

The Development of Social Media Intelligence System for Citizen Opinion and Perception Analysis over Government Policy

Rancang Bangun Sistem Social Media Intelligence untuk Analisis Opini dan Persepsi Masyarakat terhadap Kebijakan Pemerintah

Muhammad Habibi¹, Muhammad Rifqi Ma'arif², Dayat Subekti³

^{1,3} Informatika, Universitas Jenderal Achmad Yani Yogyakarta, Indonesia

² Teknik Industri, Universitas Jenderal Achmad Yani Yogyakarta, Indonesia

^{1*} muhammadhabibi17@gmail.com, ² rifqi@unjaya.ac.id, ³ dayat.subekti@gmail.com

Article's Information / Informasi Artikel

Received: December 2021

Revised: January 2021

Accepted: January 2021

Published: February 2022

Abstract

Objective: Uploads on social media regarding public perceptions of the implementation of a policy in the field are precious input for the authorities to map problems in the area quickly and accurately. This study aims to produce a platform to implement a computational approach called Social Media Intelligence.

Methods: This study proposes a computational framework for end-to-end automating the process of utilizing social media data, from data acquisition and organization to information analysis and visualization. The social media data is carried out by topic modeling, and then a sentiment analysis process is carried out.

Results: This research produces a social media intelligence platform that can be used to monitor public opinion regarding public perceptions related to the implementation of policies issued by the government, especially local governments in Indonesia. Based on the results of modeling the topic of Covid-19 vaccination cases, 11 topics of discussion were obtained. While the sentiment analysis results of the 11 topics resulted, topic 6 had the most negative sentiment value, namely regarding the development of Covid-19 vaccination in Indonesia. At the same time, topics with the most positive sentiment values are topic three and topic 10. These topics discuss the vaccination process carried out by health procedures (prokes) and government policies related to COVID-19 vaccination.

Authenticity: The developed Social Media Intelligence platform focuses on analyzing public opinion regarding government policies. The application accommodates the

community's aspirations regarding implementing guidelines issued by the government.

Keywords: Sentiment Analysis; Topic Modeling; Machine Learning
Kata kunci: Analisis sentimen;
Pemodelan topik; Machine Learning

Abstrak

Tujuan: Unggahan-unggahan di media sosial mengenai persepsi masyarakat terhadap implementasi suatu kebijakan di lapangan menjadi masukan yang sangat berharga bagi pihak yang berwenang untuk dapat memetakan masalah yang ada di lapangan secara cepat dan akurat. Penelitian ini bertujuan menghasilkan sebuah platform untuk mengimplementasikan pendekatan komputasional yang kemudian disebut dengan *Social Media Intelligence*.

Metode: Penelitian ini mengusulkan suatu kerangka kerja komputasional untuk secara *end-to-end* mengotomatiskan proses pemanfaatan data media sosial mulai dari akuisisi dan pengorganisasian data hingga analisis dan visualisasi informasi. Data media sosial dilakukan pemodelan topik (*topic modeling*) kemudian dilakukan proses analisis sentimen (*sentiment analysis*).

Hasil: Penelitian ini menghasilkan sebuah platform *social media intelligence* yang dapat digunakan untuk memantau opini publik mengenai persepsi masyarakat terkait dengan implementasi kebijakan-kebijakan yang dikeluarkan oleh pemerintah, khususnya pemerintah daerah di Indonesia. Berdasarkan hasil pemodelan topik kasus vaksinasi Covid-19 didapatkan 11 topik pembicaraan. Sedangkan hasil analisis sentimen dari 11 topik tersebut menghasilkan, topik 6 memiliki nilai sentiment negatif paling banyak yaitu mengenai perkembangan Vaksinasi Covid-19 di Indonesia. Sedangkan topik yang memiliki nilai sentiment positif paling banyak adalah topik 3 dan topik 10. Topik tersebut membahas mengenai proses vaksinasi yang dilakukan secara prosedur Kesehatan (prokes) dan Kebijakan pemerintah terkait vaksinasi COVID-19.

Keaslian: Platform *Social Media Intelligence* yang dikembangkan berfokus pada analisis opini publik terkait kebijakan pemerintah. Aplikasi tersebut dapat menampung aspirasi masyarakat terkait implementasi kebijakan-kebijakan yang dikeluarkan oleh pemerintah.

1. Pendahuluan

Dalam penyusunan maupun implementasi suatu kebijakan publik yang dikeluarkan oleh pemerintah, pelibatan opini dan persepsi masyarakat memainkan peran yang sangat vital terhadap ketercapaian tujuan dari kebijakan tersebut [1]. Kurangnya partisipasi publik dalam berbagai kebijakan membuat keputusan yang digunakan menjadi penyebab utama

ketidakaauratan terutama ketika merumuskan dan mengevaluasi kebijakan tersebut [2]. Saat ini di Indonesia, pelibatan masyarakat dalam perencanaan pembangunan dan perumusan kebijakan publik pada umumnya sudah dilakukan namun dengan skala yang terbatas. Metode biasa dilakukan adalah dengan teknik riset konvensional seperti survey, observasi dan sejenisnya. Namun, kegiatan penjangkaran aspirasi publik secara akurat dan komprehensif bukanlah pekerjaan yang mudah dan dapat selesai dalam waktu singkat apabila dikerjakan dengan pendekatan-pendekatan konvensional seperti survey, observasi lapangan dan sebagainya. Selain memerlukan waktu yang relatif lama, metode-metode tersebut tentunya membutuhkan sumber daya manusia dan finansial yang tidak sedikit ketika jumlah populasinya besar. Sehingga, pemantauan persepsi publik secara luas, masif dan dalam waktu yang cepat sangat tidak *feasible* bila dilakukan dengan pendekatan-pendekatan konvensional tersebut.

