

Analisis Sentimen Twitter Pasca Pengumuman Hasil Pilpres 2019 Menggunakan Metode *Lexicon Analysis*

Adi Pandu Rahmat Nababan, Arie S. M. Lumenta, Yaulie D. Y. Rindengan, Fransisca J. Pontoh, Yuri V. Akay
Teknik Elektro Universitas Sam Ratulangi Manado, Jl. Kampus Bahu-Unsrat Manado, 95115

17021406115@student.unsrat.ac.id, al@unsrat.ac.id, rindengan@unsrat.ac.id, fransisca@unsrat.ac.id, yuriakay@unsrat.ac.id

Submitted: Januari 2020

Received: Februari 2020

Publish: Maret 2020

Abstract - Rapid development in the field of information technology has become a new paradigm. Twitter is a technology product that allows users to convey information of events from the real world to social media so that it can be known to a wide audience through the internet media. One interesting event was the announcement of the election results by the KPU on May 21, 2019. Many tweets on Twitter became trending topics related to the incident. Therefore in this study tweet data retrieval will be performed, then analyze the data with the lexicon analysis method. Then display the data that has been analyzed. In this study there are 21 hashtag trending topics that were drawn. Furthermore, the tweet data goes through the preprocessing stages, namely tokenizing, word normalization, filtering and stemming. Data were analyzed using the lexicon analysis method and lexicon dictionary support and were run in the R language to determine positive, negative and neutral sentiments. From the results of the study the highest positive sentiment on the hashtag #JokowiAminSudahMenang by 54%. The highest negative sentiment in the hashtag #PrabowoBukanPemimpin is 51%, and the highest neutral sentiment is in the #98jagademocracy of 72%. Lexicon analysis can classify tweet data into three classes, positive, negative and neutral. And the level of accuracy of analysis is largely determined by the number of words in the Lexicon dictionary.

Keywords – KPU; lexicon analysis; sentiment analysis; R programming; presidential election; twitter.

Abstrak — Perkembangan yang cepat dalam bidang teknologi informasi menjadi sebuah paradigma baru. Twitter salah satu produk teknologi yang memungkinkan pengguna menyampaikan informasi peristiwa dari dunia nyata ke media sosial sehingga dapat diketahui khalayak luas lewat media internet. Salah satu peristiwa menarik adalah pengumuman hasil Pilpres oleh KPU pada tanggal 21 Mei 2019. Banyak kicauan di twitter menjadi trending topik yang berhubungan dengan kejadian tersebut. Karena itu pada penelitian ini akan dilakukan penarikan data tweet, kemudian menganalisis data dengan metode lexicon analysis. Lalu menampilkan data yang telah dianalisis. Pada penelitian ini terdapat 21 hashtag trending topik yang ditarik. Selanjutnya data tweet melalui tahap *preprocessing* yaitu *tokenizing*, normalisasi kata, *filtering* dan *stemming*. Data dianalisis menggunakan metode *lexicon analysis* dan dukungan kamus *lexicon* dan dijalankan dalam bahasa R untuk menentukan sentimen positif, negatif dan netral. Dari hasil penelitian sentimen positif tertinggi pada hashtag #JokowiAminSudahMenang sebanyak 54%. Sentimen

negatif tertinggi pada hashtag #PrabowoBukanPemimpin sebanyak 51%, serta sentimen netral tertinggi terdapat pada

#98jagademokrasi sebanyak 72%. *Lexicon analysis* dapat mengklasifikasikan data tweet dalam tiga kelas yaitu positif, negatif dan netral. Dan tingkat akurasi analisis sangat ditentukan oleh banyaknya jumlah kata pada kamus *lexicon*.

Kata kunci — KPU; analisis sentimen; *lexicon analysis*; pemrograman R; pilpres; twitter.

I. PENDAHULUAN

Saat ini hampir segala informasi tersedia secara daring. Orang-orang dapat dengan mudah mengekspresikan opini-opini mereka pada berbagai topik di laman media sosial. Para pengguna twitter banyak yang tertarik akan topik utama di dunia maya, dan memberikan tanggapan ataupun *retweet* terhadap topik utama. Banyaknya tanggapan dalam periode tertentu terhadap topik yang ada pada twitter akan mengakibatkan satu topik menjadi *trending* topik[1]. *Twitter* merupakan salah satu media sosial paling populer. Kegiatan yang sering dilakukan di sosial media adalah dengan melakukan *update* status, melakukan *tweet*, *comment* ataupun *re-tweet* yang merupakan kegiatan masyarakat dalam kehidupan sehari-hari yang dicerminkan melalui sosial media[2]. Hal itu merupakan persepsi atau pandangan dari para pengguna sosial media terhadap isu-isu atau peristiwa yang sedang terjadi.

Twitter merupakan situs *microblogging* yakni *blog* yang memungkinkan pengguna menuliskan teks singkat yang disebut *tweet*, *tweet* dibatasi hanya 140 karakter dan dapat dilihat oleh pengikut. Setelah layanan *microblogging* ini menggumamkan pertumbuhan pengguna yang lebih dari perkiraan kuartal pertama. Ini disebabkan meningkatnya minat pengguna terhadap berita dan politik[2].

Pesatnya pertumbuhan pengguna *twitter* tiap tahunnya menyebabkan semakin banyaknya data yang dihasilkan, fenomena ini disebut juga dengan fenomena *big data*[3]. Berdasarkan data-data tersebut banyak peneliti melakukan penelitian terhadap data yang dihasilkan oleh *twitter*. *Twitter* telah dijadikan menjadi suatu objek penelitian.

Berdasarkan penjelasan yang telah dipaparkan maka akan dilakukan sebuah penelitian tentang analisis sentimen untuk mengetahui sentimen publik terhadap isu yang berkembang di media sosial. Dengan menggunakan metode *lexicon analysis* dan sistem yang dibangun dengan bahasa pemrograman R maka tweet-tweet akan diklasifikasikan ke dalam tiga kelas yaitu positif, negatif dan netral. Klasifikasi

ini akan memudahkan untuk melihat sentimen publik terhadap trending topik.

Penelitian ini mengangkat topik tentang sentiment publik terhadap isu-isu yang berkembang pasca Pemilu 17 April 2019 dan yang lebih spesifik lagi isu-isu pasca pengumuman pemenang Pilpres pada 21 Mei 2019 oleh Komisi Pemilihan Umum Republik Indonesia. Pada penelitian ini dilakukan mulai dari proses pengumpulan data, menganalisis data twitter dan kemudian menampilkan data hasil analisis dalam bentuk diagram batang dan pie chart.

A. Penelitian Terkait

Fenomena *Trending Topic* di Twitter: Analisis Wacana Twit #Savehajilulung. Tujuan penelitian adalah untuk mendapatkan gambaran tentang analisis wacana twit #SaveHajiLulung yang menjadi *trending topic* di Twitter. Penelitian ini menggunakan metode analisis isi wacana Van Dijk yaitu struktur makro, super struktur dan struktur mikro[4].

Analisis Sentimen pada Twitter Mahasiswa Menggunakan Metode *Backpropagation*. Penelitian ini bertujuan untuk mengetahui faktor emosional dalam pengaruh terhadap motivasi belajar mahasiswa belajar dan mengajar di lingkungan kampus[5].

Analisis Sentimen pada Twitter untuk Menilai Performansi Program Televisi dengan Kombinasi Metode *Lexicon-Based* dan *Support Vector Machine*. Penelitian ini dilakukan untuk mengetahui penilaian terhadap kualitas program televisi yang selama ini dilakukan hanya dengan program rating. Pada penelitian ini metode yang digunakan adalah kombinasi dari metode *Lexicon-Based* dan *Support Vector Machine*[6].

Analisa Sentimen Menggunakan *Lexicon Based* untuk Melihat Persepsi Masyarakat Terhadap Kenaikan Harga Rokok pada Media Sosial Twitter. Penelitian ini mencoba menganalisis persepsi masyarakat kedalam kelas sentimen menggunakan metode *lexicon based* dengan SentiWordNet[7].

Eksplorasi *Trending Topik* Twitter Menggunakan *Text Mining*. Penelitian ini melakukan eksplorasi *trending* topik twitter menggunakan *text mining*. Dan penelitian ini menggunakan metode deskriptif kualitatif[8].

Text Mining dan Sentimen Analisis Twitter Pada Gerakan LGBT. Tujuan penelitian ini untuk mendeskripsikan frekuensi opini kicauan di Twitter terkait pro dan kontra terhadap gerakan LGBT[9].

Implementasi *Twitter Sentiment Analysis* untuk Review Film Menggunakan Algoritma *Support Vector Machine*. Dalam penelitian ini pengguna twitter akan memberikan opini tentang suatu hal, salah satunya film yang sedang tayang di bioskop. Opini pengguna bermanfaat bagi pengguna lain dan rumah produksi film berkaitan evaluasi film. Klasifikasi opini diperlukan untuk memudahkan pengguna dalam melihat opini dalam sentimen positif, negatif, atau netral. Algoritma yang digunakan dalam klasifikasi adalah *Support Vector Machine*[10].

B. Analisis Sentimen

Analisis sentimen adalah metode untuk menganalisis sebagian data untuk mengetahui emosi manusia. Analisis

sentimen dapat dikategorikan kedalam tiga task, yaitu *informative text detection*, *information extraction* dan *sentiment interestingness classification (emotional, polarity identification)*. *Sentiment classification* (negatif atau positif) digunakan untuk memprediksi sentiment polarity berdasarkan data sentimen dari pengguna[11]. Analisis sentimen dilakukan untuk melihat pendapat terhadap satu masalah atau peristiwa dan dapat juga digunakan untuk identifikasi kecenderungan hal yang sedang terjadi menjadi topik pembicaraan[12].

Analisis sentimen sendiri dapat dibagi pada tiga level yaitu:

1) Level Dokumen

Level dokumen menganalisis satu dokumen penuh dan mengklasifikasikan dokumen tersebut memiliki sentimen positif atau negatif. Level analisis ini berasumsi bahwa keseluruhan dokumen hanya berisi opini tentang satu entitas saja. Level analisis ini cocok diterapkan untuk membandingkan lebih dari satu entitas[13].

2) Level Kalimat

Level kalimat menganalisis satu kalimat dan menentukan tiap kalimat bernilai sentimen positif, netral atau negatif. Sentimen netral berarti kalimat tersebut bukan opini[13].

