

## Sentimen Analisis pada Sosial Media Twitter untuk Menilai Respon Masyarakat terhadap Seleksi Kartu Prakerja

Prasetyo Adi Nugroho<sup>1)</sup>, Nur Sucahyo<sup>2)\*</sup>, Ike Kurniati<sup>3)</sup>

<sup>1)2)3)</sup> Fakultas Teknologi, Institut Teknologi dan Bisnis Swadharma

<sup>\*)</sup>Correspondence Author: [n.sucahyo12@gmail.com](mailto:n.sucahyo12@gmail.com), Jakarta, Indonesia

DOI: <https://doi.org/10.37012/jtik.v9i1.862>

### Abstrak

Pemerintah selama tahun 2020 dan 2021 menyelenggarakan program kartu prakerja agar tercipta SDM yang unggul untuk Indonesia maju di masa yang akan datang sekaligus menekan angka pengangguran di Indonesia. Permasalahannya belum ada kejelasan pelaksanaan apakah program ini akan di lanjutkan atau tidak, berbagai kendala yang di hadapi baik calon peserta maupun yang sudah menjadi peserta masih kebingungan dengan program tersebut karena minimnya sosialisasi yang di dapat, sehingga banyak menimbulkan berbagai sentimen di media sosial khususnya twitter baik itu sentimen positif, negatif ataupun netral. Analisis sentimen menggunakan metode CRIPS-DM yang merupakan standar yang digunakan pada bidang data mining melalui proses-proses *business understanding*, *data understanding*, *data preparation*, *data modeling*, *evaluation* dan *deployment*. Hasil penelitian berupa visualisasi sentimen analisis prakerja yang disimpan dalam format Tableau memuat jumlah total tweets favorit, jumlah re-tweets teratas berdasarkan topik, persentase jumlah sentimen positif, netral dan negatif berdasarkan *count\_of\_hasilsentiment / count\_id*. Word cloud menampilkan deretan 10 topik yang paling banyak di pembicarakan di sosial media twitter berdasarkan *topic\_count*. Time series menampilkan tahun bulan dan tanggal pengambilan data di twitter pada periode bulan Februari sampai April tahun 2021 dalam bentuk grafik berdasarkan id pengguna. Nilai sentimen diklasifikasikan kedalam tiga kategori dengan menggunakan algoritma naïve bayes. Untuk sentimen negatif dengan total id pengguna sebanyak 5,859 dengan besaran persentase 12,98%, untuk sentimen netral dengan total id pengguna sebanyak 22,291 dengan besaran persentase 49,40%, dan terakhir untuk sentimen positif dengan total id pengguna sebanyak 16,972 dengan besaran persentase 37,61%. Dapat disimpulkan bahwa sebagian besar masyarakat tidak pro dan kontra atau berada pada posisi netral atas program kartu prakerja tersebut.

**Kata Kunci:** Sentiment Analisis , Sosial Media, Twitter, Kartu Prakerja

### Abstract

Indonesian government during 2020 and 2021 held a pre-employment card program to create superior human resources for advanced Indonesia in the future while simultaneously suppressing the unemployment rate in Indonesia. The problem is that there is no clarity in the implementation of whether this program will be continued or not, various obstacles faced by both prospective participants and those who have become participants are still confused about the program because of the lack of socialization that can be obtained, causing a lot of sentiments on social media, especially Twitter. It's a positive, negative or neutral sentiment. Sentiment analysis uses the CRIPS-DM method which is a standard used in the data mining sector through the processes of *business understanding*, *data understanding*, *data preparation*, *data modeling*, *evaluation* and *deployment*. The results of the research are visualization of pre-employment sentiment analysis stored in Tableau format containing the total number of favorite tweets, the number of top re-tweets by topic, the percentage of positive, neutral and negative sentiments based on *count\_of\_hasilsentiment / count\_id*. Word cloud displays a row of the 10 most discussed topics on Twitter based on *topic\_count*. The time series displays the year, month and date of data collection on Twitter in the period from February to April 2021 in graphical form based on the user id. Sentiment values are classified into three categories using the naïve Bayes algorithm. For negative sentiment with a total user id of 5,859 with a percentage of 12.98%, for neutral sentiment with a total user id of 22,291 with a percentage of 49.40%, and finally for positive sentiment with a total user id of 16,972 with a percentage of 37, 61%. It can be concluded that most of the people are not pro and contra or are in a neutral position on the pre-employment card program.

**Keywords:** Sentiment Analysis, Social Media, Twitter, Pre-Employment Cards

---

## PENDAHULUAN

Kehadiran sosial media di tengah masyarakat telah merubah kebiasaan cara berkomunikasi (Mustaqim and Kurniawan 2017), dahulu orang berkomunikasi secara langsung untuk bertukar informasi, pengalaman dan berita. Koran, radio dan televisi merupakan media yang digunakan oleh orang-orang untuk mencari berita dan informasi. Namun dengan adanya sosial media saat ini orang lebih memilih berbagi informasi, pengalaman, foto dan video yang dimiliki melalui media sosial. Tidak sedikit masyarakat yang menggunakan sosial media sebagai sarana mencari informasi seperti lowongan pekerjaan, peluang usaha untuk berbisnis, media pelatihan kerja dengan mengembangkan bakat melalui berbagai program yang disediakan baik oleh lembaga swasta maupun pemerintahan yang bisa di akses kapanpun dan dimanapun.

