Analisis Sentimen Twitter Debat Calon Presiden Indonesia Menggunakan Metode Fined-Grained Sentiment Analysis

¹Septian Fendyputra Pratama, ²Ricsa Andrean, ³Aryo Nugroho
¹ septian.fendyputra@fik.narotama.ac.id, ²ricsa.andrean@fik.narotama.ac.id, ³aryo.nugroho@narotama.ac.id

Program Studi Sistem Informasi, Fakultas Komputer, Universitas Narotama Surabaya

Abstract — Social media, Twitter, these days has a great impact in building public opinion, views, sentiments, and political preferences (ahead of the general election). This research is to find out the conversations on Twitter at the First Candidate Debate of the President of the Republic of Indonesia through the hashtag of the two pairs of candidates. Besides that, also to find out about the tendency of people on Twitter related to the ongoing debate tends to be positive, negative, or neutral. Conversation data on Twitter is obtained through the Twitter API taken with the R programming language and uses the Fine-grained Sentiment Analysis method. If one tweet contains more positive than negative sentences, then the overall result will be positive and worth (+1). If the number of negative sentences is greater than the positive sentence, then the overall result is negative and is worth (-1). If there is an equal number of positive and negative sentences in the input paragraph, then the result is neutral and worth (0). The results of this study indicate that sentiment tweets from both hashtags tend to be positive, more than negative and neutral sentiments.

Intisari — Media sosial, Twitter, saat ini telah banyak memberikan dampak besar dalam membangun opini, pandangan, sentimen, dan preferensi politik publik (menjelang Pemilihan Umum) berlangsung. Penelitian ini dilakukan untuk mengetahui percakapan di Twitter pada debat pertama calon presiden Republik Indonesia melalui hashtag dari kedua pasang calon. Selain itu, juga untuk mengetahui tentang kecenderungan masyarakat di Twitter terkait dengan debat vang sedang berlangsung tersebut cenderung positif, negatif, atau netral. Data percakapan di Twitter didapatkan melalui Twitter API yang diambil dengan bahasa Pemrograman R. Proses analisis sentimen ini menggunakan metode Fined-grained Sentiment Analysis yaitu, Jika satu tweet berisi lebih banyak kalimat positif daripada negatif, maka hasil keseluruhan akan positif dan bernilai (+1). Jika jumlah kalimat negatif lebih besar dari kalimat positif, maka hasil keseluruhan negatif dan bernilai (-1). Jika ada jumlah yang sama dari kalimat positif dan negatif dalam paragraf, maka hasilnya adalah netral dan bernilai (0). Hasil dari penelitian ini

menunjukkan bahwa tweet sentimen dari kedua hashtag cenderung positif, lebih banyak dari pada sentimen negatif dan netral.

Kata Kunci— Debat Calon Presiden , Twitter, Sentiment Analysis, Twitter API , Pemrograman R.

I. PENDAHULUAN

Saat ini penggunaan media sosial adalah terutama *Twitter* di Indonesia telah banyak memberikan dampak besar dalam membangun opini, pandangan, sentimen, dan preferensi politik publik terutama menjelang pemilihan umum berlangsung. Pada saat tahapan pemilihan umum seperti debat [1] kandidat, para pengguna media sosial atau para pendukung masing-masing kandidat juga memberikan opini mereka lewat media sosial tentang jalannya debat menggunakan hashtag tertentu pada *tweet* yang ada di *Twitter* [2] [3].

Untuk mengetahui dan menentukan kecenderungan pengguna *Twitter* dalam memposting *tweet*, perlu untuk melakukan Analisis Sentimen [4] [5]. Dalam konteks di media sosial, *Sentiment Analysis* adalah bagaimana melakukan analisa terhadap orang yang mengekspresikan opini-opini mereka pada berbagai topik di media sosial [6]. Analisis Sentimen telah banyak dilakukan seperti mengetahui respons opini konsumen pada suatu produk, preferensi politik, pendapat tentang film yang telah dirilis beserta prediksi pendapatan film dan lainnya [7] [8] [9].

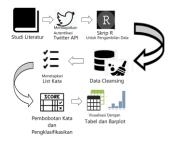
Penelitian terdahulu tentang Analisis Sentimen untuk bidang politik yang menganalisis sifat dan karakteristik wacana politik yang terjadi di Twitter selama pemilihan Presiden Amerika pada November 2016 [10]. Penelitian selanjutnya tentang Analisis Sentimen untuk bidang musik yang memprediksi keberhasilan album musik, berdasarkan komentar yang diposting di jejaring sosial selama 30 hari sebelum rilis album melalui Twitter untuk mengumpulkan komentar pengguna [11].