3) Level Entitas dan Aspek

Level aspek tidak dapat melakukan analisis pada konstruksi bahasa (dokumen, paragraf, kalimat, klausa atau frase) melainkan pada opini itu sendiri. Hal ini didasari bahwa opini terdiri dari sentimen (positif dan negatif) dan target dari opini tersebut. Tujuan dari level analisis ini adalah untuk menentukan sentimen entitas pada tiap aspek yang dibahas[13].

C. Media Sosial

Media sosial sebagai sebuah kelompok aplikasi berbasis internet yang dibangun di atas dasar ideologi dan teknologi Web 2.0, dan memungkinkan penciptaan dan pertukaran *user-generated content*. Web 2.0 menjadi *platform* dasar media sosial.

Internet dan web 2.0 menyediakan suatu *platform* yang digunakan untuk meningkatkan pelayanan yang dapat digunakan untuk: membuat dan berbagi pemikiran dan cerita (*Blogger* dan *Twitter*); berbagi informasi dan *links* (*Delicious*, *Digg* dan *Twine*); berbagi multimedia (*Youtube* dan *Flickr*); membuat dan berbagi pengetahuan (*Wikipedia*, *Yahoo Answer* dan *SlideShare*) dan membuat dan berbagi relasi (*Facebook*, *MySpace* dan *LinkedIn*) oleh grup-grup yang besar. Layanan inilah yang secara bersama-sama dikenal sebagai sosial media.

Sosial media merupakan suatu *platform* yang memberikan pelayanan dua arah yaitu membuat dan berbagi yang digunakan sebagai alat komunikasi baru di dalam era digital yang dapat membentuk jaringan pada komunitas yang memungkinkan untuk berkomunikasi secara *online* untuk membuat, mengatur, mengedit, mengomentari, menandai, mendiskusikan menggabungkan, mengkoneksikan dan bertukar informasi apapun di dalamnya. Salah satu jenis sosial media yang populer saat ini adalah Twitter. Twitter merupakan sebuah *microblogging* yang dapat mengirim pesan hingga 140

karakter secara instan melalui berbagai *platform*. 90% interaksi *Twitter* bukan berasal dari *website Twitter* melainkan dari *sms mobile*, pesan instan atau aplikasi desktop[14].

Saat ini beragam jenis media sosial seperti : *social networks, blogs, wikis, podcast, forums, content communities, microblogging*, dll. dapat digunakan untuk berbagai tujuan tertentu. Dengan mengimplementasikan teori *social presence, media richness* dan *social processes*, Kaplan dan Haenlein mengklasifikasikan media sosial menjadi enam jenis yaitu: *collaborative projects, blogs and microblogs, content communities, social networking sites, virtual game worlds, virtual communities*

Sedangkan jika dilihat berdasarkan kategorinya, media sosial dibagi menjadi empat kategori yaitu: *social networking, social coloboration, social publishing, social feedbacks*.

D. Social Media Analytics

Social Media Analytics (SMA) merupakan kegiatan yang berkaitan dengan pengembang dan evaluasi *tools* informatika dan *framework* untuk mengumpulkan, memantau, menganalisis, merangkum dan memvisualisasi data media[15]. Gartner Reasearch juga mendefinisikan SMA merupakan proses pemantauan, analisis, mengukur dan memprediksi interaksi digital, relationships, topik, ide atau konten pada media sosial. SMA bertujuan untuk melakukan proses analisis dan sintesis data media sosial sehingga menghasilkan informasi yang dapat dipergunakan oleh pihak-pihak yang memerlukan. Proses SMA dilakukan melalui tiga tahapan, yaitu : *capture, uderstand* dan *present*[16].

Tahapan *Capture* pada proses SMA merupakan proses mengumpulkan data media sosial yang relevan dengan kebutuhan dengan cara *collecting* data menggunakan *crawler tools* yang terkoneksi melalui *Application Programming Interface* (API) ke media sosial seperti: Facebook, Twitter, LinkedIn, YouTube, Pinterest, Google+, Tumblr, Foursquare, *Internet forums, blogs* dan *microblogs, Wikis, news sites, picture sharing sites, podcasts, and social bookmarking sites*, dll. Data yang dihasilkan dari proses *Capture* disimpan kedalam basis data dan dipersiapkan untuk proses berikutnya yaitu *Undetstad*. Pada tahapan ini data juga diproses untuk menghasilkan informasinya yang sesuai dengan kebutuhan, termasuk memodelkan bentuk [15].

Setelah menyelesaikan proses *Capture*, tahapan berikutnya dalah proses *Understand*. Proses *Understand* pada SMA merupakan proses pemilihan data yang relevan untuk melakukan pemodelan data, menghilangkan *noise* yang terdapat pada data, menseleksi data yang berkualitas dan melakukan proses analisis untuk memperoleh informasi yang lebih berkualitas. Proses analisis data pada tahapan ini menggunakan metode statistik, *text mining, data mining, natural language processing* (NLP), *machine tranlation, machine learning* dan *network*[17]. Beberapa teknik analisis data media sosial yang dapat digunakan untuk menghasilkan informasi spesifik antara lain : *Opinion mining (or sentiment analysis), Topic modeling, Social network analysis, Trend analysis*, dan *Visual* [15].

Tahapan terakhir dari proses SMA adalah *Present*. Proses *Present* merupakan proses untuk menampilkan atau memvisualisasikan informasi yang dihasilkan dari tahap *Understand*. Berbagai teknik visualisasi dapat digunakan untuk menampilkan informasi yang didapatkan dari proses analisis[15].

E. Twitter

Twitter merupakan media sosial yang banyak mendapatkan perhatian masyarakat Indonesia. Secara global, berdasarkan data pada Desember 2014, terdapat 284 juta pengguna aktif layanan tersebut. Jumlah pengguna Twitter di Indonesia sudah mencapai angka 50 juta. Jumlah ini diprediksi makin terus bertambah tiap tahun. Hal ini menyebabkan Indonesia merupakan pasar yang paling menguntungkan[18]. Data yang dilansir Statistika berdasarkan hasil penelitian PeerReach menunjukkan bahwa Indonesia tercatat sebagai pengguna Twitter terbanyak ketiga di dunia, dengan jumlah 6,5 persen setelah Amerika Serikat (24,3 persen), dan Jepang (9,3 persen). Sementara itu data dari perusahaan public relations, Webershandwick, menunjukkan bahwa pengguna Twitter, berdasarkan data PT. Bakrie Telecom, twitter memiliki 19,5 juta pengguna di Indonesia dari total 500 juta pengguna global. Twitter menjadi salah satu jejaring sosial paling besar di dunia sehingga mampu meraup keuntungan mencapai USD 145 juta. Kebanyakan pengguna Twitter di Indonesia adalah konsumen, yaitu yang tidak memiliki blog atau tidak pernah meng-upload video di Youtube namun sering update status di Twitter dan Facebook[19].

Chief Executive Officer Twitter, Dick Costolo, mengatakan bahwa Indonesia menjadi salah satu negara dengan pengguna Twitter terbanyak. Karena itulah Twitter pun akhirnya mendirikan kantor di Jakarta. Berdasarkan laporan Twitter di kuartal IV 2014, total pengguna aktifnya mencapai 288 juta per bulan. Menurut Costolo pengguna media sosial di Indonesia memiliki pengetahuan yang baik dengan dunia digital. Disamping itu pengguna Twitter di Indonesia dinilai sangat atraktif dan bersemangat dan dianggap sangat aktif menuliskan cuitan. Tidak jarang, hasil obrolan di lini masa menjadi *Trending topic* atau topik yang paling banyak dibicarakan di seluruh dunia[20].

F. Twitter Data Analytics

Shamanth Kumar, Fred Morstatter dan Huan Liu menyebutkan ada beberapa proses dalam melalukan *Twitter Data Analytics* yaitu: penarikan data twitter (*crawling twitter data*), penyimpanan data twitter (*storing twitter data*), analisis data twitter (*analizing twitter data*), visualisasi data twitter (*visualizing twitter data*)[21].

Berdasarkan *tweet* yang dihasilkan setiap harinya oleh pengguna, twitter dapat menjadi suatu sumber informasi sehingga dapat dilakukan proses *crawling* data Twitter dengan menggunakan *API Public* yang telah disediakan oleh Twitter. *Aplication Program Interface* (API) untuk mengakses data twitter dapat diklasifikasikan menjadi dua tipe berdasarkan desain dan metode akses, yaitu[21] :

1) REST API berdasarkan pada arsitektur REST yang sekarang digunakan mendesain web API. Tipe API ini menggunakan *pull strategy* untuk mendapatkan kembali

informasi data. Untuk mengumpulkan informasi seorang pengguna harus secara eksplisit memintanya[21].

- 2) Streaming API menyediakan aliran informasi publik yang berkelanjutan dari Twitter. Tipe API ini menggunakan *push strategy* untuk mendapatkan kembali informasi data. Satu kali melakukan permintaan informasi, *Streaming API* menyediakan aliran data yang selalu *update* tanpa *input* lebih lanjut dari pengguna[21].

Aliran API publik yang paling baik untuk digunakan dalam berbagai hal adalah *streaming API* sehingga dalam penarikan data Twitter akan digunakan tipe API ini. API Twitter hanya bisa diakses melalui permintaan otentikasi. Twitter menggunakan *Open Authentication* (OAuth) dan setiap permintaan harus dilakukan oleh pengguna Twitter yang sah. Akses ke dalam API Twitter dibatasi pada jumlah tertentu yang disebut *rate limit*. Batasan ini diterapkan pada tingkatan pengguna dan tingkatan aplikasi[21].

G. Microblogger Twitter Terminologi

Twitter merupakan salah satu media sosial yang sangat populer dan menempati ranking ke-2 pada *Alexa rank*. Lahirnya Twitter berasal dari ide Jack Dorsey pada tahun 2006 yang melihat kebiasaan orang-orang ingin berbagi kegiatan mereka saat bersama orang lain[22]. Dalam pengembangan Twitter, Jack Dorsey mengkombinasikan pola komunikasi dari satu ke banyak untuk menjadi pola dasar komunikasi yang diimplementasikan pada Twitter. Hal ini memungkinkan pengguna Twitter untuk berbagi informasi kepada banyak orang.

Neppelenbroek *et al.* menggambarkan arsitektur pengembangan Twitter dengan menggunakan model "4+1" yang dikembangkan oleh Kruchten[22]. Model ini digunakan untuk mendeskripsikan arsitektur perangkat lunak yang terfokus pada *logical*, *process*, *physical*, *development* dan *scenario view*. Dengan *Kruchten's view model*, Neppelenbroek *et al.* menggambarkan arsitektur Twitter dengan *Logical view*, *Process view*, *Physical view*, *Development view* dan *Scenario view*[22].