Program kartu prakerja adalah program pengembangan kompetensi kerja dan kewirausahaan yang ditujukan untuk pencari kerja, pekerja atau buruh yang terkena pemutusan hubungan kerja, dan pekerja yang membutuhkan peningkatan kompetensi, termasuk pelaku usaha mikro dan kecil. Program ini dirancang sebagai sebuah produk dan dikemas sedemikian rupa agar memberikan nilai bagi pengguna sekaligus memberikan nilai bagi sektor swasta. Jalan digital melalui *marketplace* dipilih untuk memudahkan pengguna mencari, membandingkan, memilih dan memberi evaluasi. Hanya dengan cara ini, program bisa terus diperbaiki, tumbuh dan relevan. Menggandeng pelaku usaha swasta, program ini adalah wujud kerjasama pemerintah dan swasta dalam melayani masyarakat dengan semangat gotong royong demi SDM unggul, Indonesia maju (Kementerian Koordinator Bidang Perekonomian Republik Indonesia 2021).

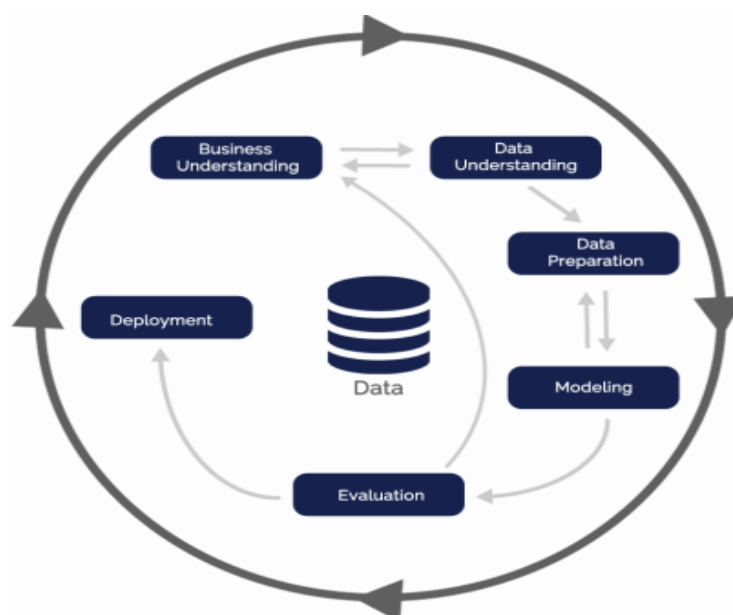
Analisis sentimen atau *opinion mining* merupakan proses memahami, mengekstrak dan mengolah data tekstual secara otomatis untuk mendapatkan informasi sentimen yang terkandung dalam suatu kalimat opini. Analisis sentimen dilakukan untuk melihat pendapat atau kecenderungan opini terhadap sebuah masalah atau objek oleh seseorang, apakah cenderung berpandangan atau beropini negatif atau positif. Salah satu contoh penggunaan analisis sentimen dalam dunia nyata adalah identifikasi kecenderungan pasar dan opini pasar terhadap suatu objek. *Opinion Mining* bisa dianggap sebagai kombinasi antara *text mining* dan *natural language processing*. Salah satu metode dari *Text Mining* yang bisa digunakan untuk menyelesaikan masalah *opinion mining* adalah *Naïve Bayes classifier* (NBC). NBC bisa digunakan untuk mengklasifikasikan opini ke dalam opini positif dan negatif. NBC bisa berfungsi dengan baik sebagai metode pengklasifikasi teks (Nurzahputra and Muslim 2016; Sudiantoro and Zuliarso 2018).

Twitter adalah layanan jejaring sosial dan *microblog* daring yang memungkinkan penggunaannya untuk mengirim dan membaca pesan berbasis teks hingga 140 karakter akan tetapi pada tanggal 07 november 2017 bertambah hingga 280 karakter yang di kenal dengan sebutan kicauan (tweet). Tingginya popularitas twitter menyebabkan layanan ini telah dimanfaatkan untuk berbagai keperluan dalam berbagai aspek, misalnya sebagai sarana protes, kampanye politik, sarana pembelajaran, dan sebagai media komunikasi darurat. Twitter juga dihadapkan pada berbagai masalah dan kontroversi seperti masalah keamanan dan privasi pengguna, gugatan hukum, dan penyensoran (Sucahyo, Nurlaela, and Waryono 2021).

Sebagian besar masyarakat menyampaikan respon melalui media sosial salah satunya Twitter. Media tersebut dipilih karena jumlah pengguna yang banyak dan kemudahan akses. Tidak sedikit masyarakat pengguna Twitter yang menyampaikan aspirasinya mengenai program kartu prakerja, ada yang meresponnya dengan baik dan ada juga merespon dengan buruk atau ada juga yang tidak peduli sama sekali atau dalam arti netral. Analisa sentimen kartu prakerja di media sosial twitter pernah dilakukan oleh (Alfikir 2020) namun menggunakan algoritma pencocokan string dan *library TextBlob* sedangkan pada penelitian ini akan menggunakan algoritma *Naïve Bayes classifier* (NBC).

