

# Pemanfaatan Sentiment Analysis Menggunakan Data Twitter untuk Melihat Tren Positif dan Negatif pada Kampanye Pemilihan Gubernur Jawa Tengah 2018

Vita Tantriyati, dan Radius Tanone

## **Abstraksi**

*Article submitted*  
2018-04-19

*Author revision submitted*  
2019-03-11

*Rini Darmastuti*  
*Editor decision submitted*  
2018-12-17

Media sosial adalah salah satu produk teknologi informasi yang banyak dipakai masyarakat untuk memaparkan pendapat mereka baik itu yang bersifat pribadi ataupun yang umum seperti pada pemilihan Gubernur (pilgub) Jawa Tengah. Twitter merupakan salah satu media sosial yang dengan mudah dimanfaatkan oleh masyarakat untuk menyampaikan pendapat mereka ataupun mengikuti perkembangan berita terkait Pilgub. Dengan memanfaatkan *hashtag* dari data Twitter dilakukan proses *sentiment analysis* untuk mengetahui respon masyarakat terhadap pilgub Jawa Tengah 2018. Hasil yang didapat dari *sentiment analysis* yaitu respon masyarakat terhadap pilgub Jateng 2018 positif. Sedangkan dukungan lebih banyak diberikan kepada pasangan calon Ganjar Pranowo dan Taj Yasin.

## **Abstract**

*Social media is one of the information technology products which is used by the society to explain broadly ones' opinions either personal or public characteristic as well, such as in the election for governor of Central Java. Twitter is one of social medias which is easily utilized by people to explain their ideas or, to update the following news development in line with the affair of election for governor known as "Pilgub". By utilizing hashtag from Twitter data sentiment analysis process to find out the public response to Pilgub of Central Java in 2018. The*

*results obtained from the sentiment analysis of the community response to the Pilgub of Central Java in 2018 is positive. While more support is given to candidate pairs Ganjar Pranowo and Taj Yasin.*

**Keywords:** *sentiment analysis, twitter, governor election of central java*

## 1. Pendahuluan

Indonesia adalah negara kesatuan yang memiliki populasi manusia nomor empat di dunia. Dalam menjalankan pemerintahan yang demokratis, Pemilihan Umum (Pemilu) menjadi bagian dalam sistem pemerintahan Indonesia. Di samping itu, ada juga Pemilihan Kepala Daerah (Pilkada) yang dilaksanakan pada sistem pemerintahan Indonesia. Pemilu atau Pilkada menjadi bagian penting dalam pesta demokrasi bagi rakyat Indonesia dalam menentukan pemimpinnya beberapa waktu ke depan. Hal ini yang membuat pemilihan pejabat publik mulai dari Presiden hingga ke Kepala Daerah (Bupati) menggunakan pemilu dan pilkada sebagai cara untuk menjalankan demokrasi di Indonesia. Dengan sistem pemerintahan yang demokratis, Indonesia membutuhkan partisipasi rakyat dalam menyampaikan pendapat terkait hal-hal yang bersifat kenegaraan (terkait pemilu atau pilkada). Ketika pemilu atau pilkada akan dilaksanakan, ada beberapa tahapan yang tentunya dilalui oleh para kandidat yang akan berjuang di pesta demokrasi. Tahapan-tahapan tersebut tertuang dalam <https://infopemilu.kpu.go.id/> yang memuat satu tahapan beberapa tahapan yang ada menyebutkan bahwa sosialisasi kepada masyarakat menjadi agenda penting dari pesta demokrasi bagi rakyat Indonesia ini.

Ketika rakyat Indonesia menjadi bagian dari pesta demokrasi ini maka tentunya ada banyak aspek yang turut berperan serta dalam mewarnai hal ini. Aspek-aspek tersebut meliputi ekonomi, sosial maupun teknologi. Sebagai negara terbesar keempat di dunia menurut <http://www.worldometers.info> untuk populasi penduduknya yang berjumlah kira-kira 266 juta lebih, maka bisa dikatakan bahwa rakyat sangat bisa memberikan "warna" bagi kegiatan pemilu atau pilkada. Hal ini dapat dilihat lebih jelas lagi pada penetrasi penggunaan teknologi informasi terkhususnya media sosial oleh rakyat Indonesia. Data dari <http://www.wearesocial.com> menunjukkan bahwa 51% penduduk Indonesia sudah menggunakan internet dalam kesehariannya. Berdasarkan pada data tersebut tentunya bisa dilihat secara jelas dan dapat dikatakan bahwa lebih dari setengah penduduk Indonesia sudah melek akan penggunaan Internet. Penggunaan internet bisa untuk berbagai kepentingan misalnya membuka *website*, mengirimkan email ataupun menggunakan sosial media seperti Twitter, Facebook atau yang lainnya untuk kepentingan-kepentingan tertentu. Bila fokus pada penggunaan media sosial maka secara lebih spesifik terkait penggunaan teknologi informasi, [statista.com](http://www.statista.com) melaporkan bahwa 38.42 % penduduk Indonesia mengakses sosial media dalam kesehariannya. Seterusnya diprediksi lagi bahwa pada 2022 jumlah penduduk pengguna sosial media di Indonesia bertambah menjadi 44.83%.

Ketika pemilu atau pilkada berjalan, tentunya ada tahapan sosialisasi kepada masyarakat dari KPU untuk memberikan pandangan kepada masyarakat terkait pesta

demokrasi tersebut. Hal ini jelas membuat masyarakat dengan mudah dapat menyampaikan penilaian baik bersifat pribadi ataupun lembaga terkait para kandidat yang akan mengikuti pesta demokrasi di media sosial. Hal-hal seperti berita, foto ataupun status-status yang sifatnya positif ataupun negatif dapat dengan mudah bertebaran di media sosial. Tentu hal ini akan menjadi dampak baik maupun buruk yang mana dapat dilihat pada penggiringan opini publik kepada hal-hal tertentu.