Lapisan *Back-end Service* dari Twitter menyimpan semua *tweets* yang diunggah oleh anggota dengan menggunakan MSQl sebagai database penyimpanan data. Pada lapisan *Search Engine*, Twitter menggunakan *Apaches Lucene*. *Search Engine* pada Twitter menggunakan metode *inverted indexing*, metode ini memisahkan *tweets* menjadi kata-kata (*words of a sentence*). Lapisan *Middle Layer* pada arsitektur Twitter pada dasarnya digunakan sebagai sistem antrian, sehingga tidak membabani *Back-end Service*. Lapisan *Middle Layer* pertama kali diimplementasikan oleh Starling dengan menggunakan bahasa pemrograman *Ruby on Rails*[22].

H. R Studio

R merupakan bahasa pemrograman statistika yang dapat digunakan untuk analisis dan manipulasi data statistika (pemodelan statistika), dan grafik. R diciptakan (*was created*) oleh Ross Ihaka dan Robert Gentleman (nama depan sama-sama diawali dengan huruf R & R, **Ross** dan **Robert**) dari departemen statistika, di Universitas Auckland,

New Zealand. Saat ini R dikembangkan oleh *R Development Core Team*. Munculnya R terinspirasi oleh bahasa statistika dengan nama S (*statistical language S*). Bahasa statistika S diciptakan oleh John Chambers dan rekan, di *Bell Labs* atau *AT&T Bell Laboratories*. Selain bahasa statistika S, bahasa pemrograman C juga dikembangkan di *AT&T Bell Laboratories* [23].

Selain gratis ada banyak kelebihan R yang lain yaitu :

- 1) R tersedia untuk berbagai sistem operasi selain Windows, seperti Macintosh, Linux dan UNIX.
- 2) Memiliki kemampuan membuat grafik yang canggih.
- 3) Sintaksnya mudah dipelajari dengan banyak fungsi-fungsi statistik yang terpasang.
- 4) R dapat dengan mudah diperluas dengan menciptakan fungsi-fungsi buatan pengguna sendiri disamping juga tersedia banyak program-program kecil sebagai tambahan (*add in*), yang disebut *package* yang juga dapat diunduh gratis.
- 5) R merupakan bahasa pemrograman komputer, sehingga bagi pemrogram menjadi lebih akrab, sedangkan bagi pemakai awal akan merupakan langkah yang mudah untuk memulai sebagai pemrogram komputer.

I. Lexicon Analysis

Lexicon Analysis adalah sebuah proses yang mendahului parsing sebuah rangkaian karakter. Ia menerima masukan serangkaian karakter (seperti dalam dokumen *plain-text* atau *source code*) dan menghasilkan deretan simbol yang masing-masing dinamakan token. Proses parsing akan lebih mudah dilakukan bila inputnya sudah berupa *token*[24].

Analisis *lexicon* terdiri dari dua tahap. Tahap pertama adalah pemindaian (*scanning*), scanner biasanya dibuat berdasarkan prinsip *Finite State Machine* ("mesin dengan jumlah keadaan terbatas"). Pada tahap ini, scanner akan membaca input karakter-ke-karakter, mengubah keadaannya sendiri berdasarkan karakter yang tengah dibaca. Setiap kondisi final (input dianggap valid) akan dicatat, bersama dengan lokasi input.

Lexicon menggunakan *dictionary* atau kamus *lexicon* untuk melakukan penilaian terhadap kata. Pada *dictionary*, kata-kata dipasangkan dengan nilai polaritasnya. Yang harus dilakukan sebelum melakukan analisis menggunakan *lexicon* adalah menentukan kata yang akan dianalisis. Pemilihan kata tersebut dapat dilakukan dengan melakukan *Part-Of-Speech Tagging* dan lalu mencari kata-kata dengan tipe yang ingin diambil, seperti misalnya *adjective* dan *adverb*.

Makna kata dapat berubah bergantung pada konteks dari kalimat, oleh karena itu, pendekatan *lexicon* terkadang tidak bisa menangkap makna sebenarnya dari kata yang diprosesnya. Akan tetapi, *lexicon* memiliki performa klasifikasi yang baik pada kasus lintas domain, dan *knowledge* dapat ditambahkan kapan saja kedalam *dictionary*[25].

Dibandingkan dengan metode lain yang sering dipakai dalam *text mining* seperti *K-NN*, *SVM*, *Hybrid Approach* dan lain-lain yang membutuhkan data latih untuk membandingkan kedekatan dan keakuratan dengan data uji, maka metode *lexicon* tidak memerlukan data pelatihan. Tapi

kelemahannya adalah maksimalnya hasil analisis tergantung pada banyaknya kata yang ada pada kamus *lexicon*.

J. Pilpres 2019

Pemilihan Umum 2019 adalah pemilihan legislatif (DPR RI, DPRD dan DPD) dengan pemilihan presiden yang diadakan secara serentak pada tanggal 17 April 2019 di seluruh wilayah NKRI.

Hal ini dilakukan berdasarkan Putusan Mahkamah Konstitusi Nomor 14 / PUU-11/2013 tentang pemilu serentak, yang bertujuan untuk meminimalkan pembiayaan negara dalam pelaksanaan pemilu, meminimalisir politik biaya tinggi bagi peserta pemilu, serta politik uang yang melibatkan pemilih, penyalahgunaan kekuasaan atau

TABEL I
TAHAPAN PELASANAAN PEMILU 2019

Tanggal	Tahapan
17 Agustus 2017 - 31 Maret 2019	Perencanaan Program dan Anggaran
1 Agustus 2017 - 28 Februari 2019	Penyusunan Peraturan KPU
17 Agustus 2017 - 14 April 2019	Sosialisasi
3 September 2017 - 20 Februari 2018	Pendaftaran dan Verifikasi Peserta Pemilu
19 Februari 2018 - 17 April 2018	Penyelesaian Sengketa Penetapan Partai Politik Peserta Pemilu
9 Januari - 21 Agustus 2019	Pembentukan Badan Penyelenggara
17 Desember 2018 - 18 Maret 2019	Pemutakhiran Data Pemilih dan Penyusunan Daftar Pemilih
17 April 2018 - 17 April 2019	Penyusunan Daftar Pemilih Di Luar Negeri
17 Desember 2017 - 6 April 2018	Penataan dan Penetapan Daerah Pemilihan (Dapil)
26 Maret 2018 - 21 September 2018	Pencalonan Anggota DPR, DPD, DPRD Provinsi dan DPRD Kabupaten / Kota Serta Pencalonan Presiden dan Wakil Presiden
20 September 2018 - 16 November 2018	Penyelesaian Sengketa Penetapan Pencalonan Anggota DPR, DPD dan DPRD Serta Pencalonan Presiden dan Wakil Presiden
24 September - 16 April 2019	Logistik
23 September 2018 - 13 April 2019	Kampanye Calon Angota DPR, DPD dan DPRD Serta Pasangan Calon Presiden dan Wakil Presiden
22 September 2018 - 2 Mei 2019	Laporan dan Audit Dana Kampanye
14 April 2019 - 16 April 2019	Masa Tenang
8 April 2019 - 17 April 2019	Pemungutan dan Perhitungan Suara
18 April 2019 - 22 mei 2019	Rekapitulasi Perhitungan Suara
Jadwal menyusul	Penyelesaian Sengketa Hasil Pemilu DPR, DPD, DPRD, Provinsi dan DPRD Kabupaten / kota
23 Mei 2019 - 15 Juni 2019	Penyelesaian Sengketa Hasil Pemilu Presiden dan Wakil Presiden
Jadwal menyusul	Pentapan Perolehan Kursi dan Calon Terpilih Tanpa Permohonan Perselisihan Hasil Pemilu
Paling lama 3 (tiga) hari setelah penetapan, putusan dismissal atau putusan makamah konstitusi dibacakan	Penetapan Perolehan Kursi dan Calon terpilih Pasca Putusan Mahkamah Konstitusi
Juli - September 2019	Peresmian Keanggotaan
Agustus - Oktober 2019	Pengucapan Sumpah /Janji

Sumber: pemilu2019.kpu.go.id

mencegah politisasi birokrasi, dan merampingkan skema kerja pemerintah. Pemilu serentak akan mempengaruhi komitmen penguatan partai politik dalam koalisi permanen untuk memperkuat basis kekuatan mereka di lembaga-lembaga negara yang tinggi sehingga dengan pemilu serentak diharapkan bisa memfasilitasi pembenahan sistem presidensial di Indonesia[26].

Pada pemilihan umum 17 April 2017 ini diikuti oleh 16 partai politik nasional dan 4 (empat) partai lokal di provinsi Nangroe Aceh Darussalam serta (dua) pasangan calon presiden/wakil presiden yaitu pasangan Joko Widodo/Maaruf Amin dan Prabowo Subianto/Sandiaga Uno[27].

Dari tahapan rencana pelaksanaa pemilu yang ada pada Tabel I, dapat dilihat bahwa proses pengumuman hasil rekapitulasi suara manual akan diumumkan pada tanggal 22 Mei 2019. Hasil penghitungan suara hasil pemilu juga bisa diakses melalui <https://pemilu2019.kpu.go.id> [27].

II. METODE

A. Identifikasi Masalah

Perkembangan medial sosial mengakibatkan kecenderungan para pengguna memberikan suatu tanggapan terhadap apa yang mereka rasakan, lihat dan alami dalam kehidupan sehari. Peristiwa besar baik politik, sosial, ekonomi dan lainnya yang sedang menjadi berita hangat di public akan menjadi *trending* topik di twitter.

Pada tanggal 21 Mei 2019 adalah pengumuman hasil Pemilu dan juga pemenang hasil Pilres 2019. Setelah adanya pengumuman tersebut ada kelompok yang merayakan kemenangan dan juga ada yang melakukan protes terhadap kekalahan yang diterima.

