

ANALISIS EFEK BUZZER PADA OPINI MASYARAKAT MENGUNAKAN METODE INDOBERTWEET DAN XGBOOST BERDASARKAN FITUR GRAF DAN TWIT

(STUDI KASUS: PEMILU PRESIDEN 2024)

Oleh:

Surya Abdillah

Davian Benito

Gaudhiwaa Hendrasto



SEMIFINAL
BIG DATA CHALLENGE
2024



Puspresnas
Pusat Prestasi Nasional

BPTI
Badan Pengembangan Talenta Indonesia

Telkom
University

Kampus
Merdeka
INDONESIA JAYA

ANALISIS EFEK BUZZER PADA OPINI MASYARAKAT MENGGUNAKAN METODE HYBRID INDOBERTWEET DAN XG-BOOST BERDASARKAN FITUR GRAF DAN TWIT (STUDI KASUS: PEMILU PRESIDEN 2024)

Abstrak

Media sosial X merupakan salah satu media sosial yang digunakan dalam melakukan kampanye Pemilu Presiden 2024. Keberadaan *buzzer* sebagai salah satu pendekatan yang banyak diamati dapat menimbulkan permasalahan, yakni tercampurnya informasi riil dengan informasi dari *buzzer*. Oleh sebab itu, penelitian ini bertujuan untuk dapat mendeteksi *buzzer* dan menganalisis efek *buzzer* pada opini masyarakat. Tahapan yang dilakukan dalam pendeteksian *buzzer* dapat terbagi menjadi beberapa bagian, yakni pembentukan data latih, persiapan data, pelatihan model, dan evaluasi. Data yang digunakan berasal dari tiga sumber, yakni fitur teks, non-teks, dan graf twit. Pada fitur teks model LLM dapat dimanfaatkan. Sedangkan, fitur non-teks dan graf dapat dimanfaatkan dengan model *ensemble learning* konvensional. Hasil akhir deteksi dapat ditentukan dengan menimbang hasil prediksi kedua model. Pada penelitian ini didapatkan model terbaik adalah IndoBERTweet dan XGBoost dengan skema penggabungan hasil prediksi dengan *soft vote* dengan nilai akurasi 0.9436. Selain itu, ditemukan bahwa penggunaan fitur graf dapat meningkatkan performa deteksi *buzzer*. Penerapan deteksi sentimen twit pada penelitian ini, dapat menunjukkan adanya pengaruh signifikan dari *buzzer* terhadap opini yang ada dimana terdapat perubahan penurunan nilai sentimen positif twit pada setiap paslon secara signifikan.

Kata kunci: Deteksi Buzzer; Topic Modeling; Prediksi Sentimen; Ensemble Learning; Pemilihan Umum Presiden 2024.

I. PENDAHULUAN

Pada masa kampanye Pemilu Presiden Indonesia 2024, penggunaan *platform* medsos (media sosial) banyak digunakan oleh masyarakat dalam bertukar opini dan pandangan. Salah satu medsos yang banyak digunakan adalah X. Jumlah pengguna medsos ini selalu mengalami kenaikan setiap tahunnya (Kemp, 2024). Ramainya pengguna X di Indonesia memberikan kesempatan besar bagi para pasangan calon terpilih Presiden dan Wakil Presiden (paslon) untuk dapat menjangkau lebih banyak masyarakat.

Salah satu pendekatan dapat diamati adalah penggunaan *buzzer*. *Buzzer* merupakan seseorang/sekelompok orang yang dibayar oleh pihak tertentu untuk meramaikan atau menyebarkan informasi yang telah ditentukan. Meski *buzzer* juga digunakan dalam konteks lain, *buzzer* dalam konteks politik memiliki konotasi yang negatif. Konotasi negatif ini merupakan imbas dari dampak negatif *buzzer* itu sendiri, yakni *misleading information*. *misleading information* merupakan penyebaran informasi palsu yang dapat menyesatkan masyarakat. Oleh sebab itu, akan lebih baik apabila masyarakat dapat mengetahui mana akun atau *tweet* *buzzer* dan tidak. Namun demikian, selama ini masyarakat masih mengalami kesulitan dan kesalahan dalam mengklasifikasi *buzzer*. Oleh sebab itu, diperlukan suatu pendekatan yang dapat mendeteksi *buzzer* pada medsos X. Pendeteksian ini juga dapat bermanfaat untuk mendapatkan analisis isu dan elektabilitas yang lebih tepat terkait opini publik yang sebenarnya (organik).