Dari latar belakang di atas, penelitian ini ingin melihat padaketerkaitan antara kegiatan pemilu atau pilkada dengan penggunaan sosial media di masyarakat Indonesia. Banyak media sosial yang digunakan oleh rakyat Indonesia seperti Twitter, Facebook, Instagram ataupun yang lainnya. Fokus pilkada yang dipilih yakni pilkada di Jawa Tengah. Namun pada penelitian ini Twitter dipilih karena media sosial ini memiliki karakteristik yang memudahkan orang untuk menulis status biasa ataupun penggunaan *hashtag* (#) dengan batas maksimal karakter adalah 280. Menggunakan teknik *sentiment analysis* maka data Twitter diambil lalu dianalisis sehingga harapannya akan muncul hasil pengaruh positif ataupun negatif dari opini masyarakat di media sosial terkait pemilu ataupun pilkada ini. Sehingga manfaat yang ingin dicapai dari penelitian ini adalah membantu pihak-pihak seperti KPU dan yang lainnya dalam melihat tanggapan publik. Penelitian ini akan mengambil beberapa sampel data (status atau *hashtag*) dari beberapa calon Gubernur pada pilkada di Jawa Tengah.

## **2. Tinjauan Pustaka**

Dalam melakukan penelitian ini, adapun penelitian-penelitian sebelumnya yang dipakai sebagai acuan. Penelitian pertama yang diambil adalah penelitian dengan judul "Twitter Data Mining for Sentiment Analysis On People Feedback Against Government Public Policy". Fokus dalam penelitian ini adalah penggunaan media sosial seperti Facebook, Twitter, Instagram, Path dan media sosial lainnya maka dilihat bagaimana respon balik dari masyarakat terhadap kebijakan yang dikeluarkan oleh Gubernur. Hasil dari penelitian ini adalah melihat tren positif atau negatif dari masyarakat terkait kebijakan yang dikeluarkan oleh Gubernur (Syaifudin, Puspitasari, 2017).

Penelitian lain yang sudah dilakukan oleh Hadi dkk pada 2017 dengan judul "Text Mining pada Media Sosial Twitter Studi Kasus: Masa Tenang Pilkada DKI 2017 Putaran 2". Pada penelitian ini, sentimen diberikan kepada data dengan menghitung jumlah kata positif dan negatif yang telah didefinisikan oleh peneliti berdasarkan observasi terhadap beberapa sampel data yang diambil secara acak. Hasil temuan ini ternyata sejalan dengan hasil pemilihan pada pilkada di DKI pada tahun 2017, yaitu kemenangan pada pasangan Anies-Sandi. Selain itu pada metode-metode *unsupervised learning* pada penelitian ini ditemukan bahwa metode *k-means* tidak dapat memberikan hasil yang merata pada setiap kelompoknya. Sebaliknya, hasil luaran dari pemodelan topik (topic modeling – Latent Dirichlet Allocation) lebih merata. Selain itu hasil pengelompokkan dari metode *k-means* dan topic modeling pada data tanggal 18 April 2017 memiliki nilai SSE (*k-means*) dan *marginal likelihood* (topic model) yang lebih baik daripada data tanggal lainnya. Hal ini disebabkan oleh karena keragaman data yang terdapat pada 18 April 2017 tersebut lebih rendah (Hadi,

dkk, 2017).

Penelitian sebelumnya dilihat sebagai acuan untuk melakukan penelitian ini. Pada penelitian ini, *sentiment analysis* dipakai sebagai cara untuk melihat tren positif, negatif ataupun netral pada opini publik di Pilkada Jateng. Data Twitter akan diambil menggunakan periode waktu tertentu dan akan dilakukan analisis untuk mendapatkan hasil tren opini masyarakat terhadap pilkada.

Twitter adalah tempat untuk menemukan apa yang sedang terjadi di dunia saat ini. Apapun yang kita sukai seperti musik, olahraga, politik, berita, selebritas atau kejadian setiap hari bisa dilihat ada media sosial ini. Secara umum Twitter memiliki lima konten utama yaitu *tweet, retweet, follow, search* dan *hashtag* (twitter, 2018).

*Sentiment analysis* adalah sebuah proses untuk menjelaskan sebuah tulisan berarti positif, negatif atau netral. *Sentiment analysis* sering juga dikenal sebagai *opinion mining* atau proses menggali opini atau sikap dari pembicara. Saat ini kasus yang banyak dijumpai adalah bagaimana hal ini dipakai untuk menggambarkan orang-orang merasakan tentang topik-topik tertentu (lexalytic.com, 2018).

Sedangkan dalam buku yang ditulis oleh Liu Bing dikatakan bahwa "*we use the term opinion as a broad concept that covers sentiment, evaluation, appraisal, or attitude and associated information such as opinion target and the person who holds the opinion, and we use the term sentiment to mean only the underlying positive or negative feeling implied by opinion*" (Liu, 2015). Hal ini berarti di dalam sebuah opini bisa terdapat nilai-nilai seperti *sentiment* dan nilai opini yang bersifat negatif atau positif.