Analisis Sentimen ini untuk mengetahui *tweet* yang diposting di lini massa *Twitter* mengandung kata positif, negatif, atau cenderung netral [4] [12]. Penerapan Analisis sentimen ini menggunakan bantuan Bahasa pemrograman

Komputasi Statistik yaitu R [13] [14]. Sebelum melakukan Analisis sentimen, harus mempunyai Twitter API sebagai sumber pengambilan data untuk bahan analisis sentimen tweet pada lini masa Twitter [15]. Dan dilanjutkan pada menjalankan Skrip pemrograman R untuk proses pengambilan data tweet di lini masa Twitter yang terkait dengan jalannya debat kandidat [16]. Setelah mendapatkan data terkait Debat [1] Pertama Kandidat Presiden dan wakil presiden, Penelitian akan mulai melakukan Analisis Sentimen terhadap pada

II. METODE PENELITIAN

- 1. Studi literatur, yaitu mempelajari Bahasa Pemrograman *R* dan penggunaannya untuk Analisis Sentimen di Media sosial *Twitter*. Literatur yang digunakan sebagai referensi adalah dari buku *The Art of R Programming* oleh Norman Matloff dan *Mastering Social Media Mining with R* yang memberikan informasi penggunaan *R* untuk Melakukan Analisis Sentimen di *Twitter*.
- 2. Mendapatkan *Twitter API* untuk bisa mengakses Data di Twitter melalui akun *Twitter* pribadi.
- 3. Menuliskan Skrip Pemrograman *R* untuk mengambil data percakapan di *Twitter* yang berkaitan dengan masingmasing Calon yang mengikuti Debat.
- 4. Melakukan Proses *Data Cleansing* untuk menghapus seperti menghapus beberapa bagian yang tidak diperlukan seperti Kata, tanda baca, *link url*, huruf, serta *emoji*, kemudian mengubah semua huruf besar menjadi huruf kecil (*tolower*) atau sebaliknya (*toupper*). Setelah proses ini selesai, simpan hasil *Data Cleansing*.
- 5. Mengambil data List of Opinion Words (positive/negative) in Bahasa Indonesia for Sentiment Analysis Originated by Liu's Opinion Words list with modification/translation to Indonesia di GitHub yang berisi masing-masing kata positif dan negatif dalam Bahasa Inggris yang telah di terjemahkan ke dalam bahasa Indonesia, untuk memudahkan proses Analisis Sentimen.
- 6. Menuliskan Skrip Pemrograman *R* Untuk melakukan Pembobotan Kata untuk menentukan *tweet* tersebut termasuk Sentimen Positif, Negatif, atau Netral.
- 7. Hasil akhir berupa Tabel yang disajikan di Pemrograman *R* berisi tiap *tweet* beserta label Positif, Negatif, atau Netral pada *masing-masing Hashtag* dari Pasangan Calon dan *Barplot* atau Diagram Batang yang menyajikan tentang pengklasifikasian positif, netral, dan negatif pada *tweet* masing-masing Pasangan Calon.



Gambar 1. Alur Penelitian

Hal pertama yang dilakukan adalah mendapatkan Autentikasi *Twitter API* terlebih dahulu melalui akun *Twitter* pribadi, untuk skrip pemrograman *R* dapat mengakses *token* dan mendapatkan sumber data *tweet* di lini masa *Twitter* [15]. Untuk pengambilan data kedua pasangan calon dilakukan di antara pukul 18.30 sampai 23.59. Setelah data dari Twitter API telah diambil, akan dilakukan Data Cleansing seperti *menghapus beberapa bagian yang tidak diperlukan seperti* Kata, tanda baca, link url, huruf, serta emoji, kemudian juga mengubah semua huruf menjadi kecil menjadi kecil (tolower) atau sebaliknya (toupper). Setelah Proses ini Selesai, simpan hasil Data Cleansing [17].

Untuk mencocokkan kata Positif, Negatif, atau Netral pada tiap tweet yaitu menggunakan List of Opinion Words (positive/negative) in Bahasa Indonesia for Sentiment Analysis Originated by Liu's Opinion Words list with modification/translation to Indonesia yang berisi masingmasing kata positif dan negatif dalam Bahasa Inggris yang telah di terjemahkan ke dalam bahasa Indonesia [18] [19].