Alternatif solusi yang lebih ekonomis dan cepat dalam menyerap aspirasi masyarakat adalah dengan menganalisis unggahan masyarakat di media sosial. Saat ini, media sosial dengan segala fiturnya yang terus berkembang telah menjadi bagian integral dalam kehidupan bermasyarakat, di segala lapisan [3], [4]. Media sosial telah memberikan kemudahan bagi masyarakat dunia untuk mendiskusikan topik-topik yang menjadi minat bersama seperti olahraga, hiburan, dan termasuk politik. Melalui sosial media, secara umum, masyarakat mengunggah sesuatu mengenai pemerintah, politik atau kebijakan baik secara sengaja maupun tidak [5]. Sehingga, data media sosial yang berasal dari unggahan-unggahan masyarakat menjadi sumber informasi yang sangat bernilai untuk mendapatkan *feedback* atas suatu kebijakan publik secara lebih komprehensif untuk merumuskan kebijakan publik memberikan layanan yang lebih baik kepada masyarakat [6], [7].

Dalam arti yang lebih luas, media sosial adalah platform komunikasi di internet yang memungkinkan interaksi dua arah dan interaktif diantara para penggunanya [8]. Media sosial tidak hanya terbatas pada *platform-platform* yang sudah terkenal seperti Facebook, Instagram, Twitter dan sebagainya, namun semua platform daring yang memungkinkan interaksi dua arah dan dapat diakses secara publik dapat dikategorikan sebagai media sosial seperti forum diskusi, forum tanya jawab, dan sebagainya [9]. Meskipun demikian tidak dapat dipungkiri bahwa jejaring sosial populer seperti Facebook, Twitter, Instagram merupakan platform jejaring sosial yang mampu menghubungkan orang-orang di dunia dalam jumlah yang sangat besar. *Platform-platform* tersebut telah secara sangat signifikan mengubah cara orang mengakses informasi, dan bagaimana informasi tersebut tersebar luas. Menurut studi yang dilakukan secara daring oleh Reuters pada tahun 2016 [10] terhadap 50.000 orang di 26 negara, *platform-platform* jejaring sosial populer tersebut merupakan sumber informasi utama untuk 51% responden, dan 12% diantaranya menjadikannya sebagai sumber informasi pertama.

Dengan daya jangkauannya yang luas ke berbagai macam dan lapisan kelompok masyarakat, media sosial menjadi sarana yang sangat efektif untuk saling berinteraksi, berbagi informasi, hingga membentuk komunitas/kelompok di dunia maya yang didasari dari kesamaan minat/hobi, profesi bahkan pandangan politik [11]. Masifnya penggunaan media sosial membuatnya memiliki pengaruh yang sangat signifikan dalam kehidupan masyarakat secara sosial [12], [13]. Peristiwa pergerakan masa yang sangat besar seperti Arab Springs sangat dipengaruhi oleh berbagai informasi yang tersebar di media sosial [14]. Dalam konteks partisipasi masyarakat dalam politik dan pemerintahan, media sosial perlahan namun pasti telah digunakan secara intensif untuk membaca opini masyarakat melalui tulisan-tulisan yang mereka

unggah [15]. Berbagai *platform* media sosial dapat dilihat sebagai sebuah *collective intelligence* dalam sebuah wadah kolaborasi antar elemen masyarakat secara daring [16]. Sehingga, kehadiran media sosial telah meningkatkan partisipasi aktif masyarakat untuk menentukan bagaimana pemerintahan berjalan hingga saat ini [17].

Dalam beberapa tahun terakhir ini, pola penggunaan media sosial oleh kalangan pemerintah bergeser dari sekedar media distribusi informasi searah menjadi upaya untuk memperkuat interaksi, transparansi dan *engagement* antara pemerintah dengan masyarakat [18]. Penelitian tentang pemanfaatan media sosial dalam pemerintahan, khususnya dari aspek analitik sejauh ini terbagi menjadi dua kategori. Kategori yang pertama adalah melakukan pengukuran efektifitas pemanfaatan media sosial melalui berbagai indikator-indikator kuantitatif seperti jumlah likes, jumlah *share*, jumlah *follower*, jumlah *comment* dan sebagainya. [19] tersebut merumuskan metrik yang disebut *citizen engagement index* untuk mengukur *engagement* pada informasi yang diunggah pemerintah di media sosial. Metrik tersebut terdiri atas tiga dimensi tersebut adalah *popularity*, *commitment* dan *virality* yang dikalkulasi dari jumlah *like*, *follower*, *share/retweet* dan *comment* masyarakat pada akun media sosial yang dimiliki suatu instansi pemerintahan. Metrik *citizen engagement index* ini kemudian menjadi landasan untuk penelitian-penelitian yang lebih aplikatif terkait dengan analitik media sosial, khususnya antara pemerintah dengan masyarakat [20]–[22]. Diluar metrik tersebut beberapa penelitian langsung mengkalkulasi tiga aspek eksplisit yakni *comment*, *likes* dan *share* untuk menghitung *engagement* [23], [24] maupun *comment* dan *share* saja [25] untuk mengkalkulasi *engagement* masyarakat pada akun media sosial instansi pemerintahan.

Kategori yang kedua untuk penelitian mengenai analitik media sosial, terkait dengan analisis pada konten unggahan. Dari sisi konten beberapa hal yang menjadi aspek analisis diantaranya adalah jenis dan topik konten yang diunggah dan pengaruhnya pada tingkat *engagement* [26], [27] serta sentimen dan emosi yang diekspresikan publik sebagai respon atas informasi yang bersumber dari akun media sosial instansi masyarakat [28], [29]. Dalam tataran yang lebih praktis, yakni studi kasus penggunaan media sosial untuk pemerintahan dan kebijakan publik, terdapat beberapa penelitian yang sudah dilakukan di berbagai negara diantaranya adalah Korea Selatan dan Jepang [30], Uni Eropa [31], Amerika Serikat [32], [33] dan berbagai negara lain. Secara lebih khusus, *social media analytics* juga digunakan secara intensif untuk melakukan analisis kebijakan publik yang berkaitan dengan penanganan pandemi Covid-19 di berbagai negara di dunia [34], [35].