Permasalahan serupa telah diangkat pada penelitian (Juzar & Akbar, 2018) yang mana mencoba melakukan pendeteksian *buzzer* menggunakan fitur graf pada data X di level pengguna. Namun, terdapat perbedaan besar pada kasus penelitian ini dimana data X yang digunakan berada di level *tweet*. Pada data X di level pengguna, peneliti (Juzar & Akbar, 2018) menggunakan informasi seperti umur pengguna, sepuluh postingan terakhir, dan keterhubungan antar pengguna yang dapat dilihat berdasarkan *followers* dan *following* dari akun pengguna tersebut. Sedangkan, pada data level *tweet* memerlukan pendekatan baru untuk mencari informasi fitur dan keterhubungan antar *tweet* yang dapat berguna nantinya. Perbedaan ini yang mendasari rumusan permasalahan dari penelitian ini, yakni bagaimana proses pendeteksian *buzzer* pada level *tweet*. Selain itu, rumusan permasalahan lainnya adalah bagaimana pengaruh dari keberadaan *buzzer* terhadap opini masyarakat terkait Pemilu Presiden 2024. Tujuan dari penelitian ini adalah mendapatkan metode deteksi *buzzer* dengan performa yang tinggi untuk mendeteksi dan menghapus *buzzer* dari data X di level *tweet*. Selain itu, penelitian ini bertujuan juga untuk menemukan dan menganalisis pengaruh keberadaan *buzzer* terhadap opini masyarakat terkait Pemilu Presiden 2024.

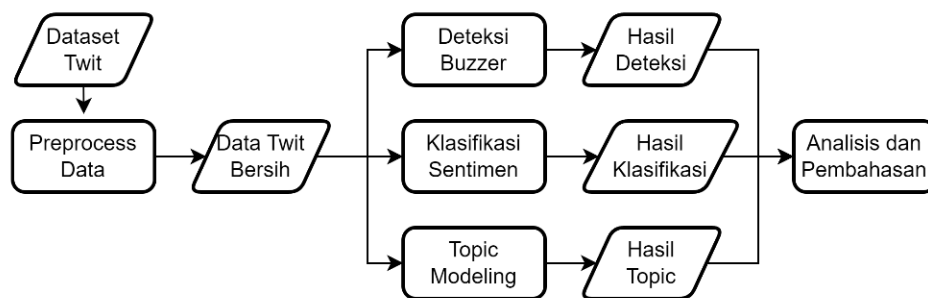
Dengan menyelesaikan permasalahan tersebut, penelitian ini diharapkan dapat bermanfaat bagi berbagai *stakeholder* yang terlibat. Bagi masyarakat umum, penelitian ini diharapkan dapat menyelesaikan permasalahan *overload information* dan *misleading information*, sehingga masyarakat bisa mendapatkan informasi yang riil dan berkualitas.

Bagi paslon dan tim sukses, penelitian ini diharapkan dapat memberikan informasi yang akurat terkait opini masyarakat yang sebenarnya pada topik dan isu yang muncul sehingga dapat menyusun strategi kampanye yang efektif. Selain itu, penelitian ini juga diharapkan dapat bermanfaat bagi lembaga survei untuk bisa mendapatkan hasil analisis elektabilitas paslon pada medsos X dengan lebih baik. Terakhir, semua hasil temuan dan metodologi yang digunakan pada penelitian ini diharapkan dapat bermanfaat pada lingkungan akademik dan peneliti lain sebagai landasan referensi untuk penelitian lainnya.

II. METODOLOGI

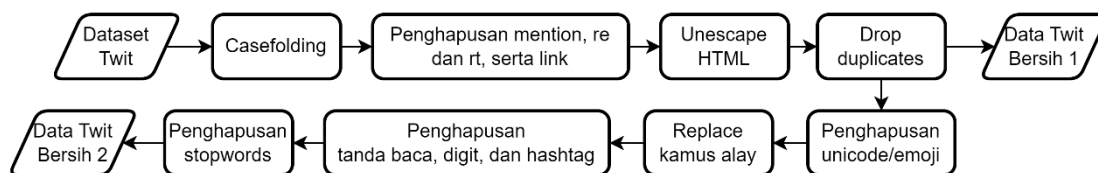
2.1. Rancangan Alur Metode

Penelitian ini akan mengikuti alur proses analisis yang menggabungkan metode dari berbagai *sub-task*, yakni deteksi *buzzer*, klasifikasi sentimen, dan *topic modeling*. Gambar 1 menampilkan diagram alir utama penelitian yang akan dilakukan.



