dan menghitung skor tweet keseluruhan untuk mengetahui sentimen yang dominan. Visualisasi pada *wordcloud* maupun *histogram* terhadap data *tweets* BPJS memuat informasi mengenai topik-topik yang paling sering dibicarakan masyarakat di jejaring sosial Twitter berkaitan dengan BPJS. Topik tersebut antara lain riba, dokter, tolak, kesehatan, dan pasien. Analisis sentimen *tweets* masyarakat pada jejaring sosial Twitter terhadap BPJS adalah sentimen negatif. Hal ini berarti program BPJS masih belum sepenuhnya memuaskan masyarakat.

Kata kunci: analisis sentimen, BPJS, program R, twitter

Abstract

Along with current technological developments, public opinion on Indonesian government programs can be viewed through Twitter's social network. Badan Penyelenggara Jaminan Sosial (BPJS) is one of the Indonesian government program organizers who often get positive and negative responses from the public. This research was conducted with the aim to know the steps in conducting sentiments analysis of Indonesian society by using Program R, knowing the topics that are often discussed by Indonesian people related to BPJS and to know the public response to the existence of BPJS. This research uses text-shaped data obtained from netizen's posts on Twitter. The data used are 1000 tweets and limited to Indonesian language tweet only. Data analysis technique in this research is done with text mining and scoring sentiment. Data retrieval, data processing, and information visualization are done with the help of R Programming. The results showed that the steps to perform sentiment analysis with the R programming is to have API for crawling data, perform transforming, filtering, and stemming in the preprocessing stage, then visualize the data in the form of histogram and word cloud, last scoring tweet and calculate score tweet the whole to know the dominant sentiment. Visualization of the word cloud and histogram of the data tweets BPJS contains information on the topics most often discussed in Twitter related to BPJS. Topics include usury, doctors, rejected, health, and patients. The analysis of the tweets sentiment of the public on Twitter's social network against BPJS is a negative sentiment. This means that the BPJS program has not fully satisfied the community.

Keywords: BPJS, R Programming, sentiment analyst, twitter

PENDAHULUAN

Badan Penyelenggara Jaminan Sosial (BPJS) adalah badan hukum yang dibentuk pemerintah Indonesia untuk menyelenggarakan program jaminan sosial di Indonesia. Tujuan keberadaan BPJS untuk memberikan jaminan terpenuhinya kebutuhan dasar hidup yang layak bagi setiap peserta BPJS dan/atau anggota keluarga peserta BPJS. Berdasarkan data yang ada sampai Februari 2016 kepesertaan BPJS mencapai 162,78 juta jiwa, kepesertaan Pekerja Bukan Penerima Upah (PBPU) atau pekerja mandiri masih minim dengan 9,052 juta jiwa, sedangkan Non-PBPU (PBI, PPU dan BP) mencapai 124,37 juta jiwa. Jumlah ini semakin meningkat pada tahun 2017 yang dilaporkan sudah mencapai 178 juta jiwa.

Seiring dengan perkembangan teknologi yang ada, opini masyarakat terhadap program pemerintah Indonesia dapat dilihat melalui jejaring sosial yang banyak digunakan oleh masyarakat, salah satunya yaitu jejaring sosial Twitter. Data yang dirilis Twitter Indonesia pada akhir 2016, menyebutkan sebutkan bahwa 77% pengguna Twitter di Indonesia merupakan pengguna aktif yang produktif menulis *tweets*. Hal tersebut dapat dilihat dari jumlah tweet yang dihasilkan sepanjang 2016 yang mencapai 4,1 miliar tweet (Herman, 2017). Banyaknya pengguna Twitter yang aktif berkomentar sudah sering dimanfaatkan oleh instansi ataupun wirausaha untuk meninjau pendapat pengguna mengenai produk yang tengah dipasarkan. Jejaring sosial memfasilitasi penggunaanya dalam memanfaatkan hak untuk melakukan penilaian dan memfasilitasi instansi ataupun wirausaha

dalam mendapatkan informasi yang dibutuhkan sesuai tujuan-tujuannya. Proses ini sering disebut analisis sentimen. Analisis sentimen merupakan sebuah penggambaran polaritas pada suatu teks atau kata. (Esuli & Sebastiani, 2006:418).