## METODE

Pelaksanaan penelitian dilakukan dengan menggunakan metode *Cross Industry Standard Process Model for Data Mining* (CRISP-DM). Berikut gambaran alur proses yang terjadi pada CRISP-DM. (Suhanda, Kurniati, and Norma 2020)



Gambar 1. Alur CRISP DM

Salah satu keuntungan dari digunakannya metode ini adalah menjelaskan langkah paling umum dalam proses melakukan *data mining*. CRIPS-DM merupakan standar yang digunakan pada bidang *data mining*, antara lain terdiri dari proses-proses sebagai berikut:

1. *Business Understanding*

Adalah tahap pertama dalam CRISP-DM dan termasuk bagian yang cukup vital dalam proyek dan selama pengerjaan pada proses selanjutnya. Pada tahap ini membutuhkan pengetahuan dari objek bisnis, bagaimana membangun atau mendapatkan data, dan bagaimana untuk mencocokkan tujuan pemodelan untuk tujuan bisnis sehingga model terbaik dapat dibangun.

2. *Data Understanding*

*Data Understanding* adalah untuk memeriksa data, sehingga dapat mengidentifikasi masalah dalam data. Tahap ini memberikan fondasi analitik untuk sebuah proyek dengan membuat ringkasan (*summary*) dan mengidentifikasi potensi masalah dalam data. Tahap ini juga harus dilakukan secara cermat dan tidak terburu-buru, seperti pada visualisasi data, yang terkadang *insight*-nya sangat sulit didapat dihubungkan dengan *summary* datanya. Jika ada masalah pada tahap ini yang belum terjawab, maka akan mengganggu pada tahap Modeling. Ringkasan atau *summary* dari data dapat berguna untuk mengkonfirmasi apakah data terdistribusi seperti yang diharapkan, atau mengungkapkan penyimpangan tak terduga yang perlu ditangani pada tahap selanjutnya, yaitu data preparation. Masalah dalam data biasanya seperti nilai-nilai yang hilang, outlier, berdistribusi spike, berdistribusi bimodal harus diidentifikasi dan diukur sehingga dapat diperbaiki dalam data preparation.

3. *Data Preparation*

*Data preparation* adalah untuk memperbaiki masalah dalam data, kemudian membuat *variabel derived*. Tahap ini jelas membutuhkan pemikiran yang cukup matang dan usaha yang cukup tinggi untuk memastikan data tepat untuk algoritma yang digunakan. Bukan berarti saat *data preparation* pertama kali dimana masalah-masalah pada data sudah diselesaikan, data sudah dapat digunakan hingga tahap terakhir. Tahap ini merupakan tahap yang sering ditinjau kembali saat menemukan masalah pada saat pembangunan model. Sehingga dilakukan iterasi sampai menemukan hal yang cocok dengan data. Tahap sampling dapat dilakukan disini dan data secara umum dibagi menjadi dua, *data training* dan *data testing*.

4. *Data Modeling*

*Modeling* adalah untuk membuat model *prediktif atau deskriptif*. Pada tahap ini akhirnya dapat menggunakan statistika dan *machine learning* untuk mendapatkan *insight* yang berguna dari data untuk mencapai tujuan proyek. Beberapa *Modeling* yang biasa dilakukan adalah *classification, scoring, ranking, clustering, finding relation, dan characterization*.

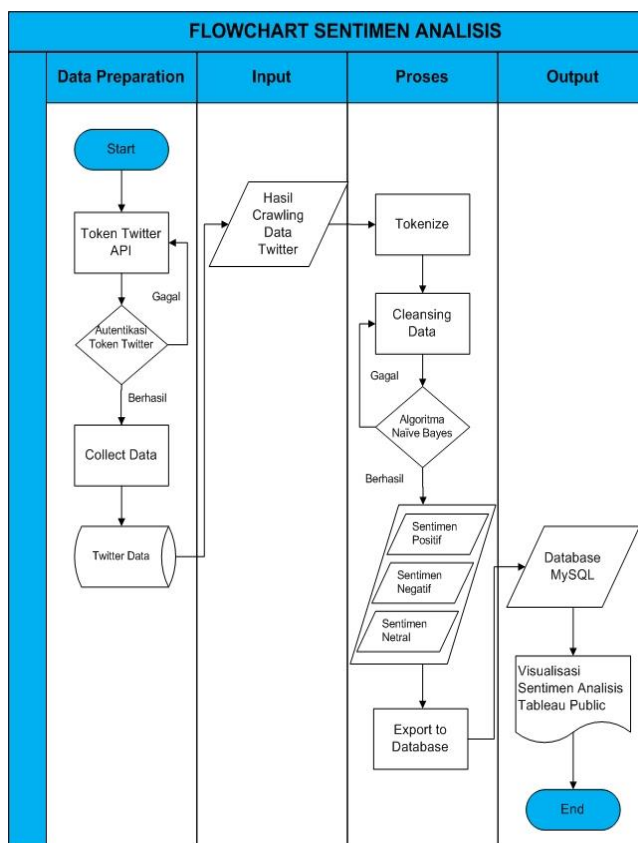
5. *Evaluation*

*Evaluation* ialah untuk menilai model agar dapat melaporkan efek yang diharapkan dari model. Setelah mempunyai model, harus ditentukan apakah data hasil olahan sesuai dengan tujuan.