Gambar 1 Diagram Alir Utama Penelitian

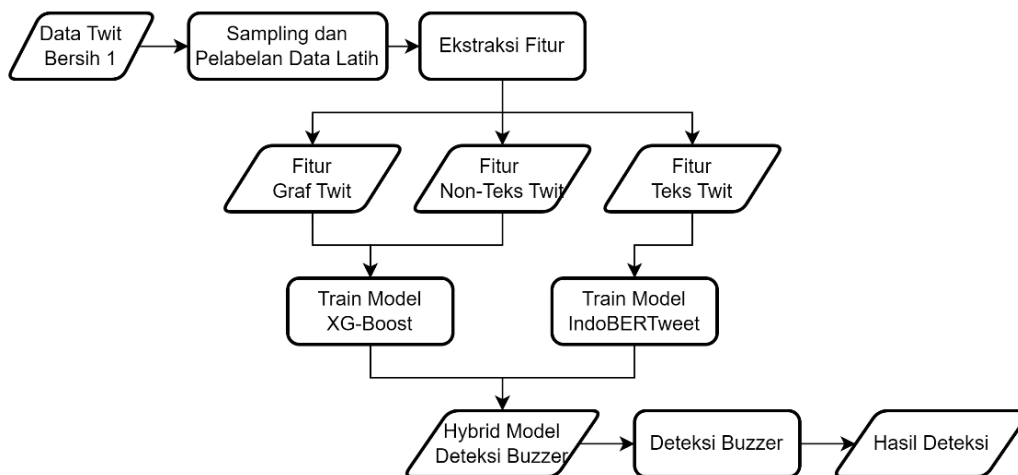
Penelitian ini juga menerapkan alur *preprocessing* data teks yang akan digunakan pada setiap tahapan alur penelitian. Gambar 2 menampilkan flowchart alur preprocess yang digunakan dalam penelitian ini. *Preprocessing* data teks bersifat sekuensial terhadap *sub-task* yang dilakukan. Pemilihan praproses data ini disesuaikan dengan kebutuhan data pada masing-masing *sub-task*. Penjelasan masing masing tahapan praproses data terdapat pada lampiran.



Gambar 2 Diagram Alir Preprocessing Teks Twit

2.2. Deteksi Buzzer

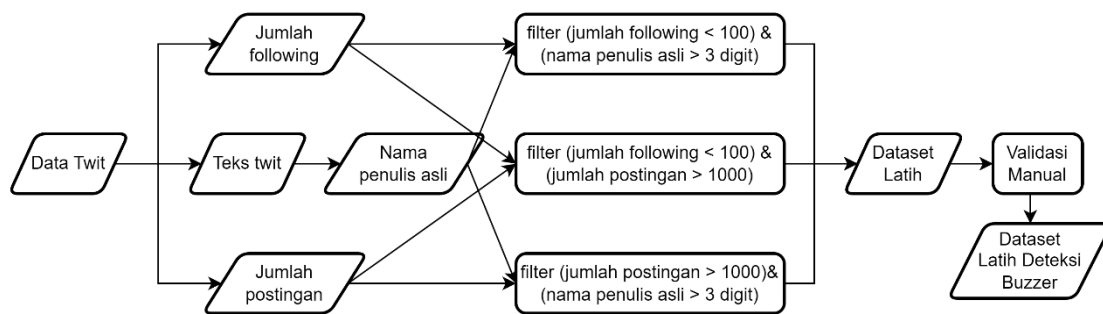
Tahap pertama pada penelitian ini adalah proses deteksi *buzzer* dalam data twit. Gambar 3 menampilkan diagram alir metode yang digunakan pada tahap ini.