### 3. Metode Penelitian

Pada penelitian ini, adapun metode penelitian yang dipakai untuk mendapatkan hasil sesuai tujuan penelitian. Menurut Sugiyono (2005: 21) menyatakan bahwa metode deskriptif adalah suatu metode yang digunakan untuk menggambarkan atau menganalisis suatu hasil penelitian tetapi tidak digunakan untuk membuat kesimpulan yang lebih luas. Menurut Whitney (1960: 160) metode deskriptif adalah pencarian fakta dengan interpretasi yang tepat. Adapun tahapan yang dilalui dalam melakukan penelitian ini sesuai metode deskriptif yaitu melakukan studi kasus. Dalam menjalankan penelitian ini, terdapat tiga langkah yakni:

- 1) Penentuan teknik pengumpulan data; Pada tahapan ini ditentukan teknik mengumpulkan data. Teknik pengumpulan data yang dipakai adalah dengan melakukan observasi langsung menggunakan media teknologi informasi berupa Twitter. Dengan menggunakan hashtag tertentu yang dipilih sebagai fokus permasalahan, kemudian dilakukan identifikasi untuk dapat lanjut ke tahapan selanjutnya.
- 2) Penyebaran alat pengumpulan data; Pada tahapan ini dilakukan penyebaran alat pengumpulan data. Pada tahapan ini yang dilakukan adalah mencari data di Twitter dengan *hashtag* (#) tertentu secara berkala (waktu).
- 3) Penganalisisan bukti studi kasus yang terkumpul. Pada tahap ini dilakukan analisa menggunakan *sentiment analysis* untuk melihat data dan fakta yang didapatkan di lapangan. Data dan fakta ini diolah untuk kemudian melihat hasil opini publik terhadap pilkada dengan menggunakan *hashtag* tertentu. Hal ini tentunya akan menunjukkan pada penilaian opini baik yang bersifat positif, negatif

maupun netral.

#### 4. Pembahasan

Sentimen dilakukan dengan cara membandingkan banyak kata positif dan negatif dalam data tersebut. Apabila jumlah kata positif lebih banyak daripada kata negatif, maka data tersebut bersentimen positif, begitu pula sebaliknya. Dengan demikian daftar kata-kata positif dan negatif sangatlah penting pada penelitian ini. Pada penelitian ini, digunakan tiga *hashtag* untuk menganalisis sentimen terhadap para pasangan calon Gubernur.

*Word cloud* merupakan salah satu metode untuk menampilkan data teks secara visual. Cara kerja dari *word cloud* ialah dengan menampilkan sekumpulan kata yang frekuensinya sering muncul ke dalam bentuk yang menarik namun tetap informatif. Semakin kata itu sering muncul maka ukuran *word cloud* akan semakin besar. Karena penelitian yang dilakukan berkaitan dengan pilgub Jateng tahun 2018, maka *hashtag* yang digunakan antara lain ialah #GanjarYasin #SudirmanSaid dan #PilgubJateng2018, dari ketiga *hashtag* tersebut didapatkan hasil seperti yang ditampilkan pada Gambar 4.1, Gambar 4.2, dan Gambar 4.3.