Penelitian dengan menggunakan metode analisis sentimen pada jejaring sosial twitter antara lain oleh Kumar & Sebastian (2012:373-375) dilakukan untuk menentukan orientasi kata pendapat di tweets dengan pendekatan menggunakan bahasa dan kamus berbasis kumpulan kata serta menyajikan studi kasus sebagai gambaran penggunaan dan keefektifan sistem yang diusulkan. Penelitian tentang analisis sentimen di Indonesia pernah dilakukan sebelumnya untuk mengetahui pendapat masyarakat terhadap produk yaitu Telkom Indihome berdasarkan opini publik menggunakan Improved K-NN, data opini publik ini juga didapatkan dari opini masyarakat lewat tweet (Herdiawan, 2012). Selain itu, Hizbullah (2015) memilih Decision Tree dalam melakukan analisis sentimen pada acara televisi di Indonesia menggunakan opini masyarakat dari jejaring sosial Twitter berupa tweet, tujuannya untuk menghasilkan suatu referensi untuk mengevaluasi kualitas acara televisi berdasarkan respon pemirsa terhadap suatu acara yang ditayangkan.

Program R dipilih sebagai program untuk melakukan proses analisis sentimen dalam penelitian ini dengan mempertimbangkan bahwa program R sudah memiliki *packages* pendukung lengkap untuk pengambilan data, analisis data maupun visualisasi hasil analisis dan alasan lain yaitu karena belum banyak penelitian di Indonesia yang menggunakan program R untuk

melakukan analisis sentimen terhadap tweet berbahasa Indonesia.

METODE PENELITIAN

Deskripsi Data

Data yang digunakan berupa data berbentuk teks yang didapatkan dari tulisan warganet di jejaring sosial Twitter. Data tersebut berupa *tweets* berisi kata sesuai topik yang sudah ditentukan untuk diteliti yaitu kata bpjs. Banyaknya *tweets data* yang diambil juga sudah ditentukan yaitu 1000 *tweets*.

Jenis dan Sumber Data Penelitian

Jenis data yang digunakan dalam penelitian ini adalah data primer, yaitu data yang diambil langsung oleh penulis dari jejaring sosial Twitter. Data tersebut berupa *tweets* yang mencantumkan kata bpjs di dalamnya.

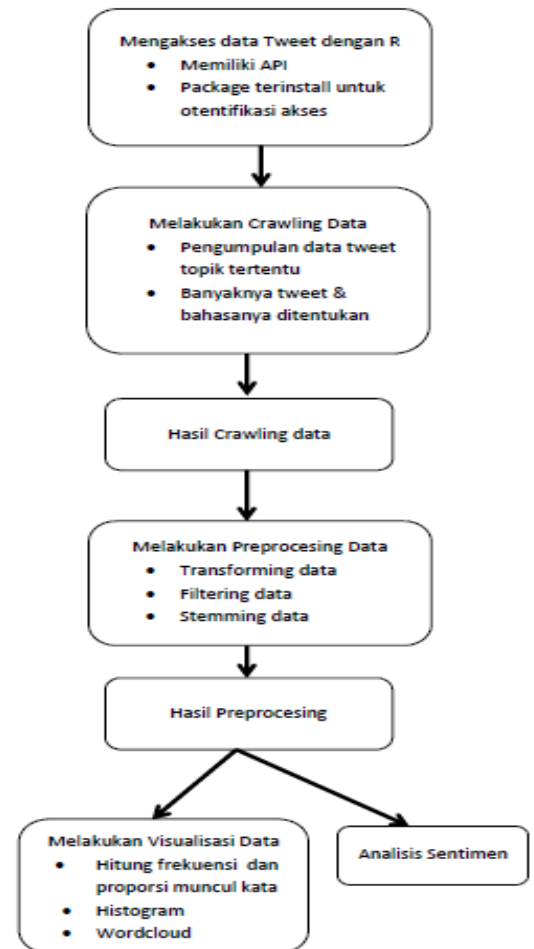
Teknik Analisis Data

Teknik analisis data pada penelitian ini menggunakan *text mining*, sedangkan aplikasi utama yang digunakan adalah program R dengan memanfaatkan fungsi Rcurl, twitterR, OAuth, tm, ggplot2 dan wordcloud. Proses mendapatkan data, mengolah data, menganalisis data dan juga memperoleh hasil visualisasi informasi dari data dilakukan dengan Program R.