Bagi netizen yang merupakan pengguna twitter yang menjadi bagian dari pendukung kelompok yang menang

TABEL II
CONTOH TOKENIZING

Sebelum <i>Tokenizing</i>	Sesudah <i>Tokenizing</i>
SBY mengucapkan selamat kepada Jokowi-Ma'ruf dan mendukung secara yudisial penuh #MenangTerhormat #JokowiAminSudahMenang https://t.co/NnO0EOA4DZ	sby mengucapkan selamat kepada jokowimaruf dan mendukung secara yudisial penuh

TABEL III
CONTOH NORMALISASI KATA

Sebelum <i>Normalisasi</i>	Sesudah <i>Normalisasi</i>
sby mengucapkan selamat kepada jokowimaruf dan mendukung secara yudisial penuh	sby mengucapkan selamat kepada jokowimaruf dan mendukung secara judisial penuh

TABEL IV
CONTOH FILTERING KATA

Sebelum <i>Filtering</i>	Sesudah <i>filtering</i>
sby mengucapkan selamat kepada jokowimaruf dan mendukung secara yudisial penuh	sby mengucapkan selamat jokowimaruf mendukung judisial penuh

TABEL V
CONTOH HASIL STEMMING

Sebelum <i>Stemming</i>	Sesudah <i>Stemming</i>
sby mengucapkan selamat kepada jokowimaruf dan mendukung secara yudisial penuh	sby ucap selamat jokowimaruf dukung judisial penuh

melampaikan kegembiraan lewat kicauan dan pendukung yang kalah juga melampaikan kekecewaan dalam bentuk protes di twitter. Untuk isu yang sama para netizen meyeritakan “#Hashtag”. Dengan *hashtag* inilah akan terjadi pengelompokan terhadap topik atau isu yang sama di twitter. Ketika *tweet* dengan *hashtag* yang sama ditanggapi dalam jumlah banyak pada periode tertentu maka akan menjadi sebuah trending topik.

B. Pengumpulan Data

Pada penelitian ini data yang dipakai adalah data *tweet* yang disediakan oleh media sosial twitter. Langkah awal yang dilakukan adalah pengecekan trending topik dengan menggunakan *Twitter Search*. Pada *twitter search* juga telah ditampilkan secara langsung 10 (sepuluh) *top trending* terkini. Selanjutnya menentukan *hashtag* yang akan diambil. Pada tahapan ini dilakukan pengecekan kesesuaian antara isi *tweet* dengan topik atau isu yang berkembang dan berhubungan dengan proses pemilu. Penulis melakukan dengan memantau perkembangan *hashtag-hashtag* yang menjadi trending topik tersebut dari tanggal 21 Mei 2019. Apabila *hashtag* tersebut menjadi trending topik dalam dua hari maka akan dilakukan penambahan data. Total *hashtag* yang menjadi trending topik pasca pengumuman hasil ada sebanyak 21 *hashtag*.

Penambahan data dilakukan dengan bahasa pemrograman R, dengan terlebih mendaftarkan akun pada *website dev.twitter.com/rest/tools/console*. Kemudian setelah mengisi syarat dan ketentuan yang berlaku maka didapatkan *access token* atau *key* yaitu *consumer_key*, *consumer_secret*, *access_token* dan *access_secret*.

API keys yang telah didapatkan kemudian ditempelkan pada kode program. Dan pada kode pencarian disertakan *hashtag* trending topik yang akan ditambang. Kemudian proses penambahan data dilakukan, hasil data *crawling* kemudian akan disimpan dalam format *.csv*.

C. Proses Data Tweet

Data yang telah dikumpulkan kemudian akan diproses dengan beberapa tahapan yaitu:

1) Tokenizing

Tahap pembuatan tokenisasi yang dilakukan adalah membuat semua huruf pada *tweet* menjadi huruf kecil, selanjutnya menghapus *mention* “@”, “#” dan URL yang ada pada *tweet*, kemudian proses selanjutnya adalah penghapusan tanda baca dan angka. Pada tahap *tokenizing* akan memotong dokumen menjadi potongan-potongan kecil yang disebut *token*[28].

- a. Langkah-langkah *Tokenizing* adalah:
- b. Tiap *file text* akan dibaca sebagai satu *tweet*.
- c. *Token* pada kalimat *tweet* akan dipisah satu sama lain dengan spasi.
- d. Kalimat *tweet* akan disimpan dari unsur *token*.

Tweet contoh hasil *tokenizing* dapat dilihat pada Tabel II.

2) Normalisasi Kata

Tahap normalisasi kata dilakukan untuk mengganti kata yang tidak baku menjadi baku. Pada langkah ini akan dibaca terlebih dahulu *tweet* hasil *tokenizing* kemudian akan diubah kata yang tidak baku menjadi baku contoh normalisasi kata dapat dilihat pada Tabel III.

3) Filtering

Merupakan tahapan penghapusan *stopwords* pada data *tweet*. Penghapusan *stopwords* dilakukan untuk menghilangkan kata-kata yang jika dihilangkan, data *tweet* masih memiliki makna. Pada Tabel IV adalah contoh proses *filtering* data *tweet*.

4) Stemming

Proses *stemming* dilakukan untuk mendapatkan kata-kata penting dengan menghapus *prefix* (awalan), *suffix* (akhiran), serta *confix* (kombinasi dari awalan dan akhiran). Pada langkah ini adalah dengan terlebih dahulu membaca *tweet* hasil *filtering* kemudian memeriksa kata jika ada pada kamus kata dasar maka kata tersebut adalah kata dasar. Jika tidak ada maka menghapus awalan atau akhiran yang ada pada kata. Untuk hasil *stemming* dapat dilihat pada Tabel V.

D. Analisis

Setelah data *tweet* melewati tahap pembersihan data, kemudian data tersebut akan disimpan sebagai data bersih atau data *cleaning*. Selanjutnya data yang telah melewati *preprocessing* atau data *cleaning* tersebut akan dimuat kembali. Kemudian pada program juga dimuat kamus kata positif dan negatif.

Melalui serangkaian algoritma dilakukan proses pengujian terhadap setiap *tweet*. Akan dihitung jumlah kata positif dan negatif yang ada pada tiap *tweet*. Kemudian akan dilakukan *scoring* data dengan menambahkan kolom negatif dan kolom positif dan dibuat skor untuk tiap *tweet* dengan melakukan perhitungan jumlah positif dikurang dengan jumlah negatif. Hasilnya akan ditampilkan pada kolom *score*. Selanjutnya akan dibuat tabel persentase kata negatif dan positif pada tiap *tweet*, dan akan dibuat sebuah tabel baru dengan nama *table_final*.

Kemudian pada tahap selanjutnya data secara keseluruhan akan diproses untuk menentukan apakah *tweet* tersebut sentimen positif, negatif atau netral. Dengan melakukan sortir data, hasil pada kolom lebih besar atau sama dengan 1 maka sentimen tersebut adalah positif. Sedangkan *tweet* dengan skor lebih kecil dari 0 adalah sentimen negatif. Pada *tweet* dengan skor sama dengan 0 adalah sentimen netral. Dan data pada hasil ini akan dibuat sebuah tabel dengan nama tabel data akhir yang merupakan data jumlah total *tweet* masing-masing berdasarkan kategori sentimen dalam satu *hashtag* trending topik.

Tahap selanjutnya adalah membuat perhitungan jumlah persentase positif, negatif dan netral dari hasil *hashtag* trending topik. Setelah didapatkan jumlah persentasenya maka akan dijadikan menjadi satu data tabel. Data persentase ini kemudian akan disimpan dalam format file *.csv*. Pada tahap ini juga akan memproses kontribusi kata negatif dan positif. Kata positif dan negatif yang berasal dari kamus kata akan dilihat sebarannya dan ditampilkan dalam *histogram*.

TABEL V
DAFTAR HASHTAG TRENDING TOPIK

No.	Hashtag	Periode	Jumlah Tweet
1	#JokowiAminSudahMenang	21 Mei 2019	8.533
2	#RakyatTolakHasilPilpres	21 Mei 2019	6.321
3	#DiskualifikasiJokowi	22 Mei 2019	4.931
4	#aksi22mei	22 Mei 2019	6.337
5	#WhatsAppDown	22 Mei 2019	16.254
6	#SaveOurDemocracy	23 Mei 2019	6.472
7	#SaveIndonesia	23 Mei 2019	12.791
8	#98jagademokrasi	23 Mei 2019	2.040
9	#HidupTakSeanjingItu	23 Mei 2019	6.982
10	#PrayForIndonesiaDemocracy	23 Mei 2019	4.226
11	#TidakAtasNamaSaya	23 Mei 2019	6.723
12	#IndonesiaNotSyria	23 Mei 2019	9.491
13	#JokowiMundurLah	23 Mei 2019	6.925
14	#JokowiPelanggarHAM	24 Mei 2019	5.645
15	#TangkapDalangKerusuhan	24 Mei 2019	10.487
16	#TangkapPRABOWO	24 Mei 2019	6.723
17	#PolriKorbanFitnah	25 Mei 2019	2.132
18	#02KasetRusak	26 Mei 2019	3.597
19	#PrabowoJanganTemuiJokowi	26 Mei 2019	6.070
20	#PrabowoBukanPemimpin	26 Mei 2019	7.189
21	#ThankYouPOLRI	26 Mei 2019	7.973

```

1 ##### CRAWLING #####
2
3 library(devtools)
4 library(twitter)
5 library(rtweet)
6
7 token <- create_token(
8   consumer_key = "H3qbjaooszj195czwQ9d7FYB",
9   consumer_secret = "k91Zzrdkwot5v8b8wbtHaQnQYdK5CMgv3Qcvxhtw8R789R8",
10  access_token = "112895941020129280-Zyeq16d6POLLA3x00Tuf1DCVPGc1b",
11  access_secret = "m10N6kwnenqAheyv3XA8QNFdUVE91uNZArv1bxu3Kt0x1")
12
13 tweet <- search_tweets(q = "#JokowiAminSudahMenang", n = 205000, include_rts = FALSE)
14
15 colnames(tweet)
16
17 tweet[,c("created_at", "screen_name", "text", "retweet_text")]
18
19 tweet
20
21 trending.21 <- tweet[,c("created_at", "screen_name", "text", "retweet_text")]
22
23 write.csv(trending.21, file = "C:/users/acer/Desktop/akhir/JokowiAminSudahMenang.csv", row.names = TRUE)

```

Gambar 1. Kode token twitter API keys

E. Visualisasi

Hasil jumlah persentase sentimen (positif, negatif dan netral) yang telah disimpan akan ditampilkan dalam bentuk Pie chart pada pie chart akan ditampilkan data per-hashtag.

Sedangkan untuk data jumlah sentimen berdasarkan data tweet masing-masing hashtag kemudian digabung dalam satu tabel. Data akan dibentuk menjadi satu dataframe. Dataframe tersebut berisi variabel “hashtag”, “sentimen”, dan “jumlah_tweet”. Kemudian data dari dataframe tersebut akan ditampilkan dalam satu diagram ggplot (barplot).