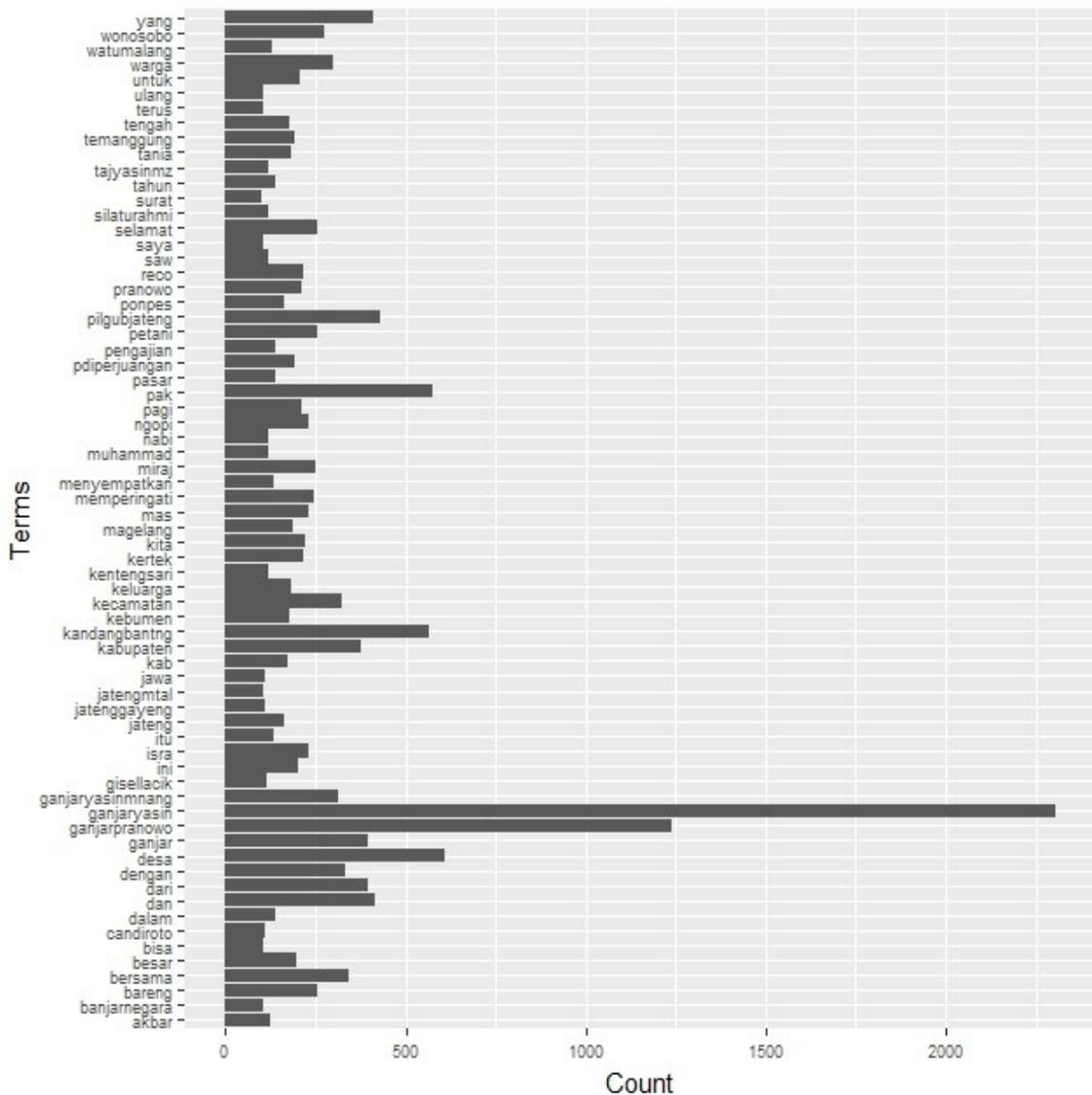
Gambar 4.1 *Word cloud* #GanjarYasin

Berdasarkan Gambar 4.1 dapat dilihat bahwa kata yang paling sering muncul ialah kata ganjaryasin, dari hal tersebut dapat disimpulkan jika kata yang paling sering digunakan dari hashtag #GanjarYasin ialah ganjaryasin itu sendiri, dan kata lain yang sering digunakan juga ialah *pilgubjateng*, dan *ganjaryasinmenang*.



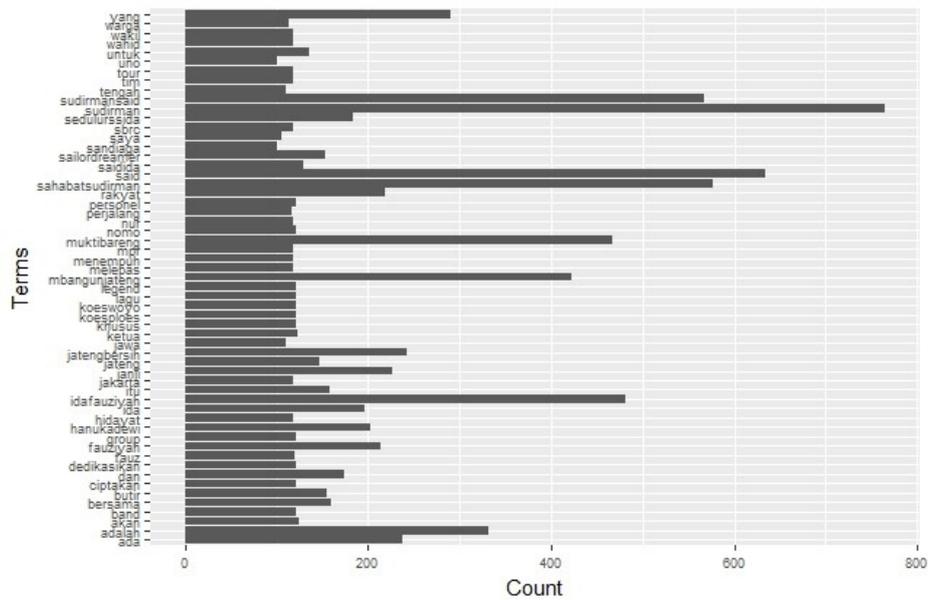
dengan frekuensi terbanyak sama yaitu *hashtag* #GanjarYasin dan #PilgubJateng2018.

Selain menggunakan *word cloud* pengolahan data *sentiment analysis* yang lain adalah menggunakan *plotting*. Dalam *plotting*, ketiga *hashtag* diolah kembali secara lebih signifikan untuk mempermudah pembacaannya, hasil dari ketiga *hashtag* yang sudah diolah dalam *plotting* dapat dilihat pada Gambar 4.4, Gambar 4.5 dan Gambar 4.6.



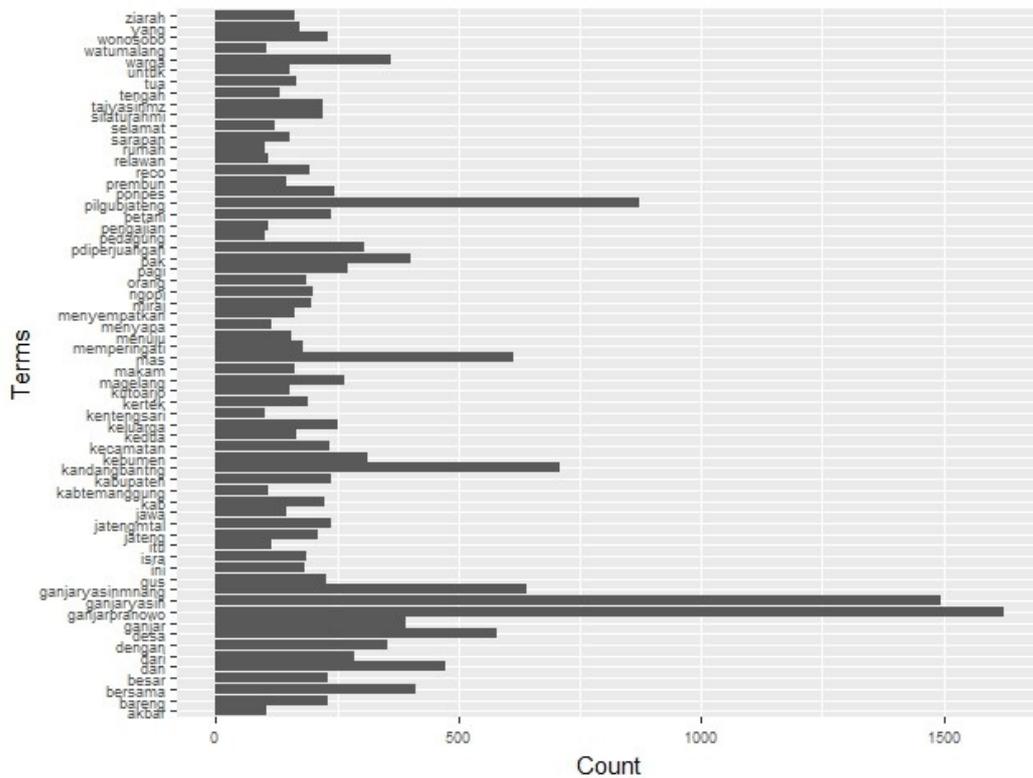
Gambar 4.4 *Plotting* kata dari *hashtag* #GanjarYasin