HASIL PENELITIAN DAN PEMBAHASAN

Analisis sentimen dalam penelitian ini dilakukan terhadap data yang diperoleh dari jejaring sosial Twitter. Data dari Twitter dapat diakses dengan menyambungkan Program R terhadap Twitter dengan memiliki API Twitter dan menjalankan packages Rcurl, twitterR, OAuth. Apabila sudah terkoneksi maka bisa dilakukan *crawling data* dengan kata kunci tertentu. Penelitian ini menggunakan kata bpjs sebagai kata kunci.

Proses pelaksanaan penelitian ini dilakukan dengan langkah-langkah seperti tampak pada gambar 1.



Gambar 1 Proses Olah Data Tweet

Crawling Data

Proses *crawling data* yaitu mengumpulkan data dari sebuah laman dan menyimpannya untuk diatur dan dianalisis lebih lanjut. Dalam penelitian ini dilakukan *crawling data* terhadap jejaring sosial Twitter dengan menggunakan RStudio sebagai alat pengumpulan sekaligus analisis data tersebut.

Hasil *Crawling Data*:

[1] "Balitoday: BPJS Kesehatan
klungkung Serahkan Kartu KIS untuk
Sulinggih ... - Tribun...
<https://t.co/65z1GUVbKc> #klungkung
#bali"

[2] "marzukialijo: BPJS Kesehatan
Menuju Cakupan Semesta

<https://t.co/BU7DFUuoBF> #aph

<https://t.co/YHhvFNi5SD>"

[3] "tafanalaras: BPJS Kesehatan Menuju Cakupan Semesta

<https://t.co/xMAR5BIeOj> #aph

<https://t.co/QFhnMWEjDI>"

[4] "LalalJuwit:

<https://t.co/9wXcpNAGki> ↔ Anggota Dewan Ini Minta Data Kepesertaan BPJS D..."

[5] "umjyy: Makanya pakai BPJS, unnie.
<https://t.co/tBnWTCZF3R>"

[6] "CetingspHakim: @BPJSKesehatanRI Mungkin perlu sosialisasi / diadakan pemberitahuan yang lebih lagi lewat RT / RW tentang informasi...

<https://t.co/Gia9eAcUDk>"

[7] "budiono_taslim: @tribunnews @Tribunwow Dr Kiki agr tdk berhadapan dg pasien ya pantasya ditempatkan di pulau terluar,yg mungkin pasien blm ada bpjs"

[8] "GoTotheJek: RT @tribunnews: Kontroversi Dokter Tolak Pasien BPJS karena Takut Dosa Riba, Begini Pendapat NU dan Para Ahli!

<https://t.co/4jEKRUxCKd> via..."

[9] "AragornArwena: @SafirSenduk BPJS"

[10] "TINBOS2: RT @tribunnews: Kontroversi Dokter Tolak Pasien BPJS karena Takut Dosa Riba, Begini Pendapat NU dan Para Ahli!

<https://t.co/4jEKRUxCKd> via..."

Preprocessing Data

Berdasarkan ketidakaturan struktur data teks, maka proses *text mining* memerlukan beberapa tahap awal yang pada intinya adalah mempersiapkan agar teks dapat diubah menjadi lebih terstruktur yaitu dengan *preprocessing* data. Proses *preprocessing* dilakukan dengan 3 tahapan yaitu *transforming data*, *filtering data* dan *stemming data*.

1. *Transforming Data* dilakukan setelah data teks diubah menjadi bentuk *corpus*/kumpulan kata. Transformasi data *tweets* dilakukan dengan

menyeragamkan semua data teks menjadi satu ukuran (huruf kecil), penghapusan tanda baca (*punctuation*), dan penghapusan spasi berlebih yang ada pada sela antara dua kata.

2. *Filtering Data*, proses penyaringan kata-kata hasil *crawling* yaitu dengan membuang kata-kata yang kurang penting untuk dianalisis. Proses ini dilakukan dengan memanfaatkan fungsi penghapusan kata-kata yang termasuk dalam daftar *stopwords*, yaitu berisi daftar kata-kata yang tidak deskriptif namun sering muncul.