Kemudian untuk tampilan front end aplikasi sentimen ini akan dibangun sebuah program dengan dukungan library shiny pada R Studio. Dengan dukungan shiny program akan bisa ditampilkan dalam browser.

Pada tampilan ini ada jendela panel field select menu. Menu yang ada pada panel itu adalah nama-nama hashtag yang menjadi trending topik. Kemudian dengan memilih menu pada panel akan ditampilkan grafik barplot (frekuensi sentimen positif, negatif dan netral) pada halaman utama.

III. HASIL DAN PEMBAHASAN

A. Pengumpulan data Twitter

Pengambilan data yang dilakukan penulis adalah mengambil trending topik twitter yang berhubungan dengan situasi paska pengumuman hasil Pilpres oleh KPU RI pada tanggal 21 Mei 2019. Ada 21 (dua puluh satu) hashtag yang menjadi trending topik twitter paska pengumuman Pilpres.

Pengumpulan data tweet dilakukan dengan menjalankan program pada R Studio. Dengan terlebih dahulu membuat akun pada developer twitter, kemudian mendapatkan kode API atau API Keys twitter. Selanjutnya kode tersebut akan ditempelkan pada kode program yang telah di buat dengan bahasa R. Data yang diambil kemudian akan disimpan dalam bentuk file dengan ekstensi .csv. Pada Gambar 1 merupakan tampilan kode program untuk melakukan proses penambangan data dengan hashtag #JokowiAminSudahMenang dengan batas maksimum data yang akan didownload 205.000 tweet.

Pada penelitian ini terdapat sebanyak 21 (dua puluh satu) hashtag yang berhasil ditambang dan untuk rincian jumlah tweet dapat dilihat pada Tabel V.

Dari hasil penambangan sebanyak dua puluh satu hashtag trending topik dapat dilihat bahwa jumlah tweet terbanyak terdapat pada hashtag #WhatsAppDown yaitu sebanyak 16.254. Sedangkan jumlah tweet paling sedikit yaitu terdapat pada hashtag #98jagademokrasi yaitu sebanyak 2.040 tweet.

Hashtag #WhatsAppDown merupakan reaksi netizen twitter atas adanya pembatasan hak akses terhadap media sosial yang dilakukan oleh kementerian Komunikasi dan Informasi. Sedangkan #98jagademokrasi adalah reaksi netizen dimedia sosial atas sikap angkatan 98 (mahasiswa demonstran 1998 yang menjatuhkan Orde Baru) untuk mengamankan demokrasi.

Pada Gambar 2 dapat dilihat hashtag trending topik yang telah berhasil di download dan disimpan pada sebuah folder penyimpanan dan dalam bentuk file .CSV.

B. Preprocessing Data

Sebelum melakukan preprocessing data, pada coding program R terlebih dahulu melakukan pengaktifan library (tm). Setelah dijalankan, program akan melakukan tokenizing untuk pembersihan data. Pada Gambar 3, dapat dilihat pada perintah script baris 10-39 adalah coding atau perintah untuk tahap tokenizing data twitter. Pada koding terdapat perintah untuk menghilangkan spasi, tanda baca, “#”, “@” dan lainnya.

Pada Gambar 4, baris 49 adalah proses stopwords untuk penghapusan tanda baca. Pada baris 52-54 adalah untuk normalisasi kata dengan dictionary slangword. Sedangkan pada baris 66-78 adalah proses stemming untuk kata-kata yang berimbuhan.

Pada Tabel VI, adalah hasil preprocessing yang dilakukan pada salah satu tweet yang ada pada hashtag #JokowiAminSudahMenang. Pada tahap preprocessing ini terlihat bagaimana perubahan konten atau isi pada tweet yang melewati preprocessing.

02KasetRusak	6/11/2019 3:42 PM	Microsoft Excel C...	669 KB
98jagademokrasi	6/11/2019 3:50 PM	Microsoft Excel C...	255 KB
aksi22mei	5/26/2019 5:42 PM	Microsoft Excel C...	1,402 KB
DiskualifikasiJokowi	6/11/2019 4:26 PM	Microsoft Excel C...	684 KB
HidupTakSeanjingtu	6/11/2019 4:31 PM	Microsoft Excel C...	735 KB
IndonesiaNotSyria	6/11/2019 4:34 PM	Microsoft Excel C...	1,004 KB
JokowiAminSudahMenang	6/11/2019 4:37 PM	Microsoft Excel C...	1,429 KB
JokowiMunduriah	6/11/2019 4:41 PM	Microsoft Excel C...	988 KB
JokowiPelanggarHAM	6/11/2019 4:44 PM	Microsoft Excel C...	937 KB
PolriKorbanFitnah	6/11/2019 4:46 PM	Microsoft Excel C...	266 KB
PrabowoBukanPemimpin	6/11/2019 4:48 PM	Microsoft Excel C...	1,475 KB
PrabowoJanganTemuiJokowi	6/11/2019 4:51 PM	Microsoft Excel C...	960 KB
PrayForIndonesiaDemocracy	6/11/2019 4:54 PM	Microsoft Excel C...	787 KB
RakyatTolakHasilPilpres	6/11/2019 4:57 PM	Microsoft Excel C...	853 KB
saveindonesia	6/11/2019 4:59 PM	Microsoft Excel C...	2,273 KB
SaveOurDemocracy	6/11/2019 5:03 PM	Microsoft Excel C...	1,147 KB
TangkapDalangKerusuhan	6/11/2019 5:05 PM	Microsoft Excel C...	1,363 KB
TangkapPRABOWO	6/11/2019 5:08 PM	Microsoft Excel C...	1,165 KB
ThankYouPOLRI	6/11/2019 5:11 PM	Microsoft Excel C...	1,016 KB
TidakAtasNamaSaya	6/11/2019 5:13 PM	Microsoft Excel C...	1,165 KB
WhatsAppDown	6/11/2019 5:16 PM	Microsoft Excel C...	2,076 KB

Gambar 2. Hashtag-hashtag yang telah ditambang

```

10 data.cleaning=data.cleaningtext
11 data.cleaning.text=corpus(vectorSource(data.cleaning))
12
13 ##### Cleaning data #####
14
15 removeURL <- function(x) gsub("http[^\s:]*", "", x)
16 clean <- tm_map(data.cleaning.text, removeURL)
17
18 removeNL <- function(y) gsub("\n", " ", y)
19 clean <- tm_map(clean, removeNL)
20
21 removepipe <- function(z) gsub("<[>+]", "", z)
22 clean <- tm_map(clean, removepipe)
23
24 remove.mention <- function(z) gsub("@\|s+", "", z)
25 clean <- tm_map(clean, remove.mention)
26
27 remove.hashtag <- function(z) gsub("#\|s+", "", z)
28 clean <- tm_map(clean, remove.hashtag)
29
30 removeamp <- function(y) gsub("&"; "", y)
31 clean <- tm_map(clean, removeamp)
32
33 removetitik3 <- function(y) gsub("[:punct:]]", "", y)
34 clean <- tm_map(clean, removetitik3)
35
36 remove.all <- function(xy) gsub("[^\s:alpha:]][:space:]]*", "", xy)
37 clean <- tm_map(clean,remove.all)
38
39 clean <- tm_map(clean, tolower)
40

```

Gambar 3. Coding untuk tokenizing

TABEL VI
HASIL PREPROCESSING

Tahapan Proses	Sebelum	Sesudah
<i>Tokenizing</i>	Presiden Turki mengucapkan selamat ats keberhasilan Pemilu Indonesia dan yg terpilihnya Jokowi Maruf.	presiden turki mengucapkan selamat ats keberhasilan pemilu indonesia dan yg terpilihnya jokowi maruf
<i>Normalisasi kata</i>	presiden turki mengucapkan selamat ats keberhasilan pemilu indonesia dan yg terpilihnya jokowi maruf	presiden turki mengucapkan selamat atas keberhasilan pemilu indonesia dan yang terpilihnya jokowi maruf
<i>Filtering</i>	presiden turki mengucapkan selamat atas keberhasilan pemilu indonesia dan yang terpilihnya jokowi maruf	presiden turki mengucapkan selamat atas keberhasilan pemilu indonesia terpilihnya jokowi maruf
<i>Stemming</i>	presiden turki mengucapkan selamat keberhasilan pemilu indonesia terpilihnya jokowi maruf	presiden turki ucap selamat hasil pemilu indonesia terpilih jokowi maruf

```

45 ##### Load stopword-ID #####
46 stopwordsID <- "C:/Users/acer/Desktop/akhir/words/ID-stopwords.txt"
47
48 ##### Membaca stopwordsID perbaris #####
49 cstopwordID<-readLines(stopwordID);
50
51 ##### Load slangword #####
52 slang <- read.csv("C:/Users/acer/Desktop/akhir/words/Slangword.csv", header=T)
53 old_slang <- as.character(slang$old)
54 new_slang <- as.character(slang$new)
55
56 ##### Load stemming #####
57 stemm <- read.csv("C:/Users/acer/Desktop/akhir/words/Stemming.csv", header=T)
58 old_stemm <- as.character(stemm$old)
59 new_stemm <- as.character(stemm$new)
60
61 ##### Load Lemmatization #####
62 Lemma <- read.csv("C:/Users/acer/Desktop/akhir/words/Lemmatization.csv", header=T)
63 old_lemma <- as.character(Lemma$old)
64 new_lemma <- as.character(Lemma$new)
65
66 stemword <- function(x) Reduce(function(x,r) gsub(stemm$old[r],stemm$new[r],x,fixed=T),
67                               seq_len(nrow(stemm)),x)
68 clean <- tm_map(clean,stemword)
69
70 slangword <- function(x) Reduce(function(x,r) gsub(slang$old[r],slang$new[r],x,fixed=T),
71                               seq_len(nrow(slang)),x)
72 clean <- tm_map(clean,slangword)
73
74 lemmatization <- function(x) Reduce(function(x,r) gsub(Lemma$old[r],Lemmanew[r],x,fixed=T),
75                               seq_len(nrow(Lemma)),x)
76 clean <- tm_map(clean,lemmatization)
77 clean <- tm_map(clean, removewords, cstopwordID)
78 writeLines(strwrap(clean[[2]]$content, 100))
79