Gambar 3 Diagram Alir Tahap Deteksi Buzzer

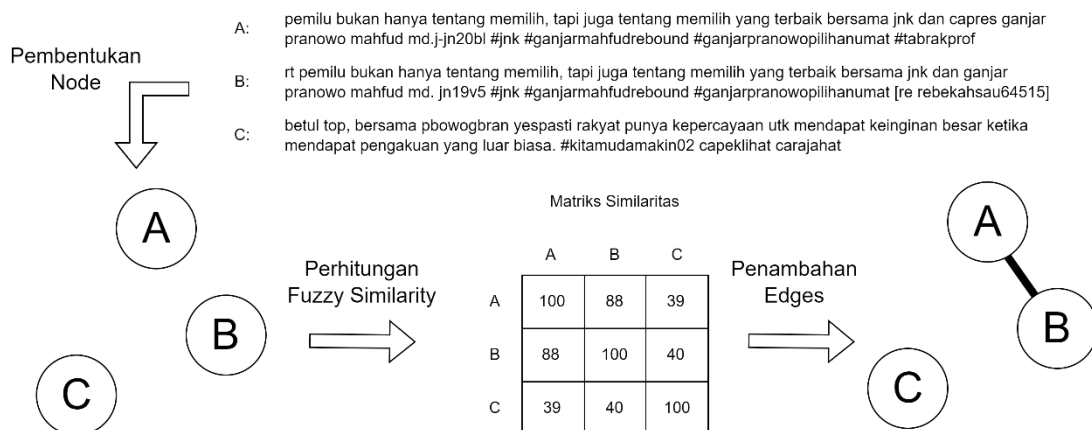
Deteksi *buzzer* merupakan jenis permasalahan *binary classification* yang bertujuan untuk mendeteksi data twit yang memiliki pola kemiripan sebagai *buzzer*. Berdasarkan penelitian (Ibrahim et al., 2016; Panatra et al., 2019; Suciati et al., 2019), terdapat beberapa pola dan ciri dari suatu twit yang memiliki probabilitas sebagai *buzzer*. Beberapa ciri tersebut merupakan twit berasal dari akun yang memiliki jumlah *follower* sedikit, tetapi memiliki jumlah aktivitas dan postingan yang sangat banyak. Selain itu, penamaan dari akun juga memengaruhi, seperti penggunaan digit yang berlebih pada username akun. Penelitian (Mustofa et al., 2023) mendapati bahwa terdapat kemiripan twit pada akun yang terdeteksi *buzzer*.

Berdasarkan hal tersebut, pada proses pembuatan *dataset* latih untuk model deteksi *buzzer*, digunakan pendekatan semi otomatis yang memanfaatkan beberapa ciri dari *buzzer* tersebut. Fitur yang akan digunakan dalam proses ini adalah jumlah postingan, jumlah *follower*, dan fitur tambahan, yakni *username* dari akun yang mengupload awal twit berjenis *retweet*. Fitur tambahan ini didapatkan dengan ekstraksi bagain *[RE ...]* pada twit. Gambar 4 menampilkan diagram alir dari pembuatan *dataset* latih dengan memanfaatkan kombinasi dari ketiga fitur tersebut. Setelah didapatkan sejumlah *dataset* latih, akan dilakukan pengecekan manual untuk memastikan kembali kualitas dari *dataset* yang telah didapatkan.



Gambar 4 Diagram Alir Pembuatan Dataset Latih Deteksi Buzzer

Pada tahap ini, terdapat tiga sumber fitur dataset yang akan digunakan. Hal ini dilakukan untuk dapat menangkap seluruh faktor dari data twit. Sumber fitur pertama berasal dari teks twit. Pada teks twit tersebut, dilakukan *preprocess* sebagaimana dijelaskan pada Gambar 2. Sumber fitur kedua berasal dari ekstraksi fitur non-teks twit. Sedangkan, sumber fitur ketiga berasal dari ekstraksi fitur graf. Fitur ini berupa beberapa nilai *centrality* yang dihasilkan dari graf yang terbentuk berdasarkan data twit. Graf yang berbentuk teks twit sebagai *node* dan *fuzzy similarity* antar twit sebagai *edges*. Dilakukan *pruning* terhadap *edges* yang terbentuk dengan memberikan nilai *threshold fuzzy similarity* di atas 80. Ilustrasi pembentukan graf dapat dilihat pada Gambar 5.



Gambar 5 Ilustrasi Proses Pembentukan Graf

Dari graf yang terbentuk akan dilakukan perhitungan nilai fitur *centrality* pada setiap *node*. Hasil fitur *centrality* yang didapatkan berguna untuk dapat menangkap keterhubungan antar data twit. Daftar fitur beserta deskripsi dapat dilihat pada Tabel 1.