Penelitian ini dimulai dengan menjalankan Skrip bahasa Pemrograman R untuk melakukan Analisis Sentimen dengan metode Fine-grained Sentiment Analysis, yaitu mengklasifikasikan tweet-tweet dalam debat termasuk dalam tweet tersebut adalah Positif, Negatif, atau Netral [20] [21]. Untuk melakukan Fine-grained Sentiment Analysis, langkah selanjutnya adalah menjalankan Skrip untuk fungsi di Pemrograman bahasa R untuk melakukan Text Scoring (Pembobotan Kata). Pembobotan kata dilakukan Jika satu tweet berisi lebih banyak kalimat positif daripada negatif, maka hasil keseluruhan akan positif dan bernilai (+1). Jika jumlah kalimat negatif lebih besar dari kalimat positif, maka hasil keseluruhan negatif dan bernilai (-1). Jika ada jumlah yang sama dari kalimat positif dan negatif dalam paragraf input, maka hasilnya adalah netral dan bernilai (0) [22]. Sebagai Contoh adalah:

- Saya benci Bau Parfum. (negatif) (-1)
- Kampus saya sangat indah sekali. (positif) (+1)
- Saya tidak Cinta. Kalimat yang memiliki makna ganda atau memuat kalimat positif maupun negatif sehingga kata tersebut bernilai netral atau nol (0).

Hashtag atau tagar yang akan jadi bahan analisis untuk melakukan Analisis Sentimen adalah #JokowiAminMenangDebat dan #PrabowoIndonesiaMenang yang tweet-nya diposting pada tanggal 17 Januari 2019.

```
positif = sum(hasil$score > 0)
negatif = sum(hasil$score < 0)
netral = sum(hasil$score == 0)</pre>
```

Gambar 2. Pembobotan Teks untuk proses Analisis Sentimen pada Pemrograman *R*.

III. HASIL DAN PEMBAHASAN.

A. Hasil Sentimen Analisis terhadap tweet #JokowiAminMenangDebat

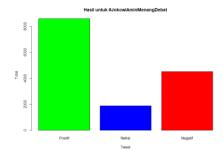
Data yang di gunakan untuk proses analisis sentimen adalah data *tweet* yang telah diambil melalui *Twitter API* saat debat belangsung dengan menyertakan *hashtag* #JokowiA minMenangDebat pada jam 18.30 sampai 23.59 yang sebanyak 15.000 percakapan. Proses Sentimen Analisis percakapan debat calon presiden di *Twitter* yang telah dilakukan ini mengunakan pemrograman *R*. Untuk proses eksekusi skrip di pemrograman *R* adalah dengan memasukkan skrip untuk fungsi melakukan proses pembobotan kata dan pengklasifikasian Sentimen dari satu per satu *tweet* (Positif, Negatif, Netral).

Hasil yang telah di hasilkan di pemrograman *R* untuk pembobotan kata dan pengklasifikasian adalah tabel yang berisi kolom skor dari pembobotan *tweet*, kolom yang berisi *tweet*, dan hasil klasifikasi berdasarkan hasil pembobotan kata yang hasilnya berupa Sentimen positif, negatif, atau netral. Dari hasil Analisis Sentimen ini dapat disimpulkan bahwa *tweet* yang menyertakan #JokowiAminMenangDebat dengan sentimen positif berjumlah 8.769 *tweet*, *tweet* dengan sentimen negatif berjumlah 4.518, dan *tweet* yang cenderung netral berjumlah 1.713.

Gambar 3. Hasil Pembobotan dan klasifikasi kata #JokowiAminMenangDebat di Pemrograman *R*

```
> count(hasil, vars = "klasifikasi")
  klasifikasi freq
1   negatif 4518
2   netral 1713
3   positif 8769
```

Gambar 4. Jumlah masing-masing Pembobotan kata #JokowiAminMenangDebat di Pemrograman R

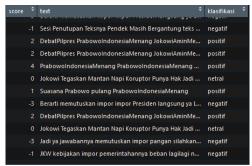


Gambar 5. *Barplot* #JokowiAminMenangDebat di Pemrograman *R*

B. Hasil Sentimen Analisis terhadap tweet #PrabowoIndonesiaMenang

Data yang di gunakan untuk proses analisis sentimen adalah data *tweet* yang telah diambil melalui *Twitter API* saat debat belangsung dengan menyertakan *hashtag* #PrabowoIndonesiaMenang pada jam 18.30 sampai 23.59 yang sebanyak 25.000 percakapan. Proses Sentimen Analisis percakapan debat calon presiden di *Twitter* yang telah dilakukan ini mengunakan pemrograman *R*. Untuk proses eksekusi skrip di pemrograman *R* adalah dengan memasukkan skrip untuk fungsi melakukan proses pembobotan kata dan pengklasifikasian Sentimen dari satu per satu *tweet* (Positif, Negatif, Netral).