Hasil sampai tahap *filtering*:

[1]bpjs kesehatan klungkung serahkan

kartu kis sulinggih tribun klungkung bali

[2]bpjs kesehatan cakupan semesta aph

[3] bpjs kesehatan cakupan semesta aph

[4] anggota dewan data kepesertaan bpjs

[5] pakai bpjs unnie

[6] bpjskesehatanri sosialisasi diadakan pemberitahuan rt rw informasi

[7] tribunnews tribunwow dr kiki agr tdk berhadapan dg pasien ya pantasya ditempatkan pulau terluar yg pasien blm bpjs

[8] rt tribunnews kontroversi dokter tolak pasien bpjs takut dosa riba pendapat nu ahli via

[9] safirsenduk bpjs

[10] rt tribunnews kontroversi dokter tolak pasien bpjs takut dosa riba pendapat nu ahli via

3. *Stemming Data*, tahap mencari kata dasar dari kumpulan kata hasil *filtering* dan kata-kata dengan penulisan yang belum baku diproses agar menjadi kata dasar yang baku. Daftar kata dan hasil *stemming* antara lain seperti pada tabel 1.

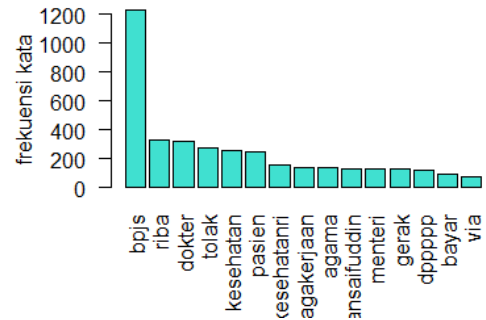
Tabel 1 Daftar Kata dalam Proses Stemming

Kata-kata	Hasil Stemming
aja	saja
bgt	banget
byk	banyak
cont	contoh
dfn	dengan
dg	dengan
dpt	dapat
enggak	tidak
hrs	harus
kalo	kalau
kes	kesehatan
klo	kalau
koq	kok
krn	karena
masy	masyarakat
mhn	mohon
min	minimal
msh	masih
org	orang
pake	pakai
sbg	sebagai
slain	selain
tak	tidak
trs	terus
udah	sudah
utk	untuk
yg	yang

Visualisasi data

Visualisasi data yaitu penyajian data dalam bentuk yang lebih sederhana dan menarik. Hal ini dilakukan untuk memudahkan dalam memahami secara langsung informasi dari data yang sudah diolah. Visualisasi data yang dilakukan dibagi menjadi dua yaitu penyajian data dalam bentuk *wordcloud* dan penyajian dalam bentuk histogram

Kata dengan frekuensi tertinggi



Gambar 1 Histogram Kata BPJS dengan Frekuensi Terurut

Histogram pada gambar 2 menunjukkan kata yang paling sering disebutkan dalam tweet BPJS. Selain itu juga 14 kata yang lain yang frekuensi kemunculannya mengikuti kata paling banyak. Kata yang tertampil pada *wordcloud* berwarna-warni dari yang paling gelap sampai warna yang paling terang dengan maksud menunjukkan urutan frekuensi munculnya kata dalam data dari frekuensi terbanyak.