Dari *plotting* yang dihasilkan pada Gambar 4.4 di atas, dapat dilihat bila lebih dari 2000 *tweet* yang menggunakan kata ganjaryasin dan lebih dari 1000 *tweet* menggunakan kata Ganjarpranowo. Hasil yang didapat dari *plotting* tersebut bisa dilihat bila kata tertingginya adalah *ganjaryasin* dan *ganjarpranowo*.



Gambar 4.5 *Plotting* dari hashtag #SudirmanSaid

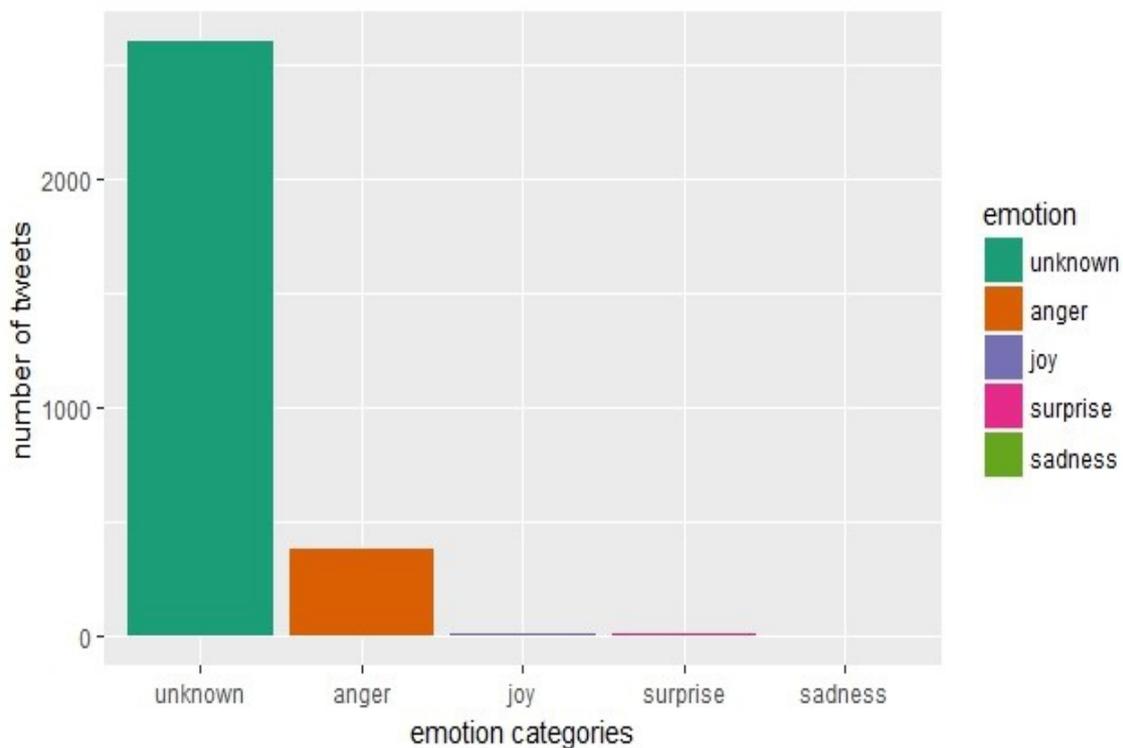
Kemudian berdasarkan Gambar 4.5 untuk *plotting* dari hashtag #SudirmanSaid, kata sudirman merupakan kata yang sering keluar di urutan pertama, karena lebih dari 700 *tweet* menggunakan kata tersebut saat menulis hashtag #SudirmanSaid, sedangkan di urutan kedua yang kata yang paling sering keluar ialah kata said yang mana lebih dari 600 *tweet* menggunakan kata tersebut.