```

Gambar 4. Coding untuk stopwords, stemming dan lemmatization

```

20 ##### KATA POSITIF #####
21
22
23 kata.positif = scan("C:/Users/acer/Desktop/akhir/words/positif.txt", what = 'character', comment.char = ';')
24 kata.negatif = scan("C:/Users/acer/Desktop/akhir/words/negatif.txt", what = 'character', comment.char = ';')
25

```

Gambar 5. Memuat kamus kata positif dan negatif

TABEL VII
KATA POSITIF DAN NEGATIF

Kata Positif	Kata Negatif
damai, indah, tolong, kasih, adil, menang, persatuan, tuhan, aman, tenang, kasihan, suka, nkri, benar, percaya, selamat, sadar, jaga, doa, maaf, belajar, bantu, terima, jaga, kondusif, mohon, berdo'a, bersatu, cinta, amin, bagus, jujur, pahlawan, enak, untung, lancar, kuat, sehat	demo, rusuh, korban, jihad, down salah, people, power, perusuh, batu mati, provokator, anarkis, ricuh, hoax, peluru, senjata, demonstran, akibat, turun, melawan, perang, petasan, curang, api, sedih, dalam, bela, meniggal, ditutup, keras, unjuk, provokasi, rezim, malu

```

27 ##### FUNGSI ANALISIS LEXICON #####
28
29 score.sentiment = function(sentences, kata.positif, kata.negatif, ..progress='none')
30 {
31   require(plyr)
32   require(stringr)
33   list=lapply(sentences, function(sentence, kata.positif, kata.negatif)
34   {
35     sentence=gsub("[:punct:]]", "", sentence)
36     sentence=gsub("[^\s:alpha:]][:space:]]", "", sentence)
37     sentence=gsub("\d", "", sentence)
38     sentence=gsub("\n", "", sentence)
39
40     sentence=tolower(sentence)
41     word.list = str_split(sentence, "\\s+")
42     words=unlist(word.list)
43     pos.matches = match(words, kata.positif)
44     neg.matches = match(words, kata.negatif)
45     pos.matches = !is.na(pos.matches)
46     neg.matches = !is.na(neg.matches)
47     pp = sum(pos.matches)
48     nn = sum(neg.matches)
49     score = sum(pos.matches) - sum(neg.matches)
50     list1 = c(score, pp, nn)
51     return(list1)
52   }, kata.positif, kata.negatif)
53
54   score_new = lapply(list1, "[[", 1)
55   ppi = lapply(list1, "[[", 2)
56   nni = lapply(list1, "[[", 3)
57
58   scores.df = data.frame(score = score_new, text=sentences)
59   positif.df = data.frame(positif = ppi, text=sentences)
60   negatif.df = data.frame(negatif = nni, text=sentences)
61
62   list_df = list(scores.df, positif.df, negatif.df)
63   return(list_df)
64 }

```

Gambar 6. Coding fungsi *lexicon analysis*

C. Kamus kata

Pada tahap berikut adalah memuat *file* kata positif dan kata negatif ke dalam program dan kemudian di bentuk menjadi objek data (dapat dilihat pada Gambar 5). *File* kata positif dan negatif yang dimuat masing-masing berisi kata-kata negatif dan positif pada masing-masing *file*. Kamus kata ini adalah *file* dalam format *file* .txt. disimpan dalam satu folder dengan program R.

Pada saat menjalankan program, kamus kata ini pada perintah program (pada Gambar 5) akan dipanggil dengan perintah scan dan menempatkan direktori dimana kamus

kata disimpan. Sebagian kata positif dan negatif tertera pada Tabel VII.

D. Fungsi Analisis Lexicon

Fungsi analisis adalah fungsi yang berupa kode program yang di buat dalam bahasa R. Fungsi ini berisi perhitungan yang telah dibuat sedemikian rupa dengan metode *lexicon*. Fungsi kemudian akan dipanggil untuk melakukan *test* pada data *clean* atau yang sudah bersih. Pada Gambar 6, adalah perintah pada fungsi *lexicon*.

E. Testing

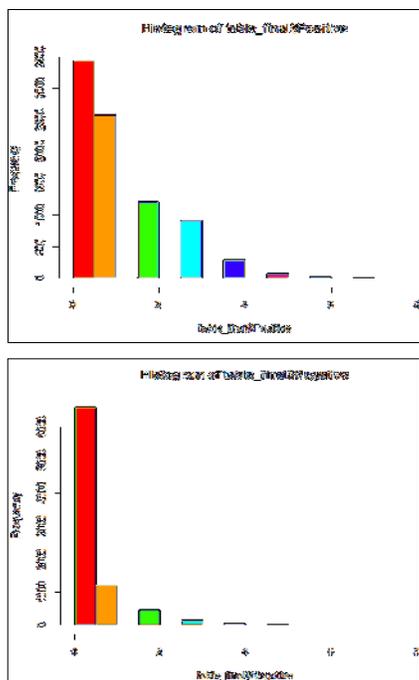
Setelah fungsi *lexicon* dijalankan maka selanjutnya melakukan pengujian pada data trending topik dengan menguji setiap data *tweet*. Gambar 7 baris ke 68-74 adalah perintah untuk membersihkan *tweet* dan menggabungkan *data frame*. Baris 78-80 membuat tiga *data frame* yang berbeda yaitu untuk skor, positif dan negatif, dan menghapus kolom teks dari *data frame*.

```

66- ##### TEST #####
67 result = score.sentiment(sample, kata.positif, kata.negatif)
68
69 library(reshape)
70
71
72 test1 = result[[1]]
73 test2 = result[[2]]
74 test3 = result[[3]]
75
76 head(result)
77
78 test1$text = NULL
79 test2$text = NULL
80 test3$text = NULL
81
82 q1 = test1[[1,]]
83 q2 = test2[[2,]]
84 q3 = test3[[3,]]
85
86 qq1 = melt(q1, , var="score")
87 qq2 = melt(q2, , var="Positive")
88 qq3 = melt(q3, , var="Negative")
89
90 qq1["score"] = NULL
91 qq2["Positive"] = NULL
92 qq3["Negative"] = NULL
93
94 table1 = data.frame(text=result[[1]]$text, score=qq1)
95 table2 = data.frame(text=result[[2]]$text, score=qq2)
96 table3 = data.frame(text=result[[3]]$text, score=qq3)
97
98 table.final = data.frame(text = table1$text, positive = table2$score.value, negative = table3$score.value)

```

Gambar 7. Coding test data dengan fungsi *lexicon*



Gambar 8. Histogram kontribusi kata

F. Hasil Pengujian

Pada penelitian ini seluruh data hastag yang ada akan melewati seluruh proses pada tahap-tahap preprosesing dan juga proses data melalui analisis lexicon. Kemudian pada seluruh data sentimen tiap *hashtag* disimpan pada folder ggplot. Kemudian seluruh data tersebut akan dipanggil kembali dan dibuat dan dijadikan objek data. Untuk prosesnya dapat dilihat pada Gambar 9.

Pada Gambar 10 dapat disimpulkan bahwa *hashtag* dengan jumlah *tweet* terbanyak yaitu #WhatsAppDown dengan komposisi sentimen adalah positif sebanyak 2.454, negatif 5.914 dan sentimen netral sebanyak 7.886.

Sedangkan *hashtag* dengan jumlah *tweet* paling sedikit yaitu #98jagademokrasi dengan komposisi sentimen positif adalah sebanyak 367, negatif 211 dan sentimen netral sebanyak 1462.

Dan hasil dari persentase sentimen pada tiap-tiap *hashtag* ditampilkan pada Gambar 11.

```

282- ##### MENGGADUNGKAN DATA (VISUALISASI)#####
283
284 kaset <- read.csv(file = "C:/Users/acer/Desktop/akhir/ggplot/02kasetrusak.csv")
285 kaset
286
287 jaga <- read.csv(file = "C:/Users/acer/Desktop/akhir/ggplot/98jagademokrasi.csv")
288 jaga
289
290 aksi <- read.csv(file = "C:/Users/acer/Desktop/akhir/ggplot/aksi22mei.csv")
291 aksi
292
293 indo <- read.csv(file = "C:/Users/acer/Desktop/akhir/ggplot/IndonesiaNotSyria.csv")
294 indo
295
296 ##### CARUNG BARI :
5518
#####
Console C:/Users/acer/Desktop/akhir/
> kaset
  X Positif Negatif Netral
1 #02kasetRusak 678 1833 1086
>
> jaga <- read.csv(file = "C:/Users/acer/Desktop/akhir/ggplot/98jagademokrasi.csv")
> jaga
  X Positif Negatif Netral
1 #98jagademokrasi 367 211 1462
>
> aksi <- read.csv(file = "C:/Users/acer/Desktop/akhir/ggplot/aksi22mei.csv")
> aksi
  X Positif Negatif Netral
1 #aksi22mei 1288 2893 2156
>
> indo <- read.csv(file = "C:/Users/acer/Desktop/akhir/ggplot/IndonesiaNotSyria.csv")
> indo
  X Positif Negatif Netral
1 #IndonesiaNotSyria 1690 3013 4788
>
> pray <- read.csv(file = "C:/Users/acer/Desktop/akhir/ggplot/PrayForIndonesiaDemocracy.csv")
> pray
  X Positif Negatif Netral

```

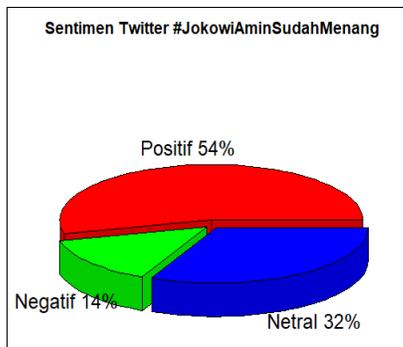
Gambar 9. Memuat kembali data *hashtag*

Hashtag	Positif	Negatif	Netral
1 #02KasetRusak	678	1833	1086
2 #98jagademokrasi	367	211	1462
3 #aksi22mei	1288	2893	2156
4 #IndonesiaNotSyria	1690	3013	4788
5 #PrayForIndonesiaDemocracy	1019	1561	1646
6 #SaveOurDemocracy	1297	2230	2945
7 #JokowiMundurlah	1059	1973	3263
8 #JokowiPelanggarHAM	899	1872	2874
9 #PolriKorbanFitnah	438	758	936
10 #PrabowoBukanPemimpin	1631	3660	1898
11 #PrabowoJanganTemuiJokowi	1310	1298	3462
12 #DiskualifikasiJokowi	606	1123	3004
13 #HidupTakSeanjingitu	773	1757	4452
14 #JokowiAminSudahMenang	4413	1172	2568
15 #RakyatTolakHasilPilpres	1107	1423	3474
16 #saveindonesia	3390	5688	3713
17 #TangkapDalangKerusuhan	1535	3222	5730
18 #TangkapPRABOWO	1188	3151	2384
19 #ThankYouPOLRI	2022	1417	4534
20 #TidakAtasNamaSaya	1188	3151	2384
21 #WhatsAppDown	2454	5914	7886