Tantangan yang kemudian muncul dari upaya pemanfaatan media sosial ini diantaranya mencakup tiga hal. Yang pertama adalah, bagaimana mendeteksi percakapan publik di media sosial yang terkait dengan respon terhadap suatu kebijakan publik. Masalah kedua adalah bagaimana data media sosial yang berjumlah besar dan tidak terstruktur (teks yang tidak baku) dapat diolah untuk mendapatkan hasil analisis secara real-time dan akurat. Dan yang ketiga adalah bagaimana agar hasil analisis dapat diakses dalam bentuk yang mudah dimengerti oleh pengambil/perumus kebijakan di instansi pemerintahan. Sebagai solusi atas ketiga masalah tersebut, penelitian ini mengusulkan suatu kerangka kerja komputasional untuk secara *end-to-end* mengotomatiskan proses pemanfaatan data media sosial mulai dari akuisisi dan pengorganisasian data hingga analisis dan visualisasi informasi. Tujuan dari penelitian ini adalah menghasilkan sebuah platform untuk mengimplementasikan pendekatan komputasional yang kemudian disebut dengan *Social Media Intelligence*. Melalui platform tersebut, instansi

pemerintahan dapat memperoleh *insight* secara real-time dari unggahan masyarakat di media sosial mengenai dinamika dan permasalahan yang muncul atas implementasi suatu kebijakan.

2. Metode

Rancang bangun model pengembangan produk berbasis *social media analytic* yang diusulkan dalam penelitian ini tersusun dari dua latar belakang teoritis (*theoritical background*), yakni pemodelan topik (*topic modeling*) dan analisis sentimen (*sentiment analysis*). Secara umum pada tahap pertama, dari data media sosial yang terkumpul (tahap 1), topik-topik pembicaraan yang terkait dengan produk akan diidentifikasi dengan menggunakan teknik pemodelan topik. Pada tahapan yang ketiga, nilai kepuasan pelanggan pada setiap topik akan dianalisis dengan menggunakan pendekatan analisis sentimen.

2.1. Rancangan Modul Akuisisi Data

Tahap yang pertama adalah pengumpulan dan pembersihan (*pre-processing*) data media sosial yang memiliki keterkaitan dengan produk yang akan dikembangkan/diperbaiki. Dengan menggunakan kata kunci tertentu yang relevan, konten yang akan diakuisisi mencakup unggahan customer pada *platform* media sosial dan review produk pada media online. Dalam tahap ini, beberapa teknik yang akan digunakan mencakup *web scrapping* [36] dan pemrograman *web service* pada *platform* yang menyediakan *Application Programming Interface* (API). Selanjutnya setelah data dikumpulkan maka data akan dibersihkan dengan menggunakan beberapa pendekatan data *preprocessing* [37].

2.2. Identifikasi Topik beserta Level Signifikansinya

Dalam tahapan ini, akan diidentifikasi fitur-fitur produk yang banyak direview oleh pengguna media sosial beserta dengan level significansinya. Proses identifikasi fitur produk dari unggahan/review pengguna dilakukan dengan menggunakan *topic modeling*. Dalam pendekatan *topic modeling* ini, eksperimen akan dilakukan dengan menggunakan algoritma Latent Dirichelet Allocation (LDA) [38]. LDA pendekatan parametric, kedua pendekatan tersebut memerlukan masukan (input) berupa jumlah topik yang diidentifikasi. Dengan asumsi bahwa jumlah topik yang menghasilkan model topik optimal belum diketahui, maka eksperimen akan dilakukan pada berbagai konfigurasi jumlah topik. Untuk menentukan hasil yang paling optimal dari jumlah topik yang dieksperimenkan, maka akan digunakan metode Elbow [39]. Topik-topik yang berhasil diidentifikasi ini kemudian akan difilter dengan menggunakan kata kunci yang sesuai untuk mendapatkan topik pembicaraan yang terkait dengan fitur produk. Tingkat signifikansi dari masing-masing fitur didapatkan dengan menggunakan pendekatan *bag of words* yakni TF/IDF [40], yang dijalankan pada term *matrix* yang menyusun model topik yang dihasilkan oleh algoritma LDA.

2.3. Mengukur Tingkat Kepuasan Publik

Langkah selanjutnya adalah menghitung tingkat kepuasan *customer* untuk masing-masing fitur produk yang sudah diidentifikasi. Pendekatan yang diterapkan untuk tahapan ini adalah dengan menggunakan analisis sentimen. Analisis sentimen merupakan metode *supervised* yang memerlukan data latih (*training*) sebagai pembentukan model klasifikasi sentimen. Dalam

penelitian ini akan digunakan dua data latih yakni data latih primer dan sekunder. Data latih primer berasal dari data unggahan media sosial terkait review produk yang sudah diberikan label sentimen secara manual oleh *expert*. Sedangkan data sekunder sebagai penguat didapatkan dari data sentimen dari unggahan umum yang sudah ada dan bisa diakses oleh umum [41]. Algoritma yang digunakan untuk analisis sentimen ini adalah *Support Vector Machine (SVM)* [42].

3. Hasil dan Pembahasan

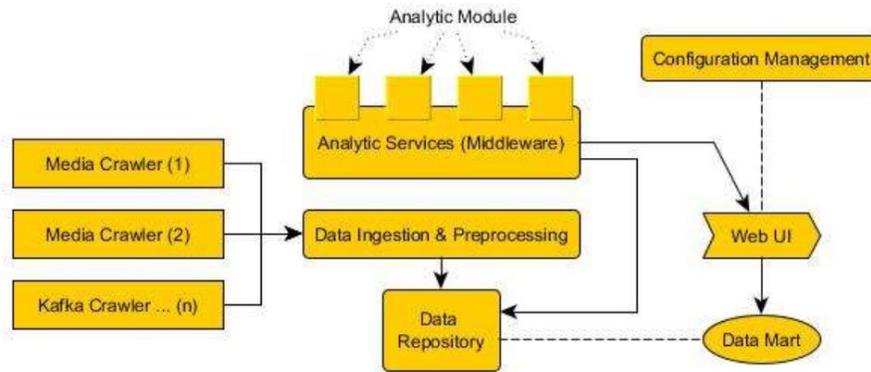
3.1. Arsitektur Sistem

Gambar 2 merupakan desain arsitektur sistem *social media intelligent* yang diusulkan dalam penelitian ini. Secara garis besar, arsitektur sistem terdiri dari 4 bagian utama yakni *Media Crawler*, *Data Ingestion & Preprocessing*, *Data Repository*, *Analytic Services*, dan *Web UI* dengan fungsi masing-masing bagian secara lebih mendetail sebagai berikut:

1. **Media Crawler.** Bagian ini terdiri dari beberapa program crawler yang berfungsi untuk pengambilan/akuisisi data dari media daring (*online*) berupa *platform* jejaring sosial. *Crawler-crawler* yang terdapat dalam bagian ini memiliki peran yang berbeda dan spesifik satu sama lain. Satu *crawler* didesain/diprogram untuk mengakuisisi data dari satu jenis media daring saja. Lebih lanjut *crawler-crawler* ini bekerja terus menerus secara paralel, sehingga setiap keterbaharuan (*update*) pada setiap platform media daring dapat di-*capture* data/informasinya.
2. **Data Ingestion & Preprocessing.** Fungsi pertama dari modul ini untuk menyatukan aliran data dari semua *crawler* yang bekerja pada bagian *Media Crawler*. Fungsi yang kedua adalah sebagai alat untuk melakukan pembersihan dan standarisasi data (*preprocessing*). Data yang berasal dari *crawler* adalah data mentah yang masih mengandung banyak noise serta memiliki format dan struktur yang berbeda-beda tergantung dari platform media daring yang menjadi sumber data. Sehingga pada modul ini data tersebut dibersihkan dan ditransformasikan kedalam format yang seragam sebelum disimpan di tempat penyimpanan data (*Data Repository*).
3. **Data Repository.** Modul ini berfungsi sebagai tempat penyimpanan data. Pendekatan yang digunakan dalam mengorganisir data dalam *repository* yang digunakan dalam arsitektur sistem adalah konsep NoSQL. Konsep NoSQL merupakan konsep pengorganisasian data yang *less schema* sehingga cocok untuk karakteristik data media daring yang berupa teks yang tidak terstruktur (*unstructured*).
4. **Analytic Services.** Bagian ini memiliki fungsi utama untuk melakukan pengolahan data media daring yang dikumpulkan menjadi informasi-informasi yang bisa memberikan insight, khususnya bagi peneliti sosial yang ingin memanfaatkan data media daring. Bagian ini dibangun secara modular dengan pendekatan *microservice*. Dengan pendekatan ini, setiap jenis *analytic* akan dikembangkan dalam modul-modul yang lebih kecil yang memungkinkan dikembangkan secara independen. Untuk merealisasikan pendekatan ini, maka bagian ini dikembangkan dengan teknologi *middleware* yang dapat mengorganisasikan fungsi dari modul-modul analitik yang dikembangkan secara terpisah. *Middleware* juga berperan sebagai penghubung modul analitik untuk

melakukan query/akses data ke *Data Repository*. Setiap modul *analytic* akan memiliki standard luaran data yang sama, yakni dalam representasi JSON (*JavaScript Object Notation*). Luarannya data yang terstandarisasi tersebut akan diakses dalam bentuk layanan/*services* oleh Modul Web UI untuk menampilkan visualisasi data. Pelaksanaan *services* tersebut memanfaatkan protokol komunikasi REST (*Representational State Transfer*).

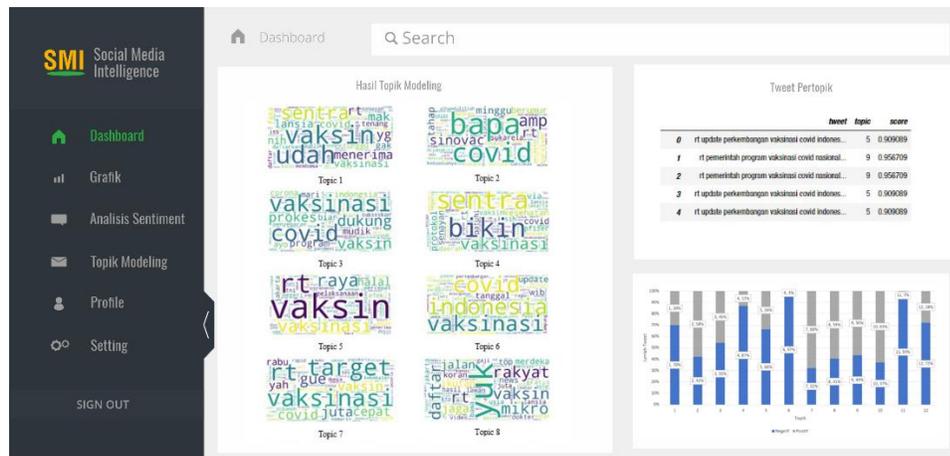
5. **Web UI.** Bagian selanjutnya dari arsitektur sistem adalah Web UI. Fungsi utama dari Web UI adalah sebagai media antar muka antara pengguna (*user*) dengan sistem/aplikasi Medi@n. Web UI akan melakukan pengaksesan data hasil olahan modul analitik melalui protokol REST. Data yang didapatkan kemudian akan diproyeksikan kedalam representasi visual melalui berbagai bentuk grafik/diagram sesuai dengan format dan maksud dari data yang diperoleh.



Gambar 2. Arsitektur sistem *social media intelligence*

3.2. Sistem *Social Media Intelligence*

Gambar 3 merupakan hasil implementasi sistem *Social Media Intelligence* yang dikembangkan. Sistem ini berfungsi sebagai *dashboard* untuk melihat hasil analitik yang sudah dilakukan terhadap data media sosial.



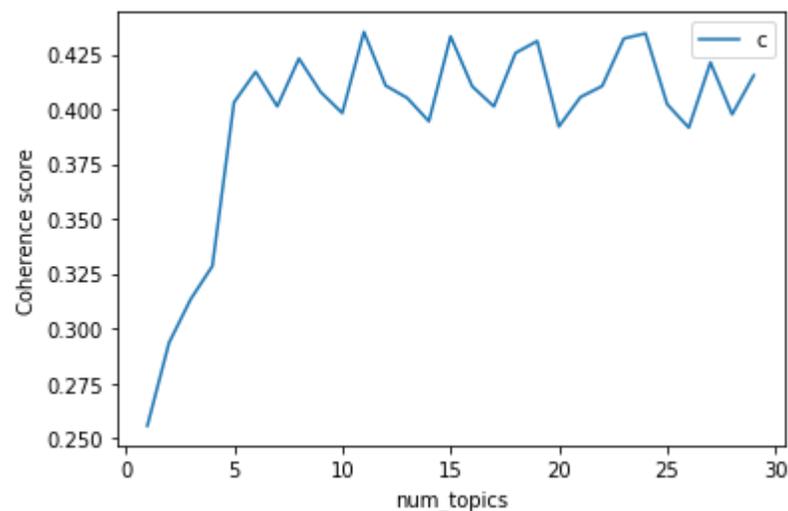
Gambar 3. Sistem *social media intelligence*