Tabel 1 Daftar Fitur Beserta Sumber Fitur

Sumber Fitur	Fitur	Deskripsi
Teks	twit	Teks twit hasil <i>preprocessing</i>
Non-teks	num_retweets	Jumlah <i>retweet</i> twit
	frn_cnt	Jumlah <i>following</i> pengguna
	flw_cnt	Jumlah <i>follower</i> pengguna
	sts_cnt	Jumlah postingan pengguna
	lst_cnt	Jumlah daftar publik pengguna
	tcode	Tipe twit
	appearance_count	Jumlah kemunculan twit dalam sebulan
Graf	DC	<i>Degree centrality node</i>
	EC	<i>Eigenvector centrality node</i>
	PR	<i>Pagerank node</i>
	TA	Jumlah nilai fitur appearance_count dari keseluruhan node tetangga

Proses selanjutnya adalah tahap pelatihan model sesuai dengan sumber fitur yang dipakai. Pada tahap ini terdapat 2 model pendekatan yang digunakan yaitu, model *Large Language Model* (LLM) untuk sumber fitur teks twit dan model *ensemble learning* konvensional untuk sumber fitur non-teks twit dan fitur graf twit. Proses pelatihan ini akan menggunakan metrik akurasi terhadap validasi dataset untuk menentukan dan mengevaluasi hasil performa setiap percobaan.

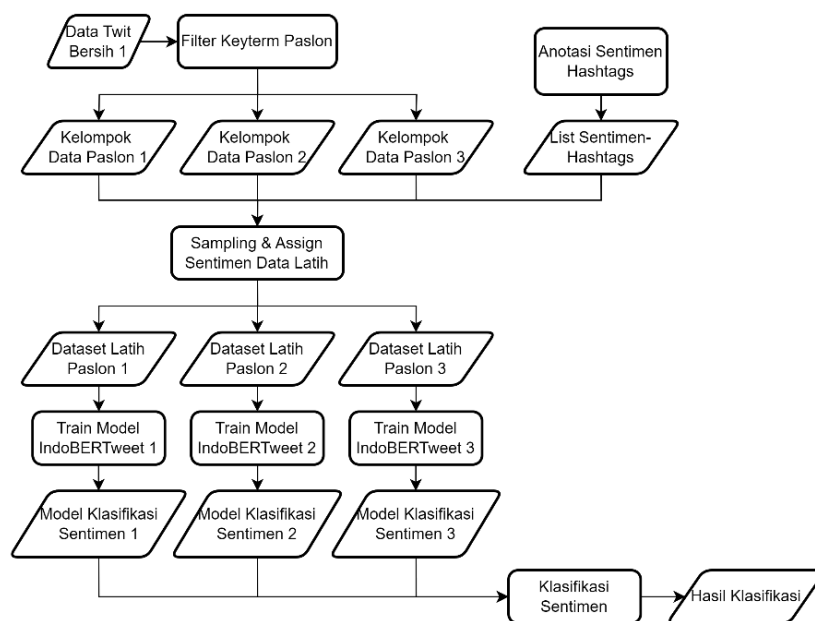
Pada pelatihan model LLM, akan dilakukan perbandingan antara penggunaan model IndoBERTweet dan IndoBERT untuk mendapatkan model dengan hasil performa yang terbaik. IndoBERTweet merupakan model *pretrained* dari IndoBERT (Wilie et al., 2020) yang telah di-*finetune* terhadap *dataset* twit Indonesia (Koto et al., 2021). Oleh sebab itu, diharapkan penggunaan model IndoBERTweet akan memiliki nilai performa yang lebih baik dibandingkan dengan IndoBERT. Selain pengujian model *pretrained*, juga akan dilakukan pengujian terhadap nilai *batch_size* yang digunakan pada tahap latih untuk mendapatkan skenario percobaan yang terbaik.

Sedangkan pada pelatihan model *ensemble learning* konvensional, akan dilakukan perbandingan antara penggunaan model XGBoost dan AdaBoost. Penggunaan model *ensemble learning* ini dipilih berdasarkan berbagai penelitian terkait, jenis model ini banyak digunakan dan terbukti memiliki nilai performa yang sangat baik terhadap fitur data tabular (Suciati et al., 2019). Penggunaan *hyperparameter tuning* juga dilakukan terhadap setiap model untuk mendapatkan hasil skenario percobaan yang terbaik.