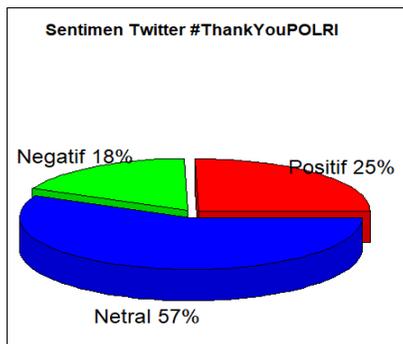
Gambar 10. Data seluruh *hashtag* dan jumlah sentimen

Hashtag	Positif	Negatif	Netral
1 KasetRusak	19	51	30
2 jagademokrasi	18	10	72
3 aksi22mei	20	46	34
4 IndonesiaNotSyria	18	32	50
5 PrayForIndonesiaDemocracy	24	37	39
6 SaveOurDemocracy	20	34	46
7 JokowiMundurlah	17	31	52
8 JokowiPelanggarHAM	16	33	51
9 PolriKorbanFitnah	21	36	44
10 PrabowoBukanPemimpin	23	51	26
11 PrabowoJanganTemuiJokowi	22	21	57
12 DiskualifikasiJokowi	13	24	63
13 HidupTakSeanjingitu	11	25	64
14 JokowiAminSudahMenang	54	14	32
15 RakyatTolakHasilPilpres	18	24	58
16 saveindonesia	27	44	29
17 TangkapDalangKerusuhan	15	31	55
18 TangkapPRABOWO	18	47	35
19 ThankYouPOLRI	25	18	57
20 TidakAtasNamaSaya	18	47	35
21 WhatsAppDown	15	36	49

Gambar 11. Persentase sentimen pada keseluruhan *hashtag*

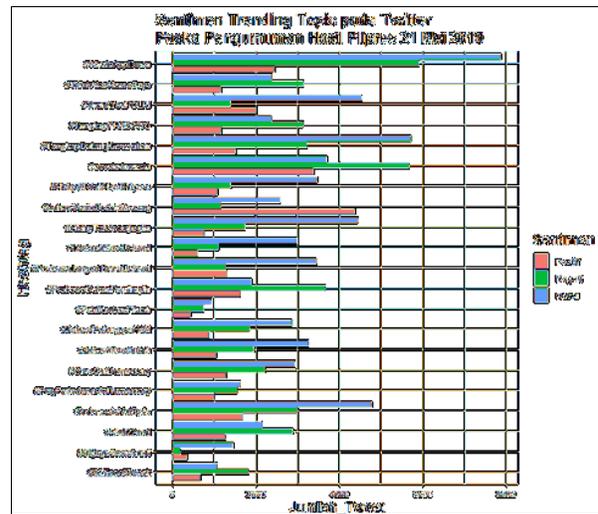


Gambar 12. Pie chart #JokowiAminSudahMenang



Gambar 13. Pie chart #ThankYouPOLRI

Dari hasil pemaparan data pada Gambar 11, #WhatsAppDown, adalah *hashtag* dengan jumlah *tweet* terbanyak memiliki persentase sentimen positif 15%, negatif 36% sedangkan persentase netral sebanyak 49%. Sedangkan pada *hashtag* #98jagademokrasi yang memiliki jumlah *tweet* paling sedikit memiliki persentase positif sebanyak 18%, negatif 10% dan sentimen netral 72%.



Gambar 14. Visualisasi data sentimen twitter

```

1 library(shiny)
2 # Server-side logic
3 # UI-side logic
4 # ...
5 # ...
6 # ...
7 # ...
8 # ...
9 # ...
10 # ...
11 # ...
12 # ...
13 # ...
14 # ...
15 # ...
16 # ...
17 # ...
18 # ...
19 # ...
20 # ...
21 # ...
22 # ...

```

Gambar 15. Coding pada form user interface

```

1 library(shiny)
2
3 shinyServer(
4   function(input, output) {
5
6
7     output$trendPlot <- renderPlot({
8
9
10      barplot(data[,input$trending],
11             main=input$trending,
12             ylab="jumlah Tweet",
13             xlab="sentimen",
14             col = c("red", "green", "blue"))
15             legend("topleft",
16                  c("positif","Negatif","Netral"),
17                  fill = c("red","green","blue"))
18
19
20    })
21
22 )

```

Gambar 16. Coding pada form server

Dilihat dari hasil persentase sentimen positif terbanyak berada pada *hashtag* pada *hashtag* #JokowiAminSudahMenang sebesar 54% dan jumlah persentase sentimen positif terendah berada pada *hashtag* #HidupTakSeanjingItu sebesar 11%.

Pada hasil persentase sentimen negatif terbanyak berada pada *hashtag* #02KasetRusak dan #PrabowoBukanPemimpin yaitu sebanyak 51%. Dan persentase sentimen negatif terbanyak berada pada *hashtag* #98jagademokrasi sebesar 10%.

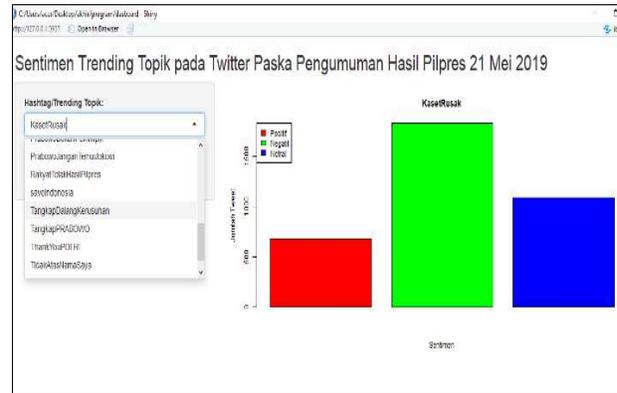
Sedangkan untuk sentimen netral terbanyak yaitu pada *hashtag* #98jagademokrasi dan untuk sentimen netral terendah berada pada *hashtag* #PrabowoBukanPemimpin dengan jumlah sentimen netral 26%.

G. Visualisasi Persentase pada Pie chart

Pada Gambar 12, *Hashtag* #JokowiAminSudahMenang muncul pada tanggal 21 Mei 2019 sebagai reaksi netizen atas kemenangan pasangan nomor urut 01 Jokowi-Ma'ruf Amin. Terdapat 8.533 tweet dengan persentase sentimen positif 54%, negatif 14% dan sentimen netral 32%.

Sedangkan pada Gambar 13, sebagai apresiasi dan terimakasih terhadap kinerja TNI dan Polri dalam menjaga keamanan, terjadi aksi pembagian bunga yang dilakukan oleh berbagai komunitas di depan gedung BAWASLU.

Dan pada twitter tanggal 26 Mei 2019 muncul *hashtag* #ThankYouPOLRI. Terdapat 7973 *tweet* dengan sentimen positif 25%, sentimen negatif 18% dan netral 57%.



Gambar 17. Front End aplikasi sentimen trending topik

H. Visualisasi Data Keseluruhan

Sebelum data sentimen ditampilkan secara keseluruhan, hasil setiap *hashtag* yang telah dianalisis akan dipanggil kembali dengan memanggil data pada tabel gabungan yang ada pada Gambar 14. Pada tampilan *data frame* yang dibuat terdapat data sebanyak 63 baris. Di mana terdapat 21 kategori *hashtag* yang ada pada sumbu y, dan Jumlah *Tweet* sebagai sumbu x serta 3 kategori sentimen menjadi legenda dalam warna.

Pada Gambar 14 dapat dilihat data dengan sentimen tertinggi yaitu sentimen netral pada *hashtag* #WhatsAppDown dengan jumlah *tweet* sebanyak 7.886. Sedangkan sentimen dengan nilai terkecil yaitu sentimen negatif pada *hashtag* #98jagademokrasi dengan jumlah *tweet* 211 *tweet*.

I. Dashboard

Dengan memanfaatkan *library* R *Shiny* yang ada pada R *Studio* maka akan dibuat suatu interface *front End* untuk menampilkan hasil data analisis hasil penelitian. Pada aplikasi ini membutuhkan dua *form* yaitu *user interface* atau UI dan *server* yang dibuat dengan bahasa pemrograman R pada R *Studio*.

Pada Gambar 15, *form user interface* terdapat perintah pada baris 1-3 untuk menyertakan *library* yaitu *shiny*, *ggplot* dan *dplyr*, *shiny* untuk menampilkan aplikasi dalam bentuk web. *Library ggplot2* untuk menampilkan data dalam bentuk grafik *ggplot*. Kemudian *library dplyr* untuk memuat data.

Pada baris 5-6 adalah perintah untuk memuat data dari direktori di mana data disimpan. Baris 9 adalah perintah menjalankan *shiny* untuk menampilkan *dashboard*. Pada baris 11 untuk memuat halaman atau *form*. Pada baris 14 untuk menampilkan judul pada halaman. Kemudian pada baris 20-25 untuk membuat panel yang berisi *select menu* dengan mengambil nilai dari kolom.

Pada baris 28 adalah untuk menampilkan halaman utama. Baris 29 perintah untuk menerima hasil dari eksekusi perintah dengan nama "trendPlot". Isi dari *trendPlot* inilah yang akan berisi data pada grafik yang akan ditampilkan.

Pada Gambar 16, *form server* disertakan *library shiny* pada baris 1. Kemudian baris 3 untuk menjalankan *shiny server*. Pada baris 4 adalah perintah input dan output. Yang mana pada *inputan* akan diterima dari panel *select* dan outputnya untuk menampilkan grafik sesuai dengan data yang ada pada kolom atau *hashtag* yang dipilih.

Pada baris 7 adalah untuk mengambil nilai pada kolom yang berisi nama *hashtag*. Di mana terdapat sebanyak 21

hashtag pada penelitian ini. Sedangkan pada baris 10-17 adalah perintah untuk membuat diagram *ggplot*. Di mana untuk nilai dari diagram diambil dari nilai sentimen trending yaitu positif, negatif dan netral.