3.3. Studi Kasus Pemanfaatan Sistem *Social Media Intelligence*

Pada bagian ini akan dibahas hasil studi kasus pemanfaatan *Sistem Social Media Intelligence* yang telah dilakukan. *Dataset* yang digunakan dalam penelitian ini diambil dalam kurun waktu yaitu dari tanggal 27 Maret 2021 sampai 11 Juli 2021 sebanyak 8.279 data *tweet*. Data *tweet* yang terkumpul menggunakan katakunci #VaksinasiNasional atau #vaksinasinasional.

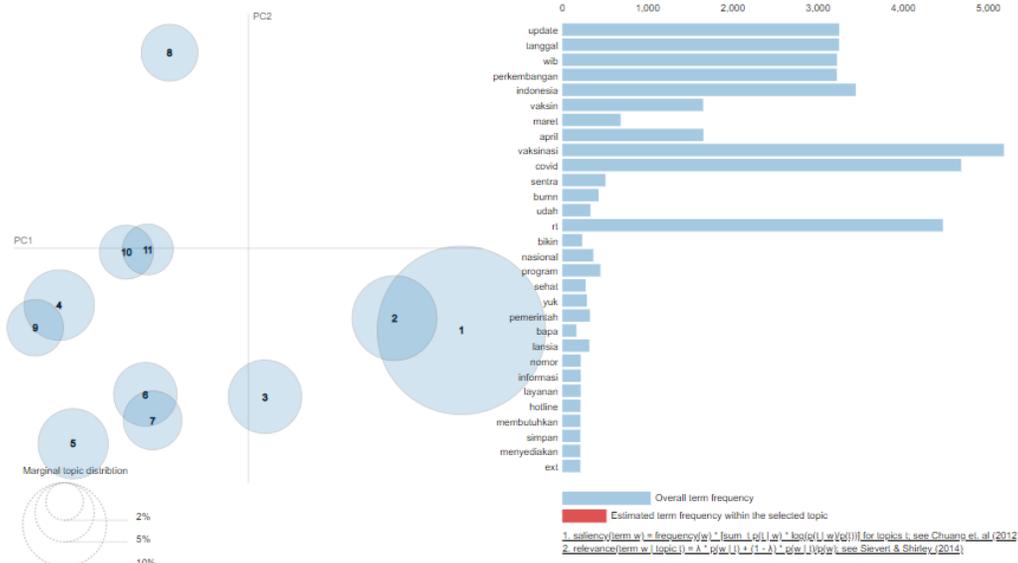
3.3.1. Pemodelan Topik

Dalam percobaan ini, kami melakukan identifikasi topik dengan data studi kasus vaksinasi Covid-19. Identifikasi topik yang dilakukan yaitu dengan menggunakan algoritma Latent Dirichlet Allocation (LDA). Pada tahapan ini dibuat visualisasi dari segmen topik dengan jumlah topik yang ditentukan di awal sebanyak 30 topik. Setelah membangun segmen topik dengan rentang yang telah ditentukan, kemudian melakukan evaluasi terhadap jumlah segmen topik yang paling sesuai untuk analisis yang lebih lanjut. Penelitian ini menggunakan *topic coherence* untuk menentukan jumlah segmen topik yang paling sesuai. Pengukuran ini membantu membedakan antara topik-topik yang dapat ditafsirkan secara semantik dan topik-topik yang merupakan hasil interpretasi manusia [43]. *Topic Coherence* merupakan cara lain untuk mengevaluasi model topik dengan jaminan yang jauh lebih tinggi pada interpretasi manusia. Gambar 4 menunjukkan hasil pengukuran jumlah kesesuaian topik menggunakan *topic coherence*.



Gambar 4. Topic Coherence

Berdasarkan Gambar 4 dapat diketahui bahwa jumlah segmen topik yang paling sesuai adalah 11 segmen topik dengan skor *topic coherence* 0.435. Oleh karena itu, pada analisis selanjutnya menggunakan 11 segmen topik. Gambar 5 merupakan visualisasi peta jarak antar topik yang dihasilkan untuk 11 segmen topik.



Gambar 5. Peta Jarak antar topik

Untuk analisis konten, kami membuat model visualisasi *wordcloud*, yaitu *wordcloud* dari setiap segmen dalam 11 segmen topik. Visualisasi *wordcloud* akan memberikan gambaran tentang apa saja kata-kata yang sering digunakan dalam percakapan tertentu. Ini berarti, kita dapat melihat kata-kata apa yang paling signifikan yang membangun segmen tertentu dari model yang dibangun. Gambar 6 merupakan *wordcloud* pada 11 segmen topik.



Topik 1



Topik 2



Topik 3



Topik 4



Topik 5



Topik 6



Topik 7



Topik 8



Topik 9



Topik 10



Topik 11

Gambar 6. Wordcloud 11 segmen topik

Berdasarkan hasil word cloud pada Gambar 6, maka kami membuat analisis topik untuk masing-masing segmen topik. Analisis topik dibuat berdasarkan hubungan kata-kata yang sering muncul pada masing-masing segmen topik. Hasil analisis topik dapat dilihat pada Tabel 1.