6. *Deployment*

*Deployment* adalah untuk rencana penggunaan model. Perencanaan untuk *deployment* dimulai selama *Business Understanding* dan harus menggabungkan tidak hanya bagaimana untuk menghasilkan nilai model, tetapi juga bagaimana mengkonversi skor keputusan, dan bagaimana untuk menggabungkan keputusan dalam sistem operasional. Model tersebut dibangun dari data yang diwakili data pada waktu tertentu, sehingga perubahan waktu dapat menyebabkan berubahnya karakteristik data. Model pun harus dipantau dan mungkin diganti dengan model yang sudah diperbaiki.

Berikut adalah gambaran alur flowchart sentimen analisis kartu prakerja :



Gambar 2. Flowchart Sentimen Analisis

---

## HASIL DAN PEMBAHASAN

Hasil dari tanggapan atau sentimen analisis di media sosial twitter terhadap seleksi kartu prakerja akan sangat bermanfaat bagi pemerintah maupun masyarakat, untuk melihat sejauh mana program tersebut berjalan dengan baik dan tercapainya tujuan yang diharapkan terutama di lembaga pemerintahan yang telah menyelenggarakan program tersebut. Dapat ditentukan apakah layak untuk di lanjutkan atau tidak demi terciptanya SDM unggul Indonesia maju dengan SDM yang berkualitas sesuai keahlian di bidangnya masing-masing. Dengan kata lain mampu menjadi sistem informasi yang strategis untuk mengontrol jalannya program seleksi kartu prakerja yang saat ini masih banyak diperbincangkan oleh masyarakat demi menekan angka pengangguran di Indonesia sesuai tujuan yang diharapkan pemerintah. Informasi yang akan di buat adalah tampilan grafik dalam format *Tableau visualitation* yang merinci informasi mengenai jumlah total data yang didapat dari proses *crawling* data twitter, selama periode tanggal Februari 2021 s/d April 2021, persentase jumlah sentimen positif, netral dan negatif, pengambilan data secara time series dan informasi pengelompokan kata kedalam sentimen positif, negatif, dan netral.

Dari hasil pengumpulan data yang telah di peroleh terdapat data *posting twitter* sebanyak 45,122 data dari 125 keyword, data yang di dapat dari berbagai *topik keyword* mengenai sentimen kartu prakerja yang 10 diantaranya adalah sebagai berikut :

1. Kartu\_Prakerja jumlah data 7139
2. Anggaran\_Program\_Kartu\_Prakerja jumlah data 779
3. Bansos\_Prakerja jumlah data 1458
4. Bantuan\_Kartu\_Prakerja jumlah data 191
5. Program jumlah data 2295
6. Gelombang\_Kartu\_Prakerja jumlah data 2716
7. Bermasalah\_Kartu\_Prakerja jumlah data 780
8. Pendaftaran jumlah data 922
9. Pandemi\_Prakerja jumlah data 890
10. Insentif\_kartu\_Prakerja jumlah data 367

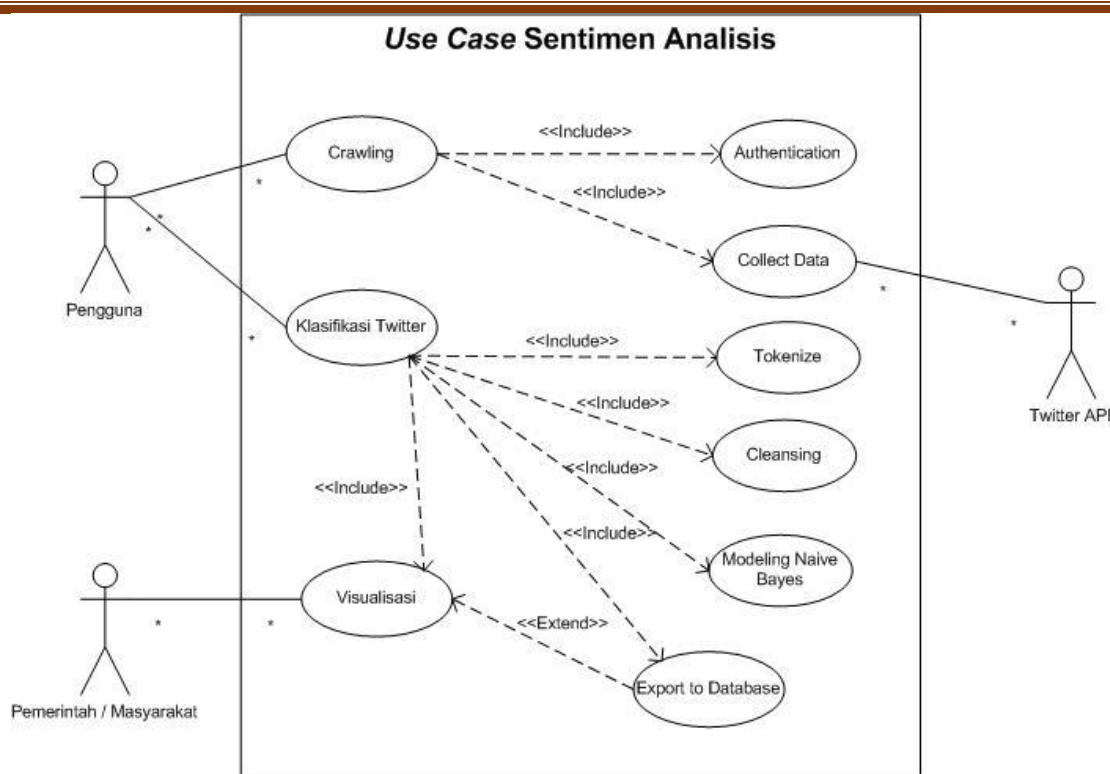
Adapun kolom *field* yang di dapat dari *twitter* yaitu *Text, favorited, favoriteCount, replyToSN, created, truncated, replyToSID, id, replyToUID, statusSource, screenName, retweetCount, isRetweet, retweeted, longitude, latitude, topik*. Berikut adalah data hasil dari pengumpulan data:

No	text	favorited	favoriteCount	replyToSN	created	truncated	replyToSID	id	replyToUID	statusSource	screenName	retweetCount	isRetweet	retweeted	longitude	latitude	topic
1	RT @rt	FALSE	0	#N/A	2/1/2021 21:22:03	FALSE	#N/A	135	#N/A	<a href="http://hellsome		1268	TRUE	FALSE	#N/A	#N/A	Kartu
2	RT @rt	FALSE	0	#N/A	2/1/2021 21:22:03	FALSE	#N/A	135	#N/A	<a href="http://hellsome		1268	TRUE	FALSE	#N/A	#N/A	Progra
3	RT @rt	FALSE	0	#N/A	2/1/2021 21:22:03	FALSE	#N/A	135	#N/A	<a href="http://hellsome		1268	TRUE	FALSE	#N/A	#N/A	Anggai
4	RT @rt	FALSE	0	#N/A	2/1/2021 21:22:03	FALSE	#N/A	135	#N/A	<a href="http://hellsome		1268	TRUE	FALSE	#N/A	#N/A	Berma
5	RT @rt	FALSE	0	#N/A	2/1/2021 21:22:03	FALSE	#N/A	135	#N/A	<a href="http://hellsome		1268	TRUE	FALSE	#N/A	#N/A	Praker
6	RT @rt	FALSE	0	#N/A	2/1/2021 21:22:03	FALSE	#N/A	135	#N/A	<a href="http://hellsome		1268	TRUE	FALSE	#N/A	#N/A	Pande
7	RT @rt	FALSE	0	#N/A	2/1/2021 21:22:03	FALSE	#N/A	135	#N/A	<a href="http://hellsome		1268	TRUE	FALSE	#N/A	#N/A	Ruang
8	RT @rt	FALSE	0	#N/A	2/1/2021 21:52:33	FALSE	#N/A	135	#N/A	<a href="http://kesamenkid		101	TRUE	FALSE	#N/A	#N/A	Praker
9	RT @rt	FALSE	0	#N/A	2/1/2021 22:12:49	FALSE	#N/A	135	#N/A	<a href="https://vostokmerah		1268	TRUE	FALSE	#N/A	#N/A	Kartu
10	RT @rt	FALSE	0	#N/A	2/1/2021 22:12:49	FALSE	#N/A	135	#N/A	<a href="https://vostokmerah		1268	TRUE	FALSE	#N/A	#N/A	Progra
11	RT @rt	FALSE	0	#N/A	2/1/2021 22:12:49	FALSE	#N/A	135	#N/A	<a href="https://vostokmerah		1268	TRUE	FALSE	#N/A	#N/A	Anggai
12	RT @rt	FALSE	0	#N/A	2/1/2021 22:12:49	FALSE	#N/A	135	#N/A	<a href="https://vostokmerah		1268	TRUE	FALSE	#N/A	#N/A	Berma
13	RT @rt	FALSE	0	#N/A	2/1/2021 22:12:49	FALSE	#N/A	135	#N/A	<a href="https://vostokmerah		1268	TRUE	FALSE	#N/A	#N/A	Praker
14	RT @rt	FALSE	0	#N/A	2/1/2021 22:12:49	FALSE	#N/A	135	#N/A	<a href="https://vostokmerah		1268	TRUE	FALSE	#N/A	#N/A	Pande
15	RT @rt	FALSE	0	#N/A	2/1/2021 22:12:49	FALSE	#N/A	135	#N/A	<a href="https://vostokmerah		1268	TRUE	FALSE	#N/A	#N/A	Ruang
16	RT @rt	FALSE	0	#N/A	2/1/2021 22:15:13	FALSE	#N/A	135	#N/A	<a href="http://r_dan_a		1268	TRUE	FALSE	#N/A	#N/A	Kartu
17	RT @rt	FALSE	0	#N/A	2/1/2021 22:15:13	FALSE	#N/A	135	#N/A	<a href="http://r_dan_a		1268	TRUE	FALSE	#N/A	#N/A	Progra
18	RT @rt	FALSE	0	#N/A	2/1/2021 22:15:13	FALSE	#N/A	135	#N/A	<a href="http://r_dan_a		1268	TRUE	FALSE	#N/A	#N/A	Anggai
19	RT @rt	FALSE	0	#N/A	2/1/2021 22:15:13	FALSE	#N/A	135	#N/A	<a href="http://r_dan_a		1268	TRUE	FALSE	#N/A	#N/A	Berma
20	RT @rt	FALSE	0	#N/A	2/1/2021 22:15:13	FALSE	#N/A	135	#N/A	<a href="http://r_dan_a		1268	TRUE	FALSE	#N/A	#N/A	Praker
21	RT @rt	FALSE	0	#N/A	2/1/2021 22:15:13	FALSE	#N/A	135	#N/A	<a href="http://r_dan_a		1268	TRUE	FALSE	#N/A	#N/A	Pande