Gambar 4.6 *Plotting* dari hashtag #PilgubJateng2018

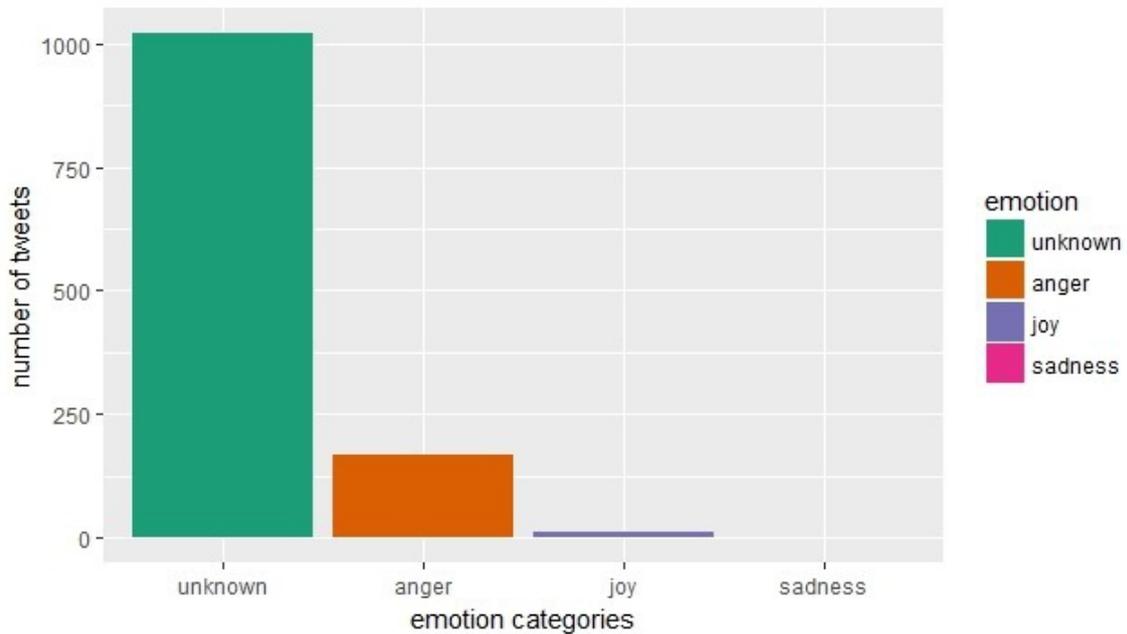
Hasil dari *plotting* dari *hashtag* #PilgubJateng2018 pada Gambar 4.6 di atas menunjukkan hasil seperti berikut, yaitu dari *plotting* yang sudah dilakukan dapat dilihat jika lebih dari 1700 *tweet* menggunakan kata ganjarpranowo pada *hashtag* tersebut sehingga menyebabkan kata tersebut menjadi kata pertama yang sering muncul pada *hashtag* #PilgubJateng2018, setelah itu kata kedua yang muncul adalah *ganjaryasin*, yaitu nama gubernur Jawa Tengah tersebut sering disebut oleh masyarakat di Twitter pada *hashtag* tersebut, hampir 1500 *tweet* menyebut nama *ganjaryasin* pada *hashtag* #PilgubJateng2018.

Teknik selanjutnya yang dilakukan adalah menilai emosi dari *tweet* yang ditulis oleh masyarakat mengenai Pilgub Jateng 2018. Terdapat bermacam-macam kategori emosi yang dapat ditampilkan dari hasil *emotion analysis* ini, yaitu *unknown* (tidak mengekspresikan apapun), *anger* (kemarahan), *joy* (senang), *sadness* (kesedihan), *surprise* (terkejut), *disgust* (tidak menyenangkan) dan lain-lain. Langkah selanjutnya yang dilakukan yaitu memproses *hashtag* yang ada dengan *emotion analysis*. Hasil pemrosesannya dapat dilihat pada Gambar 4.7, Gambar 4.8 dan Gambar 4.9 di bawah ini.



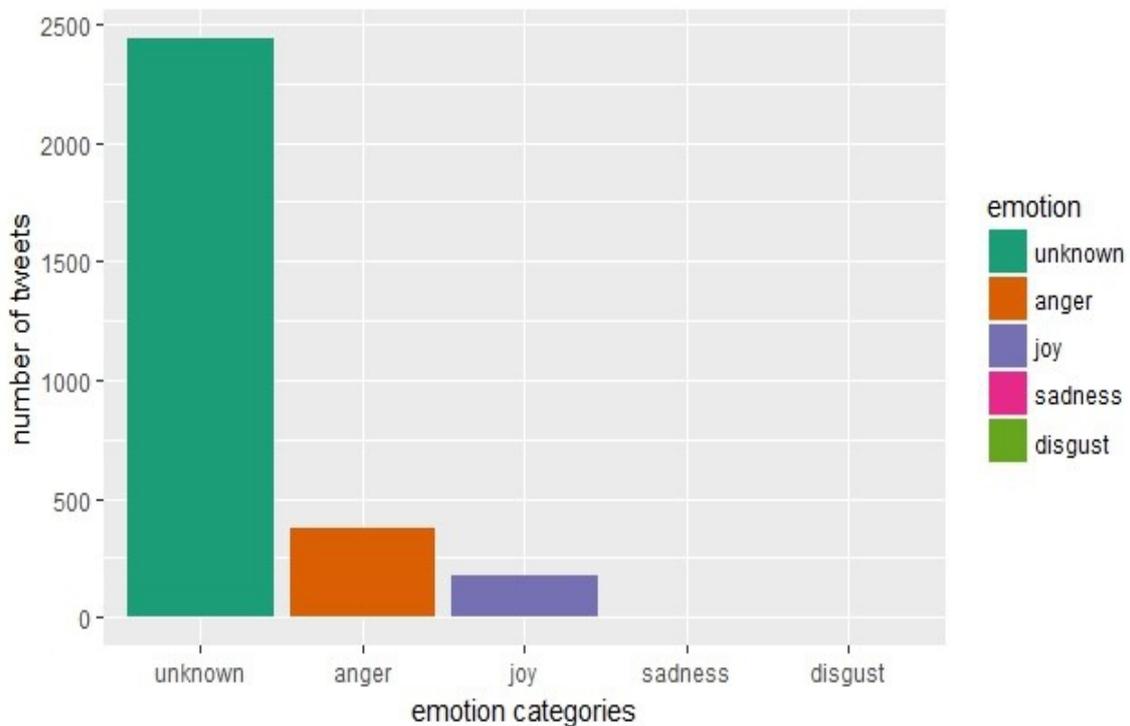
Gambar 4.7 Emosi dengan *hashtag* #GanjarYasin

Gambar 4.7 merupakan hasil penilaian emosi terhadap *hashtag* #GanjarYasin. Pada hasil tersebut dapat dilihat bahwa lebih dari 2000 kata yang tidak mengekspresikan apapun terhadap *hashtag* tersebut, dan ada kurang dari 500 kata yang mengekspresikan kemarahan serta ada sedikit kata yang mengandung kesenangan dan emosi terkejut terhadap *hashtag* tersebut.