Tabel 1. Analisis Topik pada setiap segmen topic

Segmen Topik	Kata Kunci	Analisis Topik
Topic 1	vaksin, udah, sentra, menerima, rt, mak, yg, lansia, covid, nih	Vaksinasi lansia
Topic 2	bapa, covid, amp, sinovac, rt, minggu, tahap, berumur, sukarela, dijangkiti	Vaksinasi sinovac
Topic 3	vaksinasi, covid, vaksin, dukung, prokes, program, mudik, indonesia, biar, ayo	Vaksinasi secara prokes
Topic 4	bikin, sentra, vaksinasi, kesehatan, protokol, senayan, covid, rt, via, pfizer	Sentra Vaksinasi Covid
Topic 5	vaksin, rt, vaksinasi, raya, halal, pelaksanaan, efek, aman, jakarta, negeri	Vaksin Halal
Topic 6	indonesia, covid, vaksinasi, update, tanggal, wib, perkembangan, rt, april, mei	Perkembangan Vaksinasi Covid
Topic 7	vaksinasi, rt, target, gue, juta, covid, vaksin, cepat, yah, rabu	Target Vaksinasi
Topic 8	yuk, vaksin, rt, rakyat, daftar, mikro, jalan, jaga, ikutan, merdeka	Seruan untuk Vaksinasi
Topic 9	vaksin, rt, lansia, vaksinasi, covid, aman, divaksin, yuk, divaksinasi, astrazeneca	Vaksinasi astrazeneca
Topic 10	maret, vaksinasi, covid, bumh, rt, nasional, program, sehat, pemerintah, nomor	Kebijakan pemerintah terkait Vaksinasi
Topic 11	vaksinasi, rt, tni, ri, panglima, meninjau, sentra, ikuti, kota, penularen	Peninjauan Vaksinasi oleh panglima TNI

3.3.2. Analisis Sentimen

Setelah ditemukan topik-topik yang terkandung dalam kasus vaksinasi Covid-19, Langkah berikutnya adalah melakukan analisis sentimen terhadap setiap topik. Analisis sentimen yang dilakukan yaitu dengan mengklasifikasikan sentimen kedalam dua kelas, yaitu kelas positif dan negatif. Sehingga pada setiap topik akan diketahui seberapa banyak masyarakat yang memiliki sudut pandang positif atau negatif.

Analisis sentimen dilakukan dengan menggunakan algoritma Support Vector Machine (SVM) yang sudah dilakukan proses pelatihan (*training*) dengan akurasi 81.73%. Data *training* yang digunakan dilakukan pelabelan manual sebanyak 1442 *tweet*, data *tweet* tersebut terdiri dari *tweet* negatif sebanyak 722 *tweet* dan 720 *tweet* positif. Sedangkan data *testing* yang digunakan sebanyak 350 data *tweet*. Dari proses analisis sentimen didapatkan hasil kelas sentimen per topik seperti pada Tabel 2.

Tabel 2. Hasil analisis sentiment per topik

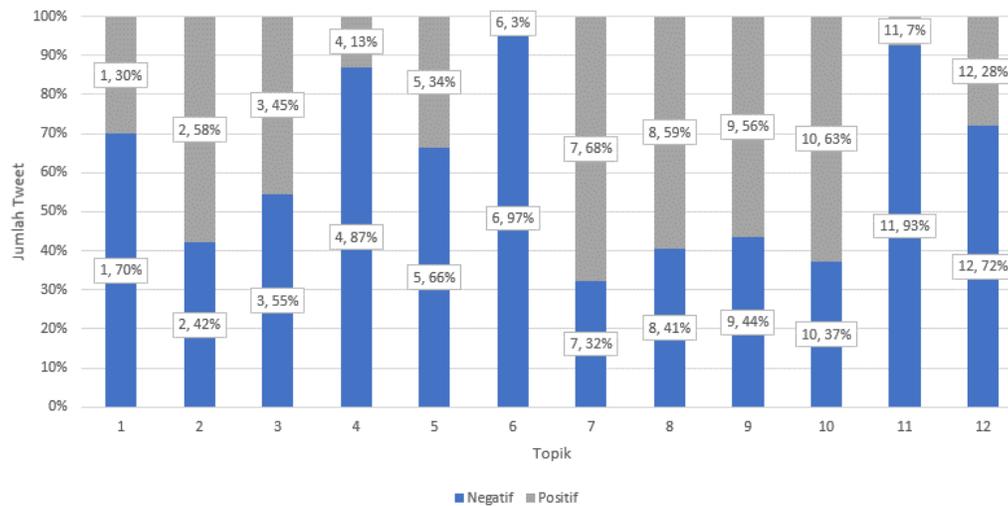
Topic	Kelas Sentimen	Jumlah Tweet
1	Negatif	299
	Positif	128
2	Negatif	99
	Positif	136
3	Negatif	426
	Positif	354
4	Negatif	340
	Positif	51

Tabel 2. Hasil analisis sentiment pertopik

Topic	Kelas Sentimen	Jumlah Tweet
5	Negatif	269
	Positif	137
6	Negatif	3493
	Positif	120
7	Negatif	68
	Positif	142
8	Negatif	97
	Positif	142
9	Negatif	184
	Positif	309
10	Negatif	356
	Positif	28
11	Negatif	177
	Positif	69

Gambar 7 merupakan grafik hasil analisis sentimen pertopik. Berdasarkan hasil analisis sentimen pertopik, dapat diketahui jumlah tweet kelas sentimen positif atau negatif pada tiap topik. Topik yang memiliki kelas sentimen positif lebih banyak dari pada kelas negative adalah topik 10, topik 9, topik 7, dan topik 8. Sedangkan topik yang memiliki kelas sentimen negatif lebih banyak dari pada kelas positif adalah topik 6, topik 3, topik 11, topik 4, topik 1 dan topik 5.

Topik 6 memiliki nilai sentimen negatif paling banyak dibandingkan topik yang lain. Topik tersebut mengenai perkembangan Vaksinasi Covid-19 di Indonesia. Hal tersebut dimungkinkan karena perkembangan proses vaksinasi di Indonesia belum begitu signifikan di awal tahun 2021. Sehingga banyak masyarakat yang merasa belum puas dengan proses vaksinasi yang sudah dilakukan oleh pemerintah. Sedangkan topik yang memiliki nilai sentiment positif paling banyak adalah topik 3 dan topik 10. Topik-topik tersebut membahas mengenai proses vaksinasi yang dilakukan secara prosedur kesehatan (prokes) dan kebijakan pemerintah terkait vaksinasi COVID-19. Hal tersebut mengindikasikan bahwa kepuasan masyarakat terkait standar pelayanan vasinasi sesuai prokes sangat tinggi. Serta kebijakan pemerintah terkait dengan penyelenggaraan vaksinasi Covid-19 sudah bagus [44].