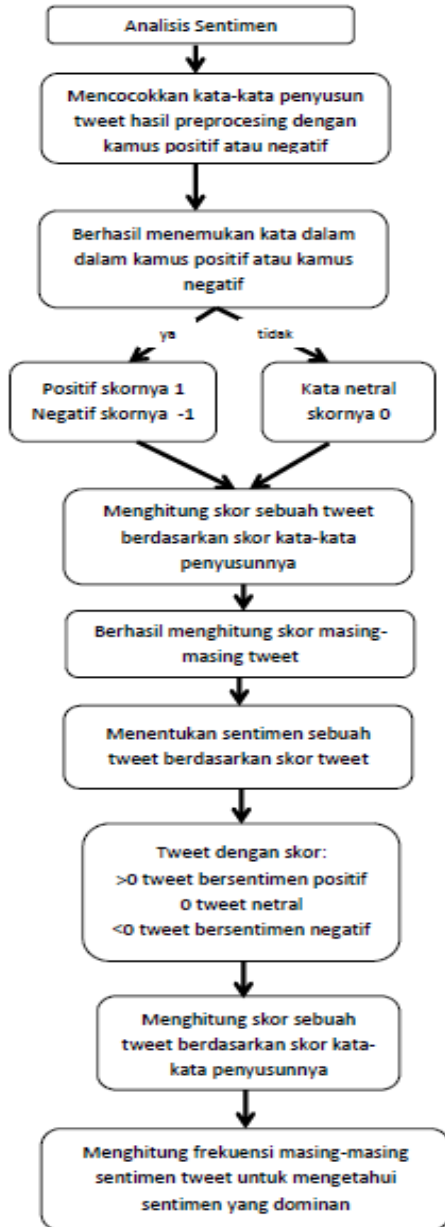
Gambar 2 Wordcloud Tweets BPJS

Kata dengan frekuensi muncul paling tinggi berwarna paling gelap dan berukuran paling besar. Frekuensi kemunculan kata dalam tweets ini menunjukkan informasi mengenai topik-topik yang paling sering dibicarakan masyarakat di jejaring sosial Twitter berkaitan dengan BPJS. Topik tersebut antara lain riba,

dokter, tolak, pasien, kesehatan dan bpjskesehatanri.

Analisis Sentimen

Proses analisis sentimen dilakukan dengan langkah-langkah seperti tampak pada gambar 4.



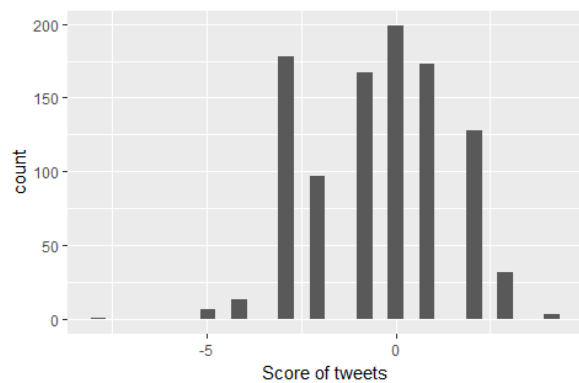
Gambar 3 Proses Analisis Sentimen

Hasil dari menjalankan fungsi `score.sentiment` terhadap tweet BPJS menghasilkan penghitungan skor sentimen untuk masing-masing tweet seperti tampak pada gambar 5.

score	text
1	BPJS Kesehatan Klungkung Serahkan Kartu KIS untuk S...
2	1 BPJS Kesehatan Menuju Cakupan Semesta https://t.co...
3	1 BPJS Kesehatan Menuju Cakupan Semesta https://t.co...
4	-1 https://t.co/9wXcpNAGKi <U+2190><U+2190> #bi...
5	0 Makanya pakai BPJS, unnie. https://t.co/tBnWTCZF3R
6	0 @BPJSKesehatanRI Mungkin perlu sosialisasi / diadaka...
7	-2 @tribunnews @TribunWow Dr Kiki agr tdk berhadapan...
8	-5 RT @tribunnews: Kontroversi Dokter Tolak Pasien BPJS...
9	0 @SafirSenduk BPJS
10	-5 RT @tribunnews: Kontroversi Dokter Tolak Pasien BPJS...
11	1 @CetingspHakim hrs ada surat dr kepolisian dan sura...
12	-5 tribunnews: Kontroversi Dokter Tolak Pasien BPJS kare...
13	0 BPJS RT @SafirSenduk "Budget Pas-pasan, Jiwa Sosialita."
14	1 BPJS Kesehatan Menuju Cakupan Semesta https://t.co...
15	1 BPJS Kesehatan Menuju Cakupan Semesta https://t.co...