Setelah program dijalankan maka kemudian muncul tampilan dalam bentuk *web* dengan menggunakan browser. Kemudian pada bagian sebelah kiri akan ditampilkan panel yang berisi *select menu*. Dengan memilih data yang ada pada *select menu* maka akan bisa dilihat data sentimen trending topik dalam diagram *ggplot*. Untuk tampilan *interface* akan ditampilkan dalam browser web. Setelah program dijalankan progra, dapat dilihat pada Gambar 17.

Setiap tampilan grafik atau *barplot* berasal dari pilihan pada *select menu*. Pada grafik *barplot* untuk sumbu y yaitu jumlah atau frekuensi sentimen. Sedangkan pada sumbu x adalah kategori dari sentimen. Pada legenda juga terdapat tanda kategori sentimen dalam warna.

IV. KESIMPULAN DAN SARAN

A. Kesimpulan

Berdasarkan penelitian yang dilakukan penulis memberikan kesimpulan bahwa Analisis *lexicon* dapat mengklasifikasikan sentimen Media sosial twitter dalam tiga kelas sentimen yaitu positif, negatif dan netral. Dan akurasi terhadap analisis *lexicon* ditentukan oleh jumlah kata yang ada pada kamus kata positif dan kamus kata negatif.

Sentimen positif terbanyak berada pada *hashtag* #JokowiAminSudahMenang sebesar 54% dan terendah berada pada *hashtag* #HidupTakSeanjingItu sebesar 11%. Sedangkan sentimen negatif terbanyak berada pada *hashtag* #02KasetRusak dan #PrabowoBukanPemimpin yaitu sebanyak 51% dan negatif terendah pada *hashtag* #98jagademokrasi sebesar 10%. Untuk sentimen netral tertinggi pada *hashtag* #98jagademokrasi, netral terendah berada pada *hashtag* #PrabowoBukanPemimpin dengan jumlah 26%.

B. Saran

Penelitian ke depannya penulis menyarankan untuk menambah parameter untuk validasi *tweet* untuk mengantisipasi data *tweet* dari pengirim yang sama berulang oleh akun yang sama pula. Dan juga membuat link untuk menghubungkan aplikasi *front end* dengan web selain dari *RPubs*.

V.KUTIPAN

- [1] N. M. S. Hadna, P. I. Santosa, and W. W. Wing, "STUDI LITERATUR TENTANG PERBANDINGAN METODE UNTUK PROSES," *SENTIKA*, vol. 18, no. 1, pp. 57–64, 2016.
- [2] M. Prabowo and I. Irwansyah, "Trending Topics Vs Agenda-Setting: Pengaruh Trending Topics Politik sebagai Reversed Agenda-Setting dan Haluan Politik Pemilik Terhadap Berita Politik di Televisi," *J. Komun. Indones.*, vol. 5, no. 1, pp. 5–15, 2018.
- [3] N. Anusha, "Sentiment Analysis of Twitter Data through Big Data," vol. 6, no. 06, pp. 307–309, 2017.
- [4] C. Juditha, "Fenomena Trending Topic Di Twitter: Analisis Wacana Twit #Savehajilulung," *J. Penelit. Komun. dan Pembang.*, vol. 16, no. 2, p. 138, 2018.
- [5] R. Habibi and D. B. Setyohadi, "ANALISIS SENTIMEN PADA TWITTER MAHASISWA," vol. 12, no. 1, pp. 103–109, 2016.
- [6] Tiara, M. K. Sabariah, and V. Effendy, "Analisis Sentimen pada Twitter untuk Menilai Performansi Program Televisi dengan Kombinasi Metode Lexicon-Based dan Support Vector Machine," *e-Proceeding Eng.*, vol. 2, no. 1 April 2015, pp. 1237–1247, 2015.
- [7] I. I. N. Kusumawati, P. S. Informatika, F. Komunikasi, D. A. N. Informatika, and U. M. Surakarta, "ANALISA SENTIMEN MENGGUNAKAN LEXICON BASED KENAIKAN HARGA ROKOK PADA MEDIA SOSIAL TWITTER," 2017.
- [8] E. Mesak, Y. N. Kunang, and R. Andryani, "ESKPLORASI TRENDING TOPIK TWITTER MENGGUNAKAN TEXT MINING," *SENTIKOM*, vol. 1, no. 1, pp. 1–8, 2017.
- [9] Hartanto, "TEXT MINING DAN SENTIMEN ANALISIS TWITTER PADA GERAKAN LGBT," *J. P. Ilm.*, vol. 9, no. 1, pp. 18–25, 2017.
- [10] J. I. Polinema *et al.*, "IMPLEMENTASI TWITTER SENTIMENT ANALYSIS UNTUK REVIEW FILM MENGGUNAKAN ALGORITMA SUPPORT," pp. 93–100, 2016.
- [11] Z. Pan, S., Ni, X., Sun, J., Yang, Q., & Chen, "Cross-domain sentiment classification via spectral feature alignment. International World Wide Web Conference Committee," pp. 751–760, 2010.
- [12] Y. Tresnawati, *ANALISIS SENTIMEN PADA TWITTER MENGGUNAKAN PENDEKATAN AGGLOMERATIVE HIERARCHICAL CLUSTERING*. 2017.
- [13] B. Liu, "Sentiment Analysis: A Multi-Faceted Problem," *IEEE Intell. Syst.*, vol. In Proceed, pp. 76–80, 2010.
- [14] E. S. Negara, R. Andryani, and P. H. Saksono, "Twitter Data Analytics: Geospatial Data Extraction and Analysis," *J. INKOM*, vol. 10, no. 1, pp. 27–36, 2016.
- [15] W. F. and M. D. Gordon, "The Power of social media analytics," *Commun. ACM*, vol. 57, 2010.
- [16] D. Zheng, H. Chen, R. Lusch, and SHLi, "Social media analytics and intelligence," *Intell. Syst. IEEE*, vol. 25, no. 6, pp. 13–16, 2014.
- [17] and Z. Z. W.Fan, L. Wallace, S.Rich, "Tapping the power of text mining," *Commun. ACM*, vol. 49, no. 9, pp. 76–82, 2006.
- [18] Deliusno, "Pengguna Twitter di Indonesia Capai 50 Juta," *26 Maret*, 2015. [Online]. Available: <https://tekno.kompas.com/read/2015/03/26/16465417/Pengguna.Twitter.di.Indonesia.Capai.50.Juta>.
- [19] Y. Permana and T. Kurnia, "PENGARUH AKTIVITAS MEDIA TWITTER BRI SYARIAH TERHADAP KEPUASAN NASABAH BRI SYARIAH," *J. Nisbah*, vol. 1, no. 1, pp. 38–54, 2015.
- [20] A. Amalina, "PENGARUH SOCIAL MEDIA MARKETING TERHADAP BRAND LOYALTY DENGAN BRAND TRUST SEBAGAI VARIABEL INTERVENING (STUDI PADA FOLLOWER TWITTER MIZONE @MIZONEID)," *J. Ilmu Manaj. UNESA*, vol. 4, no. 3, pp. 1–12, 2016.
- [21] S. N. Edi, R. Andryani, and P. H. Saksono, "Analisis Data Twitter: Ekstraksi dan Analisis Data Geospasial," *INKOM*, vol. 10, no. 1, pp. 56–63, 2016.
- [22] T. de B. M. Neppelenbroek, M. Lossek, R. Janssen, *Twitter an architectural review*. 2011.
- [23] W. Rahayu, "MENINGKATKAN KETERAMPILAN MENGGUNAKAN SOFTWARE R SEBAGAI SOLUSI UNTUK MENINGKATKAN INOVASI PEMBELAJARAN BAGI GURU-GURU MATEMATIKA SMA DAN SMK DI JAKARTA TIMUR," *Sarwahita*, vol. 12, no. 2, pp. 134–140, 2015.
- [24] S. Baccianella, A. Esuli, and F. Sebastiani, "SENTIMENT: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining," vol. 0, pp. 2200–2204, 2008.
- [25] S. U. Masruroh and L. K. Wardhani, "PERBANDINGAN

- [26] KINERJA ALGORITMA NAIVE BAYES DAN K-NN PENDEKATAN LEXICON PADA ANALISIS SENTIMEN DI MEDIA," no. September, pp. 978–979, 2019.
- [27] R. Solihah and U. Padjadjaran, "Peluang dan tantangan pemilu serentak 2019 dalam perspektif politik," vol. 3, no. 1, pp. 73–88, 2019.
- [28] KPU, "Tahapan Pemilu Tahapan pada Pemilu 2019," 2017. [Online]. Available: <https://infopemilu.kpu.go.id/>.
- [29] C. D. Manning, P. Raghavan, and H. Schütze, *An Introduction to Information Retrieval*, (C) 2009. Cambridge: Cambridge UP, 2009.



Adi Pandu Rahmat Nababan Penulis Lahir di Tapanuli Sumatera Utara merupakan mahasiswa Fakultas Teknik Universitas Sam Ratulangi Manado, Jurusan Teknik Elektro Program Studi Teknik Informatika. Memulai studi pada tahun 2017 dan menyelesaikan pada tahun 2019 melalui program studi lanjutan dari Diploma. Pada saat akan menyelesaikan perkuliahan penulis mengambil penelitian dalam bidang text mining dengan judul tugas akhir "Analisis Sentimen Trending Topik pada Twitter Menggunakan Lexicon Analysis". Penelitian yang dilakukan fokus pada sentimen publik pada twitter pasca pengumuman hasil Pilpres pada tanggal 21 Mei 2019 oleh Komisi Pemilihan Umum Republik Indonesia.

Ir. Arie S. M. Lumenta, ST., MT, adalah dosen pada Jurusan Teknik Elektro Universitas Sam Ratulangi Manado. Menyelesaikan studi strata dua teknik elektro dari Institut Teknologi Bandung pada tahun 2004 dan memiliki bidang penelitian pada Computer Architecture, Computer Networks, E-Learning, Database dan Internet of Things. Dan sekarang ini menjabat sebagai Ketua Jurusan Teknik Elektro Universitas Sam Ratulangi.

Yaulie Deo Y. Rindengan, ST, MSc, MM adalah dosen di Fakultas Teknik Universitas Sam Ratulangi Manado Jurusan Teknik Elektro. Menyelesaikan strata dua dari Institut Pertanian Bogor pada tahun 2008 dan memiliki bidang keahlian dan penelitian pada teknologi informasi khususnya Decision Support Systems, e-Government dan Smart City. Dan sekarang ini selain sebagai dosen juga aktif dalam berbagai seminar nasional di bidang IT.