Setelah didapatkan model LLM dan *ensemble learning* terlatih, dilakukan penggabungan prediksi dari kedua model tersebut atau dengan istilah lain *Hybrid Method Model Prediction*. Pada metode ini akan dilakukan pengujian sistem penggabungan prediksi untuk menentukan hasil prediksi model yang terbaik. Pengujian dilakukan antara penggunaan *unanimous voting* dan *soft voting*. *Unanimous voting* adalah penggabungan prediksi dimana kedua prediksi harus bernilai benar untuk prediksi akhir bernilai benar. Sedangkan, *soft voting* adalah penggabungan prediksi dimana salah satu atau kedua prediksi bernilai benar untuk prediksi akhir bernilai benar. Model hibrida ini lalu digunakan untuk memprediksi *buzzer* terhadap keseluruhan data.

2.3. Klasifikasi Sentimen

Tahap selanjutnya adalah klasifikasi sentimen tweet terhadap setiap paslon. Gambar 6 menampilkan diagram alir metode yang digunakan pada tahap klasifikasi sentimen ini.



Gambar 6 Diagram Alir Tahap Klasifikasi Sentimen

Pada tahap ini, data tweet yang telah melalui tahap *preprocess* akan diklasifikasikan ke dalam kelompok sentimen positif dan negatif. Sebelum itu, proses pengelompokan data tweet berdasarkan *keyterm* setiap paslon dilakukan untuk memisahkan *sub-task* klasifikasi sentimen ini untuk masing-masing paslon. Hal ini dilakukan untuk menangani situasi dimana sebuah tweet dapat mengandung *keyterms* beberapa paslon. Kondisi ini dapat menyebabkan kekeliruan prediksi sentimen terhadap paslon terkait. Dengan menggunakan pemisahan *sub-task*, model dapat lebih yakin menangkap sentimen yang

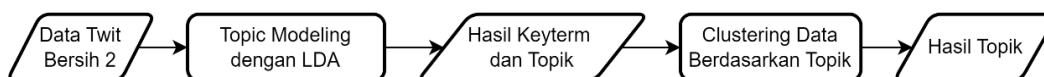
memang ditujukan secara spesifik terhadap paslon terkait. Daftar lengkap *keyterms* dapat dilihat pada lampiran 2.

Pembuatan *dataset* latih akan menggunakan sampel data setiap kelompok paslon dan memanfaatkan *hashtag* yang ada pada beberapa *tweet* untuk menentukan sentimen terkait *tweet* tersebut. *Hashtag* merupakan kutipan teks khusus pada sebuah *tweet* yang mencerminkan topik dan sentimen terkait keseluruhan *tweet* tersebut. Proses pengecekan dan pelabelan manual sentimen dilakukan terhadap berbagai *hashtag* yang sering muncul di *dataset* untuk setiap paslon. Lalu, dilakukan pengambilan data *tweet* yang mengandung *hashtag* terkait dan diberi label sentimen sesuai dengan *hashtag*-nya.

Setelah didapatkan *dataset* latih, dilakukan pelatihan model klasifikasi sentimen untuk masing-masing paslon menggunakan model IndoBERTweet. Setelah didapatkan ketiga model sentimen klasifikasi untuk setiap paslon, dilakukan prediksi klasifikasi sentimen terhadap data *tweet* terkait.

2.4. Topic Modeling

Tahap selanjutnya adalah identifikasi topik atau isu berdasarkan hasil *topic modeling* data *tweet*. Gambar 7 menampilkan diagram alir metode yang digunakan.



Gambar 7 Diagram Alir Topic Modeling

Metode yang digunakan pada tahap ini adalah *Latent Dirichlet Allocation* (LDA). LDA merupakan metode statistik yang digunakan untuk menemukan berbagai topik dari kumpulan dokumen. Penggunaan LDA berdasarkan (Yang et al., 2014) memiliki performa yang bagus dan efisien terhadap volume data teks yang besar sehingga cocok digunakan untuk permasalahan ini.

Dari penggunaan LDA, akan dihasilkan *keyterm* untuk setiap topik yang diidentifikasi. Hasil *keyterm* tersebut dapat mencerminkan topik atau isu yang ada pada kelompok data *tweet* terkait. Penggunaan metrik koherensi *c_v* dan *u_mass* dipakai untuk melihat kualitas dari topik yang dihasilkan. Pengujian terhadap jumlah topik yang terbentuk dilakukan untuk mencari hasil *topic modeling* yang terbaik.