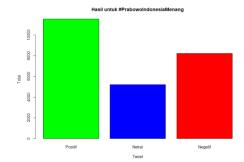
Hasil yang telah di hasilkan di pemrograman R untuk pembobotan kata dan pengklasifikasian adalah tabel yang berisi kolom skor dari pembobotan tweet, kolom yang berisi tweet, dan hasil klasifikasi berdasarkan hasil pembobotan kata yang hasilnya berupa Sentimen positif, negatif, atau netral. Hasil akhir dari proses Sentimen Analisis menggunakan Pemrograman R pada Data di Twitter yang menyertakan hashtag #PrabowoIndonesiaMenang pada tweet mereka adalah, sentimen positif berjumlah 11.551 tweet, lalu sentimen negatif berjumlah 8.238 tweet, dan jumlah tweet netral 5.211.



Gambar 6. Hasil Pembobotan dan klasifikasi kata #PrabowoIndonesiaMenang di Pemrograman *R*

```
> count(hasil, vars = "klasifikasi")
  klasifikasi freq
1   negatif 8238
2   netral 5211
3   positif 11551
```

Gambar 7. Jumlah masing-masing Pembobotan kata #PrabowoIndonesiaMenang di Pemrograman *R*



Gambar 8. Hasil *Barplot* #PrabowoIndonesiaMenang di Pemrograman *R*

IV. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang dilakukan tentang analisis sentimen percakapan di *Twitter* pada saat debat pertama calon presiden di Indonesia berlangsung pada tanggal 17 Januari 2018 yang menggunakan data dari *Twitter* pada jam 18.30 hingga 23.59 dapat disimpulkan bahwa, sentimen yang di tunjukan publik dunia maya pada saat debat berlangsung cenderung bersentimen positif.

Berdasarkan hasil pada tabel dan *barplot* pada pemrograman *R*, hasil persentase, baik *tweet* yang menggunakan *hashtag* #JokowiAminMenangDebat maupun *hashtag* #PrabowoIndonesiaMenang menunjukkan cenderung mengandung sentimen positif lebih banyak dari sentimen negatif dan netral pada tiap *tweet*. Untuk sentimen negatif pada kedua *hashtag*, jumlahnya cukup banyak pada setiap *tweet* yang diposting saat debat berlangsung, walaupun memang tak sebanyak sentimen positif.

Hasil sentimen netral pada kedua *hashtag* ini, bisa dikatakan cukup banyak dan bisa disimpulkan juga bahwa, terdapat pengguna *Twitter* yang tidak terlalu berorientasi pada salah satu calon presiden dan wakil presiden tertentu dan cenderung menjadi pemilih yang belum menentukan pilihan mereka saat pemilihan umum nanti yang menjadikan debat sebagai cara meyakinkan mereka untuk kepada siapa mereka akan menentukan pilihannya dan memposting *tweet* mereka di *Twitter* saat debat berlangsung.

Terdapat hasil yang tidak jauh berbeda antara kedua hashtag pada setiap hasil analisis sentimen dapat di sebabkan karena pada tweet dengan sentimen positif, negatif, ataupun netral para pengguna Twitter saat debat berlangsung menyertakan kedua hashtag tersebut bersamaan. Untuk perolehan jumlah data yang didapatkan berbeda untuk masingmasing hashtag, karena Pemrograman R dalam memproses Twitter API tidak mampu memproses data dalam waktu bersama pada satu waktu.

Saran-saran untuk untuk penelitian lebih lanjut untuk menutup kekurangan penelitian. Tidak memuat saran-saran diluar untuk penelitian lanjut.