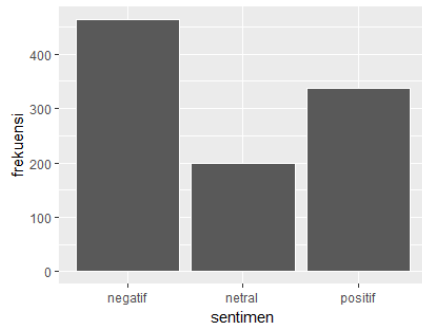
Gambar 4 Data Frame Hasil Skor Sentimen Setiap Tweet BPJS

histogram skor berdasarkan sentimen seperti tampak pada gambar 6



Gambar 5 Histogram Frekuensi Masing-masing Skor Sentimen Tweet BPJS

Tweet dengan skor 0 frekuensinya paling banyak, namun bukan berarti bahwa tweet netral yang mendominasi. Sentimen dominan dapat diketahui dengan menghitung persebaran tweet. Rentang skor untuk tweet BPJS yaitu dari skor paling minimal -8 hingga skor paling maksimal 4 dengan 801 tweet mengandung sentimen. Tweet dengan skor <0 termasuk dalam tweet negatif, tweet dengan skor >0 termasuk tweet bersentimen positif.



Gambar 6 Histogram Persebaran

Karena skor tweet bernilai negatif lebih banyak maka sentimen yang dominan dalam tweet-tweet dengan topik BPJS yaitu sentimen negatif.

SIMPULAN DAN SARAN

Simpulan

Hasil penelitian menunjukkan, langkah-langkah untuk melakukan analisis sentimen dengan program R yaitu dengan memiliki API untuk bisa mengakses data /melakukan *crawling data*, kemudian melakukan *transforming*, *filtering*, dan *stemming* pada tahap *preprocessing*, lalu memvisualisasikan data dalam bentuk histogram dan *wordcloud*, terakhir *scoring* tweet dan menghitung skor tweet keseluruhan untuk mengetahui sentimen yang dominan. Visualisasi pada *wordcloud* maupun *histogram* terhadap data *tweets* BPJS memuat informasi mengenai topik-topik yang paling sering dibicarakan masyarakat di jejaring sosial Twitter berkaitan dengan BPJS. Topik tersebut antara lain riba, dokter, tolak, kesehatan, dan pasien. Hasil analisis sentiment terhadap *tweets* masyarakat pada jejaring sosial Twitter mengenai BPJS diperoleh bahwa respon masyarakat dominan sentimen negatif.. Hal ini berarti program BPJS masih belum sepenuhnya memuaskan masyarakat

Saran

Penelitian selanjutnya diharapkan akan lebih baik jika data yang diambil untuk digunakan sebagai data analisis lebih banyak karena diharapkan dapat lebih merepresentasikan pendapat masyarakat, proses *stemming* pada penelitian selanjutnya diharapkan menggunakan metode atau fungsi tertentu agar dapat dilakukan dengan lebih akurat dan efisien. Hasil penelitian didapatkan dari menghitung polaritas tiap kata secara satu per satu kata, sehingga perhitungan polaritas sentimen pada kalimat yang mengandung kata negasi belum tercakup. Hasil diharapkan akan lebih akurat dengan menambahkan kamus ketiga sebagai pertimbangan yaitu kamus negasi untuk penelitian lebih lanjut. Program R masih tergolong baru untuk analisis sentimen berbahasa Indonesia, namun selain dari jejaring sosial Twitter, R sudah memiliki *packages* yang dapat memfasilitasi untuk pengambilan data dari jejaring sosial yang lainnya seperti Facebook, Instagram ataupun WhatsApp.

DAFTAR PUSTAKA

- Berry, M. W., & Kogan, J. (2010). *Text Mining Application and Theory*. Chicester: WILEY.
- eBizMBA Inc. (2017, July 1). *Top 15 Most Popular Social Networking Sites | July 2017*. Retrieved July 7, 2017, from publishes eBizMBA: <http://www.ebizmba.com/articles/social-networking-websites>
- Esuli, A., & Sebastiani, F. (2006). Sentiwordnet: A Publicly available lexical resource for opinion mining. *In Proceedings of LREC*, 417-422.
- Febri S, B. (2014, Juny 30). *[Pengenalan] Apa Itu Twitter API dan Pembuatan Consumer Key dan Consumer Secret ?*