III. HASIL DAN PEMBAHASAN

3.1. Hasil Deteksi Buzzer

Proses pembuatan dan anotasi data latih yang dilakukan pada tahap deteksi *buzzer* mendapatkan sejumlah 6000 data twit dengan distribusi label seimbang. Ekstraksi setiap sumber fitur dilakukan dan didapatkan fitur non-teks dan teks twit untuk setiap data twit dengan mengambil dan mengolah data twit yang tersedia. Sedangkan, pada sumber fitur graf twit, dilakukan pembuatan graf representasi setiap data twit. Dari data twit yang ada, didapatkan graf dengan jumlah *node* dan *edge* sebanyak 2161199 dan 1956593.

Selanjutnya, dilakukan proses training terhadap setiap sumber fitur. Tabel 2 menampilkan perbandingan hasil performa pengujian pelatihan model LLM terhadap sumber data teks twit. Didapatkan model IndoBERTweet dengan batch size 32 dan learning rate 1e-5 memiliki nilai performa terbaik dengan akurasi sebesar 0.8294.

Tabel 2 Perbandingan Hasil Performa Pelatihan Model LLM

		IndoBERT (batch size)			IndoBERTweet (batch size)		
lr	batch_size	16	24	32	16	24	32
	1e-5	0.8133	0.8150	0.8125	0.8200	0.8150	0.8294
	3e-5	0.8142	0.8133	0.8100	0.8158	0.8158	0.8133

Sedangkan Tabel 3 menampilkan perbandingan hasil performa pengujian pelatihan model terhadap sumber data non-teks twit dan graf twit. Diperoleh model XGBoost dengan *max depth* 2 dan penambahan fitur graf twit memiliki nilai performa terbaik dengan akurasi sebesar 0.9241.

Tabel 3 Perbandingan Hasil Performa Pelatihan Model Ensemble Learning

		AdaBoost (max_depth)			XGBoost (max_depth)		
Fitur	Max depth	2	3	4	2	3	4
	non-teks	0.9100	0.9125	0.9016	0.9200	0.9216	0.9141
	non-teks + graf	0.9150	0.9141	0.9133	0.9241	0.9233	0.9225

Model terbaik dari kedua jenis, ini yakni model LLM IndoBERTweet (*batch size* 32 dan *learning rate* 1e-5) dan model XGBoost (*max depth* 2 dan fitur non-teks + graf)

akan digunakan dalam model penggabungan. Penggabungan dilakukan dengan *unanimous voting* dan *soft voting*. Didapatkan bahwa *soft voting* mendapatkan nilai akurasi yang tertinggi, yakni sebesar 0.9436. Dengan membandingkan nilai akurasi terbaik pada model LLM, ensemble learning, dan penggabungan dua model, dapat terlihat bahwa penggabungan dua model mendapatkan nilai performa terbaik. Prediksi *buzzer* terhadap keseluruhan data twit mendapatkan hasil twit terdeteksi *buzzer* sebanyak 4054361 dari 9817355 twit. Contoh hasil deteksi dapat dilihat pada lampiran 5.

3.2. Hasil Klasifikasi Sentimen

Proses pembuatan *dataset* latih dilakukan dengan memanfaatkan hasil anotasi sentimen *hashtag* untuk setiap paslon. Daftar sentimen *hashtag* setiap paslon dapat dilihat pada lampiran 3. Dari keseluruhan data sentimen yang didapatkan, akan digunakan 3000 data untuk setiap paslon, baik sentimen positif maupun negatif. Data yang didapatkan ini akan digunakan sebagai data latih untuk model prediksi sentimen IndoBERTweet. Pada pelatihan ini digunakan nilai parameter *batch size* 32 dan *learning rate* 1e-5 yang merupakan parameter terbaik pada deteksi *buzzer* dengan pemakaian fungsi callbacks terhadap nilai akurasi validasi. Hasil pelatihan ini dapat dilihat pada Tabel 4.

Tabel 4 Perbandingan Hasil Performa Pelatihan Model Klasifikasi Sentiment

Model	Epoch Terbaik	Val Akurasi
IndoBERTweet (Paslon 1)	5	0.9508
IndoBERTweet (Paslon 2)	4	0.9767
IndoBERTweet (Paslon 3)	2	0.9985

Hasil klasifikasi sentimen untuk setiap kelompok data paslon akan didapatkan dengan menggunakan prediksi dari ketiga model tersebut. Tabel 5 menampilkan distribusi hasil prediksi untuk setiap model terhadap data twit terkait masing-masing paslon.