Gambar 7. Hasil analisis sentiment pertopik

4. Kesimpulan

Dalam penelitian ini kami melakukan rancang bangun *Platform Social Media Intelligence* sebagai perangkat analisis persepsi masyarakat untuk peningkatan kualitas kebijakan publik yang berkelanjutan. Untuk melakukan analisis persepsi tersebut, penelitian ini melakukan dua tahapan analisis yaitu pemodelan topik (*topic modeling*) dan analisis sentimen (*sentiment analysis*). (1) Hasil pemodelan topik, didapatkan 11 topik pembicaraan terkait vaksinasi covid-19. Sedangkan hasil analisis sentiment dari 11 topik menghasilkan, Topik 6 memiliki nilai sentiment negative paling banyak yaitu mengenai perkembangan Vaksinasi Covid-19 di Indonesia. (2) Sedangkan topik yang memiliki nilai sentiment positif paling banyak adalah topik 3 dan topik 10. Topik-topik tersebut membahas mengenai proses vaksinasi yang dilakukan secara prosedur kesehatan (prokes) dan kebijakan pemerintah terkait vaksinasi COVID-19. Hasil analisis yang didapatkan diharapkan dapat dijadikan bahan evaluasi untuk peningkatan kualitas kebijakan publik. Untuk pengembangan penelitian selanjutnya, kami berencana menambahkan fitur *Social Network Analysis* (SNA) untuk pemetaan akun-akun yang secara aktif mengaspirasikan kebijakan-kebijakan pemerintah pada suatu jaringan sosial.

Daftar Pustaka

- [1] A. Ramdhani, ; Muhammad, and A. Ramdhani, “Konsep Umum Pelaksanaan Kebijakan Publik,” Feb. 2017.
- [2] Y. Lee and H. L. Schachter, “Exploring the Relationship between Trust in Government and Citizen Participation,” *Int. J. Public Adm.*, vol. 42, no. 5, pp. 405–416, Apr. 2019.
- [3] R. Fletcher and R. K. Nielsen, “Are people incidentally exposed to news on social media? A comparative analysis,” *New Media Soc.*, vol. 20, no. 7, pp. 2450–2468, Jul. 2018.
- [4] J. A. Hall, “When is social media use social interaction? Defining mediated social interaction,” *New Media Soc.*, vol. 20, no. 1, pp. 162–179, Jan. 2018.
- [5] A. Haro-de-Rosario, A. Sáez-Martín, and M. del Carmen Caba-Pérez, “Using social

- media to enhance citizen engagement with local government: Twitter or Facebook?," *New Media Soc.*, vol. 20, no. 1, pp. 29–49, Jan. 2018.
- [6] A. Androutsopoulou, Y. Charalabidis, and E. Loukis, "Policy Informatics in the Social Media Era: Analyzing Opinions for Policy Making," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2018, vol. 11021 LNCS, pp. 129–142.
- [7] A. Lovari and C. Valentini, "Public Sector Communication and Social Media," in *The Handbook of Public Sector Communication*, Wiley, 2020, pp. 315–328.
- [8] J. H. Kietzmann, K. Hermkens, I. P. McCarthy, and B. S. Silvestre, "Social media? Get serious! Understanding the functional building blocks of social media," *Bus. Horiz.*, vol. 54, no. 3, pp. 241–251, May 2011.
- [9] A. Perrin, "Social Media Usage: 2005-2015 | Pew Research Center," *Pew Research Center*, 2015. [Online]. Available: <https://www.pewresearch.org/internet/2015/10/08/social-networking-usage-2005-2015/>. [Accessed: 19-Apr-2021].
- [10] N. Newman, R. Fletcher, D. Levy, and R. K. Nielsen, *Reuters Institute Digital News Report 2016*. Reuters Institute for the Study of Journalism, 2016.
- [11] M. Anderson and J. Jiang, "Teens, Social Media & Technology 2018 | Pew Research Center," *Pew Research Center*, 2018. [Online]. Available: <https://www.pewresearch.org/internet/2018/05/31/teens-social-media-technology-2018/>. [Accessed: 19-Apr-2021].
- [12] S. Siddiqui and T. Singh, "Social Media its Impact with Positive and Negative Aspects," *Int. J. Comput. Appl. Technol. Res.*, vol. 5, no. 2, pp. 71–75, 2016.
- [13] N. S. Hawi and M. Samaha, "The Relations Among Social Media Addiction, Self-Esteem, and Life Satisfaction in University Students," *Soc. Sci. Comput. Rev.*, vol. 35, no. 5, pp. 576–586, Oct. 2017.
- [14] J.-Y. Jung, "Social Media, Global Communications, and the Arab Spring: Cross-Level and Cross-Media Story Flows," in *Mediated Identities and New Journalism in the Arab World*, Palgrave Macmillan UK, 2016, pp. 21–40.
- [15] S. C. McGregor, "Social media as public opinion: How journalists use social media to represent public opinion:," <https://doi.org/10.1177/1464884919845458>, vol. 20, no. 8, pp. 1070–1086, May 2019.
- [16] J. C. Bertot, P. T. Jaeger, and J. M. Grimes, "Promoting transparency and accountability through ICTs, social media, and collaborative e-government," *Transform. Gov. People, Process Policy*, vol. 6, no. 1, pp. 78–91, Mar. 2012.
- [17] R. Díaz-Díaz and D. Pérez-González, "Implementation of Social Media Concepts for e-Government: Case Study of a Social Media Tool for Value Co-Creation and Citizen Participation," *J. Organ. End User Comput.*, vol. 28, no. 3, 2016.
- [18] E. Falco and R. Kleinhans, "Beyond technology: Identifying local government challenges for using digital platforms for citizen engagement," *Int. J. Inf. Manage.*, vol. 40, pp. 17–

20, Jun. 2018.