**Gambar 3.** Hasil Pengumpulan Data

Pada tahap ini alur dari sentimen analisis di gambarkan dan jelaskan dalam bentuk *use case diagram* analisis. *Use case diagram* terlihat pada gambar 4. *Use case* sentimen analisis menggambarkan urutan proses sentimen mulai dari pengambilan data di *twitter*, proses algoritma dan visualisasi mengenai sentimen analisis pada sosial media *twitter* untuk menilai respon masyarakat terhadap seleksi program kartu prakerja yang di bahas dalam penelitian ini. Adapun permasalahan pada penelitian ini adalah belum adanya kejelasan untuk gelombang di bulan berikutnya pada tahun 2021 ini apakah akan di lanjutkan atau tidak. Berbagai kendala yang dihadapi baik calon peserta kartu prakerja maupun yang sudah menjadi peserta program kartu prakerja yang masih kebingungan dengan program tersebut karena minimnya sosialisasi yang didapat dan kurang tersampaikan oleh masyarakat yang antusias dalam mengikuti program tersebut, sehingga banyak menimbulkan berbagai sentimen di media sosial baik itu sentimen positif, negatif ataupun netral dari gelombang 1 sampai dengan gelombang 17 di tahun 2020-2021 ini.

Pada tahap berikutnya, hasil olahan data yang sudah melalui berbagai macam proses pengolahan data sebelumnya untuk selanjutnya akan di eksekusi dengan *algoritma naïve bayes*. Data yang perlu di tambahkan pada proses ini adalah *lexicon* kartu prakerja sebagai pemanggilan data untuk memunculkan hasil sentimen positif, negatif dan netral di simpan dalam 1 format *excel.xlsx* dengan diberi judul *hasil\_sentiment\_prakerja*.



**Gambar 4.** Use Case Sentimen Analisis

Tampilan yang akan dimunculkan berupa *dashboard* yang berisi grafis mengenai hasil penelitian yang telah selesai dan di visualisasikan melalui aplikasi program *Tableau*. *Dashboard Tableau* terdiri dari kumpulan *sheet* olahan grafis dan penjelasan tampilan akan dipaparkan berdasarkan *sheet* yang dibuat sebelum dijadikan bentuk *dashboard* dalam ukuran *desktop* standar.

Setiap *sheet* di desain sedemikian rupa untuk mempermudah bagi pemerintah dan masyarakat dalam memahami dan membaca grafis dari hasil *dashboard* visualisasi *tableau*. Agar visualisasai sentimen bisa di akses oleh pemerintah ataupun masyarakat maka perlu di buatkannya suatu media yang sifatnya publikasi, pada media tersebut akan di ditampilkan hasil-hasil dari pengolahan data sentimen yang sudah di visualisasikan melalui media *website tableau public*. Pada situs *web* tersebut di beri nama dengan Sentimen Analisis pada Sosial Media *Twitter* untuk menilai respon masyarakat terhadap seleksi Kartu Prakerja yang di tujukan kepada pemerintah maupun masyarakat mengenai jalannya program kartu prakerja tersebut.