DAFTAR PUSTAKA

- [1] A. Y. Rahman, M. Sa'adah, F. W. Setiawan, and E. Supriyanto, "Vice Presidential Election Prediction Period 2019- 2024 using Simple Additive Weighting," pp. 56–60, 2018.
- [2] E. L. Damanik, "Hashtag #2019GantiPresiden: Sentimen Anti-petahana dan Orientasi Politik Pemilih Pemula menghadapi Pilpres 2019 di Pematangsiantar," JPPUMA J. Ilmu Pemerintah. dan Sos. Polit., 2018.
- [3] R. Hidayatillah, M. Mirwan, M. Hakam, and A. Nugroho, "Levels of Political Participation Based on Naive Bayes Classifier," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 13, no. 1, pp. 73–82, 2019.
- [4] L. A. Nugroho and Alves, "Analisis Sentimen Publik Terhadap Kembalinya Partai Fretilin Memimpin Timor-Leste," 2017.
- [5] C. W. Park and D. R. Seo, "Sentiment analysis of Twitter

- corpus related to artificial intelligence assistants," pp. 495–498, 2018.
- [6] V. N. Patodkar and S. I.R, "Twitter as a Corpus for Sentiment Analysis and Opinion Mining," *IJARCCE*, vol. 5, no. 12, pp. 320–322, 2016.
- [7] V. Ikoro, M. Sharmina, K. Malik, and R. Batista-Navarro, "Analyzing Sentiments Expressed on Twitter by UK Energy Company Consumers," 2018 Fifth Int. Conf. Soc. Networks Anal. Manag. Secur., pp. 95–98, 2018.
- [8] U. R. Hodeghatta, "Sentiment Analysis of Hollywood Movies on Twitter," *Proc. 2013 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Min.*, pp. 1401–1404, 2013.
- [9] R. Sandoval-Almazan and D. Valle-Cruz, "Facebook Impact and Sentiment Analysis on Political Campaigns," Proc. 19th Annu. Int. Conf. Digit. Gov. Res. Gov. Data Age, pp. 561–567, 2018.
- [10] U. Yaqub, S. A. Chun, V. Atluri, and J. Vaidya, "Sentiment based Analysis of Tweets during the US Presidential Elections," *Present. Proc. 18th Annu. Int. Conf. Digit. Gov. Res.*, pp. 149–156, 2017.
- [11] C. V. S. Araujo, R. M. Neto, F. N. G., and E. F. Nakamura, "Predicting Music Success Based on Users' Comments on Online Social Networks," *Proc. 23rd Brazillian Symp. Multimed. Web*, pp. 1–10, 2017.
- [12] Sharma P and Moh T, "Prediction of Indian election using sentiment analysis on Hindi Twitter," pp. 1996–1971, 2016.
- [13] W. Budiharto and M. Meiliana, "Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis," *J. Big Data*, vol. 5, no. 1, 2018.
- [14] N. Matloff, "The Art of R Programming," p. 404.
- [15] and J. S. H. H. Kaur, M. Talluri, "Get Twitter information: A collaborative Android application for big data analysis," 2015 Int. Conf. Collab. Technol. Syst., pp. 483–484, 2015.
- [16] S. K. Ravindran and V. Garg, "Mastering social media mining with R: extract valuable data from social media sites and make better business decisions using R," 2015.
- [17] E. de Jonge, "An introduction to data cleaning with R," p. 53.
- [18] D. H. Wahid and A Sn, "Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 10, no. 2, pp. 207–218, 2016.
- [19] B. Liu, M. Hu, and J. Cheng, "Opinion observer: analyzing and comparing opinions on the Web," *Proc.* 14th Int. Conf. World Wide Web WWW, p. 342, 2005.
- [20] H. Wang and and Y. Z. F. Wu, X. Li, S. Tang, J. Shao, "Indeks Kualiti Air (IKA) Sg. Selangor pasca pembinaan Water Quality Index (WQI) of the Selangor River after the construction of the Selangor River Dam," *Jointly Discov. Fine-grained Coarse-grained Sentim. via Top. Model.*, pp. 913–916, 2014.

JOINTECS Vol. 4, No.2, Mei 2019 DOI: 10.31328/jointecs.v4i2.1004

- [21] and J. C. T. Z. Wang, C. S. Chong, L. Lan, Y. Yang, S. B. Ho, "Fine-grained sentiment analysis of social media with emotion sensing," 2016 Futur. Technol. Conf., pp. 1361–1364, 2016.
- [22] I. Percy and and S. S. A. Balinsky, H. Balinsky, "Text Mining and Recommender Systems for Predictive Policing," *Proc. ACM Symp. Doc. Eng. 2018*, pp. 151–154, 2018.

JOINTECS Vol. 4, No.2, Mei 2019
DOI: 10.31328/jointecs.v4i2.1004
p-ISSN: 2541-3619
e-ISSN: 2541-6448