Gambar 4.8 Emosi dengan *hashtag* #SudirmanSaid

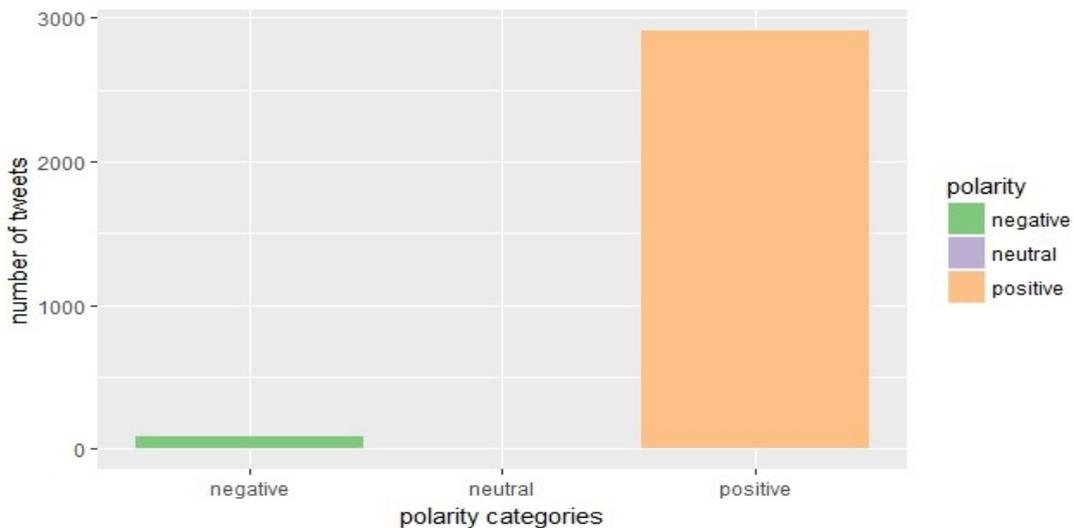
Sedangkan pada Gambar 4.8 adalah hasil dari pemrosesan *tweet* untuk *hashtag* #SudirmanSaid, ada lebih dari 1000 kata yang tidak mengekspresikan apapun terhadap *hashtag* tersebut, kemudian terdapat kurang dari 250 kata yang menunjukkan kemarahan terhadap *hashtag* tersebut dan sisanya adalah ekspresi senang.



Gambar 4.9 Emosi dengan *hashtag* #PilgubJateng2018

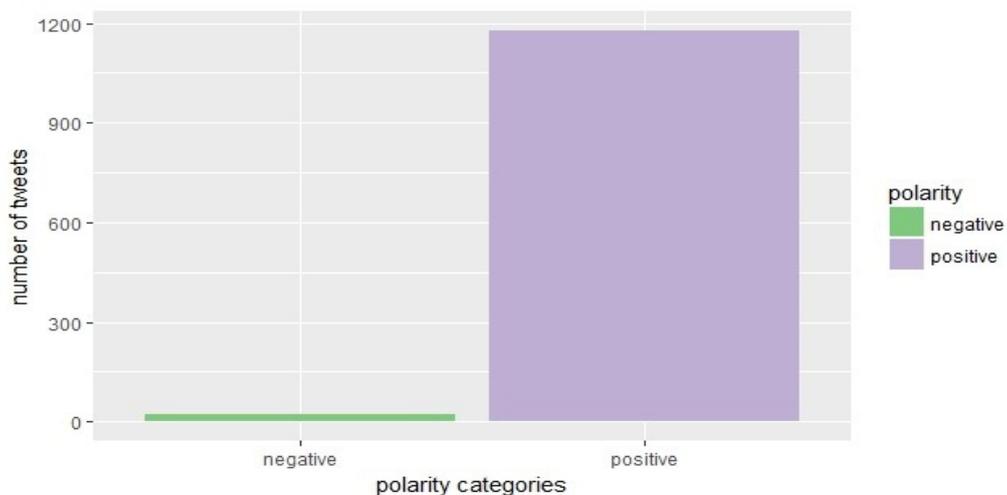
Untuk emosi yang ditunjukkan pada *hashtag* #PilgubJateng2018, terdapat kurang dari 2500 kata yang tidak menunjukkan emosi apapun. Kemudian kurang dari 500 kata menunjukkan kemarahan dalam *tweet* yang ditulis, dan terdapat kurang dari 100 kata yang menunjukkan emosi senang dalam *tweet*-nya, sisanya adalah emosi kesedihan dan tidak menyenangkan

Teknik selanjutnya yang dilakukan adalah menilai polaritas dari *tweet* yang ditulis, polaritas ini dilakukan untuk menilai apakah *tweet* yang disampaikan mengandung unsur positif negatif atau netral, hal ini dilakukan untuk menghindari kedepannya apakah *tweet* yang ditulis mengandung unsur negatif seperti *black campaign*.