- [19] M. M. Skoric, Q. Zhu, D. Goh, and N. Pang, "Social media and citizen engagement: A meta-analytic review," *New Media Soc.*, vol. 18, no. 9, pp. 1817–1839, Oct. 2016.
- [20] M. del M. Gálvez-Rodríguez, A. Haro-de-Rosario, and M. del C. Caba-Pérez, "A Comparative View of Citizen Engagement in Social Media of Local Governments from North American Countries," in *Handbook of Research on Citizen Engagement and Public Participation in the Era of New Media*, 2017, p. 18.
- [21] D. LETCHER, "ONLINE POLITICAL PARTICIPATION, COLLECTIVE ACTION EVENTS, AND MEANINGFUL CITIZEN ENGAGEMENT: SOCIAL MEDIA USE DURING MASS PROTESTS," *Geopolit. Hist. Int. Relations*, vol. 10, no. 2, pp. 70–75, 2018.
- [22] İ. Hatipoğlu, M. Z. Sobaci, and M. F. Korkmaz, "Determinants of the Citizen Engagement Level of Mayors on Twitter: The Case of Turkey," in *Digital Government and Achieving E-Public Participation: Emerging Research and Opportunities*, 2020, p. 16.
- [23] P. J. Brubaker and C. Wilson, "Let's give them something to talk about: Global brands' use of visual content to drive engagement and build relationships," *Public Relat. Rev.*, vol. 44, no. 3, pp. 342–352, Sep. 2018.
- [24] C. Kim and S. U. Yang, "Like, comment, and share on Facebook: How each behavior differs from the other," *Public Relat. Rev.*, vol. 43, no. 2, pp. 441–449, Jun. 2017.
- [25] M. Habibi and P. W. Cahyo, "A social network analysis: identifying influencers in the COVID-19 vaccination discussion on twitter," *Compiler*, vol. 10, no. 2, pp. 99–108, Nov. 2021.
- [26] S. Bhattacharya, P. Srinivasan, and P. Polgreen, "Social media engagement analysis of U.S. Federal health agencies on Facebook," *BMC Med. Inform. Decis. Mak.*, vol. 17, no. 1, Apr. 2017.
- [27] H. Hendriks, B. van den Putte, W. A. Gebhardt, and M. A. Moreno, "Social drinking on social media: Content analysis of the social aspects of alcohol-related posts on facebook and instagram," *J. Med. Internet Res.*, vol. 20, no. 6, p. e9355, Jun. 2018.
- [28] R. B. Hubert, E. Estevez, A. Maguitman, and T. Janowski, "Examining government-citizen interactions on twitter using visual and sentiment analysis," *ACM Int. Conf. Proceeding Ser.*, May 2018.
- [29] R. Sandoval-Almazan and D. Valle-Cruz, "Facebook impact and sentiment analysis on political campaigns," in *ACM International Conference Proceeding Series*, 2018, pp. 1–7.
- [30] S. Park, D. Chung, and H. W. Park, "Analytical framework for evaluating digital diplomacy using network analysis and topic modeling: Comparing South Korea and Japan," *Inf. Process. Manag.*, vol. 56, no. 4, pp. 1468–1483, Jul. 2019.
- [31] J. J. Schoenefeld and A. J. Jordan, "Towards harder soft governance? Monitoring climate policy in the EU," <https://doi.org/10.1080/1523908X.2020.1792861>, pp. 774–786, 2020.

- [32] A. H. Juma'h A and Y. Alnsour, "Using social media analytics: The effect of President Trump's tweets on companies' performance," *Account. Manag. Inf. Syst.*, vol. 17, no. 1, pp. 100–121, 2018.
- [33] P. Singh, Y. K. Dwivedi, K. S. Kahlon, R. S. Sawhney, A. A. Alalwan, and N. P. Rana, "Smart Monitoring and Controlling of Government Policies Using Social Media and Cloud Computing," *Inf. Syst. Front.*, vol. 22, no. 2, pp. 315–337, Apr. 2020.
- [34] M. S. Islam *et al.*, "COVID-19–Related Infodemic and Its Impact on Public Health: A Global Social Media Analysis," *Am. J. Trop. Med. Hyg.*, vol. 103, no. 4, pp. 1621–1629, Oct. 2020.
- [35] E. Chen, K. Lerman, and E. Ferrara, "Tracking social media discourse about the COVID-19 pandemic: Development of a public coronavirus Twitter data set," *JMIR Public Heal. Surveill.*, vol. 6, no. 2, Apr. 2020.
- [36] A. Bradley and R. J. E. James, "Web Scraping Using R;," <https://doi.org/10.1177/2515245919859535>, vol. 2, no. 3, pp. 264–270, Jul. 2019.
- [37] S. Rastogi, D. Bansal -Farhan, W. Kurniawan, W. Maharani, A. F. Hidayatullah, and M. R. Ma'arif, "Pre-processing Tasks in Indonesian Twitter Messages," *J. Phys. Conf. Ser.*, vol. 801, no. 1, p. 012072, Jan. 2017.
- [38] H. Jelodar *et al.*, "Latent Dirichlet allocation (LDA) and topic modeling: models, applications, a survey," *Multimed. Tools Appl.*, vol. 78, no. 11, pp. 15169–15211, Jun. 2019.
- [39] H. Latipa Sari *et al.*, "Integration K-Means Clustering Method and Elbow Method For Identification of The Best Customer Profile Cluster," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 336, no. 1, p. 012017, Apr. 2018.
- [40] K. Chen, Z. Zhang, J. Long, and H. Zhang, "Turning from TF-IDF to TF-IGM for term weighting in text classification," *Expert Syst. Appl.*, vol. 66, pp. 245–260, Dec. 2016.
- [41] R. Ferdiana, F. Jatmiko, D. D. Purwanti, A. S. T. Ayu, and W. F. Dicka, "Dataset Indonesia untuk Analisis Sentimen," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 8, no. 4, pp. 334–339, Nov. 2019.
- [42] S. Fatima and B. Srinivasu, "Text Document categorization using support vector machine," *Int. Res. J. Eng. Technol.*, vol. 4, no. 2, 2017.
- [43] K. Stevens, P. Kegelmeyer, D. Andrzejewski, and D. Buttler, "Exploring topic coherence over many models and many topics," in *Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP)*, 2012, no. July, pp. 952–961.
- [44] M. Habibi, A. Priadana, and M. R. Ma'arif, "Sentiment Analysis and Topic Modeling of Indonesian Public Conversation about COVID-19 Epidemics on Twitter," *IJID (International J. Informatics Dev.)*, vol. 10, no. 1, pp. 23–30, Jun. 2021.