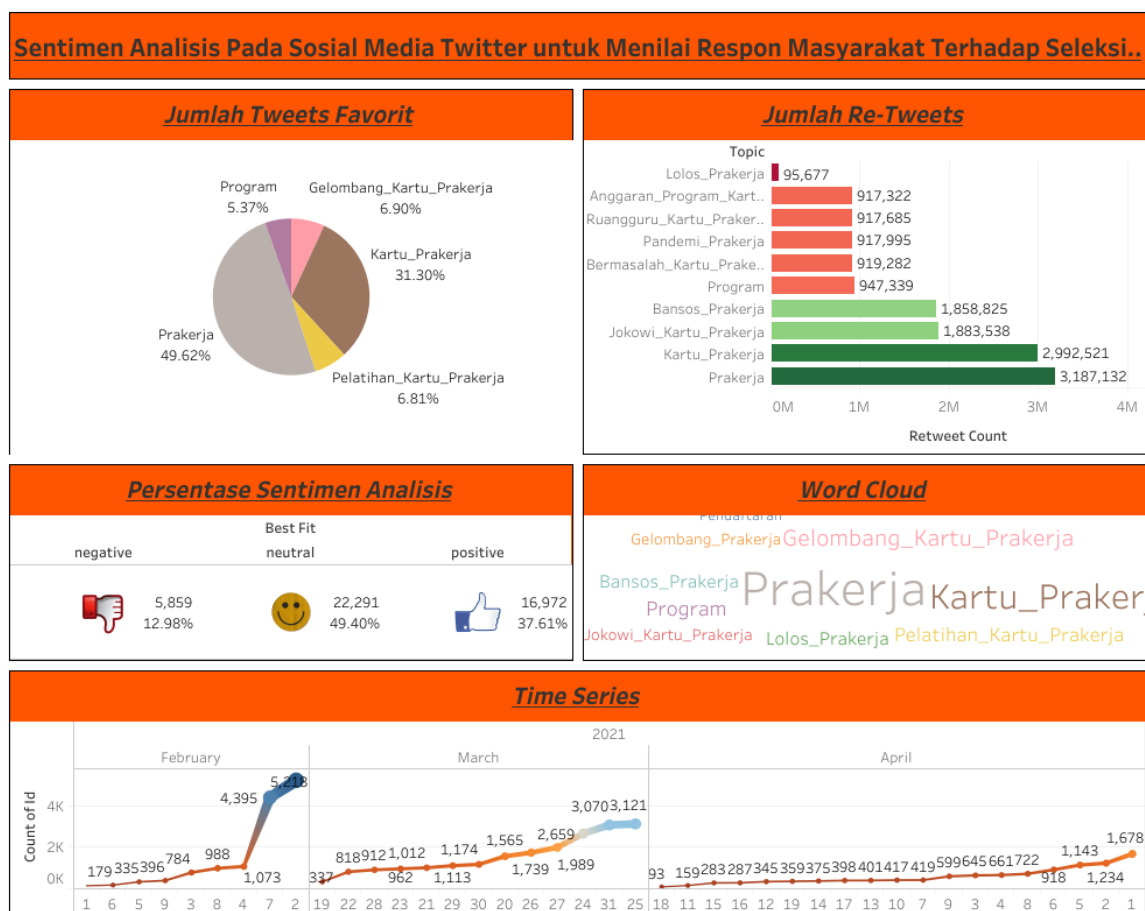
Untuk tampilan *dashboard* visualisasi sentimen analisis terhadap seleksi kartu prakerja terlihat pada gambar 5. Pada sheet pertama di aplikasi *tableau* yang menggambarkan jumlah *tweets favorite* terbanyak di twitter pada masa periode tersebut. Jumlah total *favorite count* sebanyak 19,398 untuk topik prakerja pada posisi pertama dengan persentase 49.62%, pada posisi kedua diperoleh sebanyak 12,235 untuk topik kartu prakerja dengan persentase



31.30%, pada posisi ke tiga diperoleh sebanyak 2,698 untuk topik gelombang\_kartu\_prakerja dengan persentase 6.90%, pada posisi ke empat diperoleh data sebanyak 2,663 untuk topik pelatihan\_kartu\_prakerja dengan persentase 6.81%, dan di posisi ke lima diperoleh sebanyak 2,098 untuk topik program dengan persentase 5,37%. Pada sheet ke dua di aplikasi tableau yang menggambarkan jumlah re-tweets terbanyak di aplikasi twitter pada masa periode tersebut. Jumlah total *re-tweets count* sebanyak 3,187,132 untuk topik prakerja pada posisi pertama, pada posisi kedua diperoleh sebanyak 2,992,521 dengan topik kartu\_prakerja, untuk posisi ke tiga diperoleh sebanyak 1,883,538 dengan topik jokowi\_kartu\_prakerja, untuk posisi ke empat diperoleh sebanyak 1,858,825 dengan topik bansos\_prakerja, untuk posisi ke lima diperoleh sebanyak 947,339 dengan topik program, untuk posisi ke enam diperoleh sebanyak 919,282 dengan topik bermasalah\_kartu\_prakerja, untuk posisi ke tujuh diperoleh sebanyak 917,995 dengan topik pandemi\_prakerja, untuk posisi ke delapan diperoleh sebanyak 917,685 dengan topik ruangguru\_kartu\_prakerja, untuk posisi ke sembilan diperoleh sebanyak 917,322 dengan topik anggaran\_program\_kartu\_prakerja, dan untuk posisi ke sepuluh diperoleh sebanyak 95,677 dengan topik lolos\_prakerja. Pada *sheet* ketiga untuk menunjukkan berapa persentase nilai sentimen dengan diklasifikasikan kedalam 3 kategori diantaranya kategori *positive*, *negative*, dan netral dengan menggunakan algoritma *Naïve Bayes*. Untuk sentimen *negative* dengan total id pengguna sebanyak 5,859 dengan besaran persentase 12,98%, untuk sentimen *neutral* dengan total id pengguna sebanyak 22,291 dengan besaran persentase 49,40%, dan terakhir untuk sentimen *positive* dengan total id pengguna sebanyak 16,972 dengan besaran persentase 37,61%.

Bulan pertama yaitu pada bulan Februari 2021, pada bulan Februari data di ambil secara acak mulai dari tanggal 1 s/d 9 Februari 2021 dari gambaran grafik *time series* pada gambar 5 disimpulkan bahwa di bulan Februari sentimen mengenai kartu prakerja terjadi lonjakan yang signifikan dan merupakan bulan yang paling ramai memperbincangkan mengenai program kartu prakerja dengan total id pengguna sebanyak 5,218 id pengguna pada tanggal 2 Februari 2021.

Bulan kedua yaitu bulan Maret 2021 di ambil dari tanggal 19 s/d 31 Maret 2021, pada bulan Maret sentimen mengenai program kartu prakerja di sosial media twitter masih banyak diperbincangkan pada bulan tersebut terjadi naik turun grafik yang menandakan bahwa sentimen mengenai program kartu prakerja masih ramai di bahas di akun sosial media twitter kicauan yang paling banyak di bahas terdapat pada tanggal 25 Maret 2021 dengan total id pengguna 3,121.