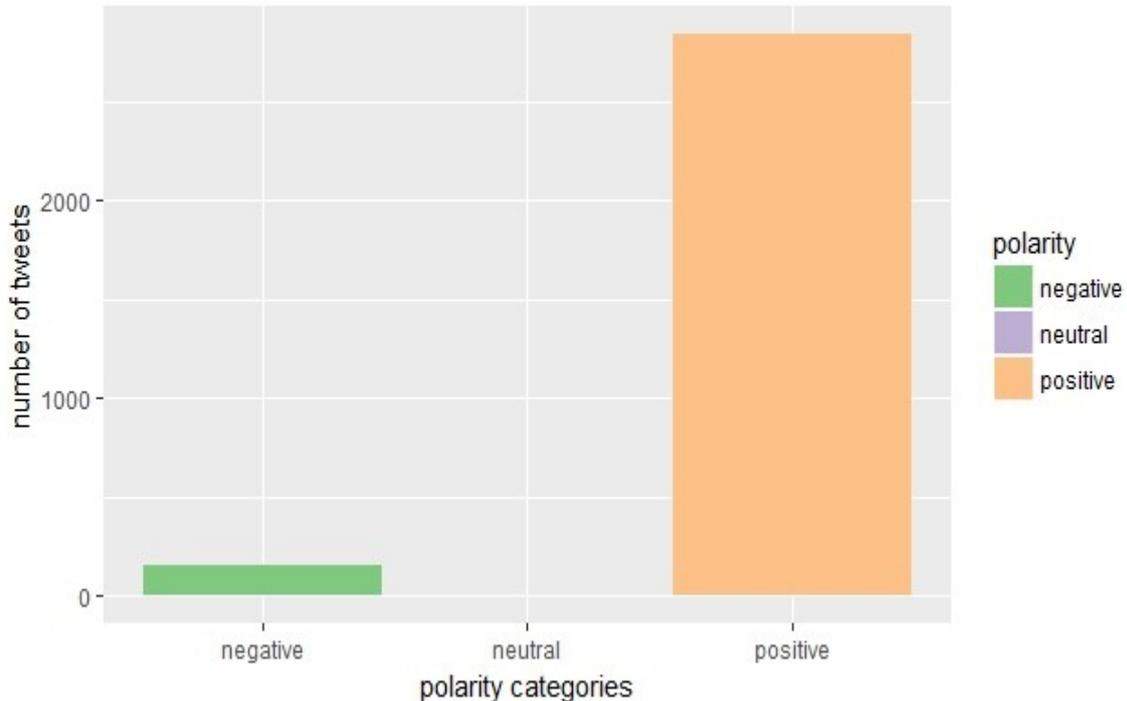
Gambar 4.10 Polaritas dengan *hashtag* #GanjarYasin

Gambar 4.10 di atas merupakan hasil pemrosesan dari polaritas dengan *hashtag* #GanjarYasin, dapat dilihat bahwa hampir 3000 *tweet* memberikan respon positif dari *tweet* dengan *hashtag* #GanjarYasin tersebut, dan ada beberapa respon negatif dari *hashtag* tersebut.



Gambar 4.11 Polaritas dengan *hashtag* #SudirmanSaid

Gambar 4.11 merupakan hasil pemrosesan untuk melihat polaritas dari *hashtag* #SudirmanSaid yang menunjukkan banyak respon positif yang diberikan untuk *hashtag* tersebut, bisa dilihat kurang lebih sekitar 1200 *tweet* yang memberikan respon positif terhadap *hashtag* tersebut. Sisanya menunjukkan respon negatif terhadap *hashtaa* itu.



Gambar 4.12 Polaritas dengan *hashtag* #PilgubJateng2018

Selanjutnya Gambar 4.11 merupakan polaritas untuk *hashtag* #PilgubJateng2018, dapat dilihat dari hasilnya yaitu lebih dari 2500 *tweet* menunjukkan respon positif untuk *hashtag* tersebut, dan sisanya menunjukkan respon negatif serta netral. Dari hal ini dapat dilihat bahwa hampir semua masyarakat memberikan respon positif terhadap Pilgub Jateng 2018 ini baik untuk calon Ganjar Pranowo-Taj Yasin dan juga Sudirman Said-Ida Fauziyah.

## 5. Kesimpulan

Berdasarkan hasil yang sudah didapatkan, disimpulkan bahwa respon masyarakat terhadap Pilgub Jateng 2018 sebagian besar mengarah kepada pasangan calon Ganjar Pranowo-Taj Yasin. Hal tersebut dapat dilihat pada hasil pemrosesan di *word cloud* serta *plotting* untuk *hashtag* #PilgubJateng2018. Sedangkan respon emosi yang diberikan di *tweet* yang ditulis lebih ke arah yang tidak menunjukkan emosi apapun atau netral dan respon netral tersebut bisa dikatakan respon yang positif, hal tersebut dapat dilihat dari polaritas yang dapat dilihat pada hasil pemrosesan polaritas untuk masing masing *hashtag*. Sebagian besar *hashtag* menunjukkan respon positif dan mungkin hanya beberapa orang yang merespon negatif terhadap Pilgub Jateng 2018.

### **Daftar Pustaka**

- Syaifudin and Puspitasari, "Twitter Data Mining for Sentiment Analysis On People Feedback Againsts Government Public Policy," *MATTER: International Journal of Science and Technology*, vol. 3, no. 1, pp. 110-122, 2017.
- Hadi, "Text Mining pada Media Sosial Twitter Studi Kasus: Masa Tenang Pilkada DKI 2017 Putaran 2," *Seminar Nasional Matematika dan Aplikasinya Universitas Airlangga*, pp. 325-331, 2017.
- Twitter, Inc, "Help Center," Twitter, Inc, 2018. [Online]. Available: <https://help.twitter.com/en/twitter-guide>. [Accessed 14 April 2018].
- Lexalytics, "Lexalytics," Lexalytics, 2018. [Online]. Available: <https://www.lexalytics.com/technology/sentiment>. [Accessed 14 April 2018].
- L. B, "Sentiment Analysis: Mining Opinions, Sentiment and Emotions," Cambridge University Press, Cambridge, June 4th, 2015.
- Sugiyono, "Memahami Penelitian Kualitatif," Alfabet, Bandung, 2005.
- F. Whitney, "The Element of Research," Hall, Inc, New York, 1960.

**Vita Tantriyati**, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana. Email: [vita.tantri@gmail.com](mailto:vita.tantri@gmail.com)

**Radius Tanone**, Fakultas Teknologi Informasi, Universitas Kristen Satya Wacana, Email: [radiustanone@gmail.com](mailto:radiustanone@gmail.com)