- Diakses July 2, 2017, dari Jago Coding: http://jagocoding.com/tutorial/427/Pengenalan_Apa_Itu_Twitter_API_dan_Pembuatan_Consumer_Key_dan_Consumer_Secret
- Feldman, R., & Sanger, J. (2007). *The Text Mining Handbook: Advanced Approach in Analyzing Unstructured Data*. New York: Cambridge University Press.
- Hepburn, A. (2010, May 12). *Infographic: Twitter Statistic, Facts & Figures*. Diakses July 2, 2017, dari Digital Buzz Blog: <http://www.digitalbuzzblog.com/infographic-twitter-statistics-facts-figures/>
- Herdiawan. (2012). Analisis Sentimen terhadap Telkom Indihome Berdasarkan Opini Publik Menggunakan Metode Improve K-Nearest Neighbor. *Komputa*.
- Herman. (2017, May 3). *Indonesia Masuk Lima Besar Pengguna Twitter*. Diakses July 12, 2017, dari BeritaSatu.com: <http://www.beritasatu.com/digital-life/428591-indonesia-masuk-lima-besar-pengguna-twitter.html>
- Hizbullah, M. (2015, September 14). *Digital library - Perpustakaan Pusat Unikom - Knowledge Center*. Diakses July 8, 2017, dari ELIB UNIKOM: <http://elib.unikom.ac.id/gdl.php?mod=browse&op=read&id=jbptunikompp-gdl-hizbullahm-32767&q=Mulsy,%20Hizbullah.%202015>
- Hotho, A., Andreas, N., & Gerhard, P. (2005). A brief survey of text mining. *In Ldv Forum*, 1-37.
- Kadir, M. Y. (2012). *Ekstraksi dan Visualisasi Informasi Lalu Lintas Kota Bandung dari Twiter*. Bandung: Universitas Pendidikan Indonesia.
- Kumar, A., & Sebastian, T. M. (2012). Sentiment Analysis on Twitter. *IJCSI International Journal of Computer Science Issues*, 9(4), 372-378.
- Ling, J., Kencana, I. P., & Oka, T. B. (2014). Analisis Sentimen Menggunakan Metode Naive Bayes Classifier dengan Seleksi Fitur Chi Square. *E-Jurnal Matematika*, 3(3), 92-99.
- Liu, B. (2011). *Web Data Mining: Supervised Learning (Chapter 3) in Exploring Hyperlinks, Contents and Usage Data*. Berlin Heidelberg: Springer.
- Liu, B. H. (2005). Opinion observer: analyzing and comparing opinions on the Web. *International World Wide Web Conference* . 5, hal. 342-351. Chiba: IW3C2.
- Nurmansyah. (2013, Juny 12). *R, Bahasa Pemrograman untuk Analisis Data dan Statistik*. Diakses July 18, 2017, dari Kompasiana: https://www.kompasiana.com/nurmansyah/r-bahasa-pemrograman-untuk- analisis-data-dan- statistik_552fab4f6ea8349b138b456f
- Pak, A., & Paurobek, P. (2010). *Twitter as a Corpus for Sentiment Analysis and Opinion Mining*. France: Universite de Paris, Laboratoire LIMSI-CNRS.
- Putri, A. E. (2014). *Paham BPJS Badan Penyelenggara Jaminan Sosial*. Jakarta: Kantor Perwakilan Indonesia.
- Qomaruddin. (2012). Badan Hukum Publik Badan Penyelenggara Jaminan Sosial dan Transformasinya Menurut Undang-Undang Nomor 24 Tahun 2011 tentang Badan Penyelenggara Jaminan Sosial. *Journal Legilasi Indonesia*, 9(2), 230.
- Rozi, I., Pramono, S., & Dahlan, E. (2012). Implementasi Opinion Mining (Analisis Sentimen) untuk Ekstraksi Data Opini Publik pada Perguruan Tinggi. *Jurnal EECCIS*, 6(1), 37-43.
- Sumartini, N. W. (2011). *Text Mining Classifier dengan Metode Naive Bayes dan Support Vector Machines untuk Sentiment Analysis*. Denpasar: Universitas Udayana.

Tala, F. (2003). *A Study of Stemming Effect on Information Retrieval in Bahasa Indonesia*. Amsterdam: Universiteti van Amsterdam The Netherlands.

Yudistira, A. (2008, May 3). *Mengenal Bahasa R*. Diakses July 16, 2017, dari R blogs / lang: <https://www.r-bloggers.com/lang/indonesian/50>