Gambar 5. Hasil Dashboard Visualisasi

Terakhir pada bulan April 2021 dari tanggal 1 s/d 19 april 2021 pembahasan di sosial media twitter mengenai program kartu prakerja sudah mulai melandai dimana jika dilihat dari grafik rata-rata kicauan dari sosial media twitter di bawah dari 1000 id pengguna, data kicauan twitter yang paling banyak hanya terdapat pada tanggal 1 april 2021 dengan total id pengguna 1,678.

Seluruh informasi ditampilkan secara visualisasi melalui situs web Tableau Public dengan URL:

<https://public.tableau.com/app/profile/arif.rahman7878/viz/VisualisasiSentimenAnalisisKartuPrakerja/Dashboard>

---

## KESIMPULAN DAN REKOMENDASI

Dari hasil analisis sentimen pada sosial media twitter untuk menilai respon masyarakat terhadap seleksi kartu prakerja maka dapat disimpulkan sebagai berikut :

1. Pemerintah pada tahun 2020 s/d 2021 ini sedang menyelenggarakan program kartu prakerja dengan tujuan dan harapan terciptanya SDM yang unggul untuk Indonesia maju di masa yang akan datang guna menekan angka pengangguran di Indonesia. Berbagai upaya pemerintah telah dilakukan demi lancarnya keberlangsungan program tersebut. Dengan melihat respon pengguna Twitter terhadap sentimen analisis program kartu prakerja yang menjadi sumber data penelitian, dapat dirancang dan divisualisasikan *dashboard* sentimen analisis mengenai program kartu prakerja. Adapun hasil dari sentimen analisis kartu prakerja didapatkan sebanyak 22.291 total id pengguna dengan besaran persentase 49.40% yang menyatakan netral, dalam artian bahwa masyarakat sebagian besar tidak pro dan kontra atas program kartu prakerja tersebut.
2. Visualisasi sentimen analisis prakerja disimpan dalam format *Tableau* yang memuat jumlah total tweets favorit dengan menampilkan 5 tweets favorit teratas berdasarkan *favorite\_coun*, jumlah re-tweets dengan menampilkan 10 re-tweets teratas berdasarkan topik, persentase jumlah sentimen positif, netral dan negative berdasarkan *count\_of\_hasilsentiment / count\_id*. *Word cloud* menampilkan deretan 10 topik yang paling banyak dibicarakan di sosial media twitter berdasarkan *topic\_count*. *Time series* menampilkan tahun bulan dan tanggal pengambilan data di twitter pada periode bulan Februari, Maret, April tahun 2021 dalam bentuk grafik berdasarkan id pengguna.

## REFERENSI

- Alfikir, Muhammad Zunan. (2020). Strategi Algoritma Analisis Sentimen Twitter Terhadap Kartu Prakerja Di Tengah Pandemi COVID-19 Menggunakan Algoritma Pencocokan String Dan Library TextBlob. Bandung. <https://informatika.stei.itb.ac.id/~rinaldi.munir/Stmik/2019-2020/Makalah/Makalah-Stima-2020-031.pdf>.
- Kementerian Koordinator Bidang Perekonomian Republik Indonesia. (2021). Tentang Kartu Prakerja. <https://www.prakerja.go.id/tentang-prakerja> (August 30, 2021).
- Mustaqim, Ilmawan, and Nanang Kurniawan. (2017). Pengembangan Media Pembelajaran Berbasis Augmented Reality. *Jurnal Edukasi Elektro* 1(1): 36–48.
- Nurzahputra, Aldi, and Much Aziz Muslim. (2016). “Analisis Sentimen Pada Opini Mahasiswa Menggunakan Natural Language Processing.” In *Seminar Nasional Ilmu Komputer (SNIK)*, Semarang: Jurusan Ilmu Komputer, Universitas Negeri Semarang,

114–18.

[https://ilkom.unnes.ac.id/snik/prosiding/2016/18.SNIK\\_366\\_Analisis\\_Sentimen.pdf](https://ilkom.unnes.ac.id/snik/prosiding/2016/18.SNIK_366_Analisis_Sentimen.pdf).

Sucahyo, Nur, Lela Nurlaela, and Robbi Rekto Waryono. (2021). Analisis Sentimen Masyarakat Jakarta Terhadap Kebijakan Perluasan Dan Perpanjangan Ganjil Genap Di Media Sosial Twitter. *Jurnal Teknologi Informatika dan Komputer* 7(2): 97–111. <http://journal.thamrin.ac.id/index.php/jtik/article/view/506>.

Sudiantoro, Adhi Viky, and Eri Zuliarso. (2018). Analisis Sentimen Twitter Menggunakan Text Mining Dengan Algoritma Naive Bayes Classifier. *Dinamika Informatika* 10(2): 69–73.

Suhanda, Yogasetya, Ike Kurniati, and Siti Norma. (2020). “Penerapan Metode Crisp-DM Dengan Algoritma K-Means Clustering Untuk Segmentasi Mahasiswa Berdasarkan Kualitas Akademik.” *Jurnal Teknologi Informatika dan Komputer MH Thamrin* 6(2): 12–20.