Tabel 5 Distribusi Hasil Prediksi Sentimen Keseluruhan Data Twit

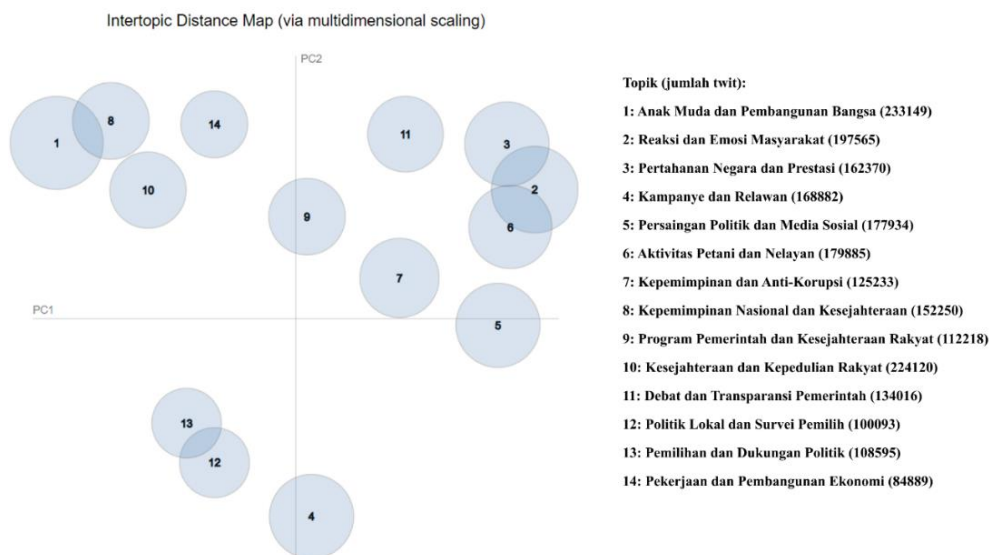
Data	Total	Positif	Negatif
Paslon 1	4686059	4671046	15013
Paslon 2	3269533	2381175	888358
Paslon 3	2874525	2779775	94750

3.3. Hasil Topic Modeling

Pada tahap *topic modeling*, digunakan algoritma LDA untuk mengidentifikasi topik yang terdapat pada keseluruhan data twit dan mengelompokkan twit berdasarkan topik yang sesuai. Penggunaan metrik evaluasi berupa nilai koherensi c_v dan u_mass juga dipakai untuk menentukan kualitas dari topik yang dihasilkan. Grafik pergerakan nilai koherensi c_v dan u_mass dapat dilihat pada lampiran 7.

Koherensi c_v yang semakin tinggi menandakan bahwa topik yang terbentuk semakin baik. Sedangkan, pada koherensi u_mass nilai negatif yang semakin dekat dengan nilai 0 merupakan nilai terbaik. Pada pengujian ini, didapatkan jumlah topik terbaik adalah 14 dengan nilai koherensi c_v dan u_mass sebesar 0.5343 dan -4.7826 secara berurutan.

Dari *keyterms* yang didapatkan (dapat dilihat pada lampiran 4) dapat ditarik kesatuan topik sebagaimana tersedia pada Gambar 10. Dapat terlihat bahwa jarak antar kluster topik cukup terlihat. Meski terdapat tumpang tindih pada kluster 1 dan 8; 2, 3, dan 6; serta 12 dan 13, penarikan kesatuan topik masih bisa dilakukan untuk masing-masing topik. Pada data twit yang digunakan, tiga topik atau isu yang paling ramai dibahas adalah anak muda dan pembangunan bangsa, reaksi dan emosi masyarakat, serta perhatian negara dan prestasi.

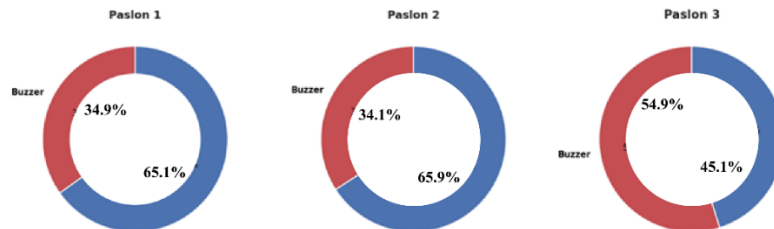


Gambar 8 Visualisasi Kluster Topik LDA

3.4. Analisis dan Pembahasan

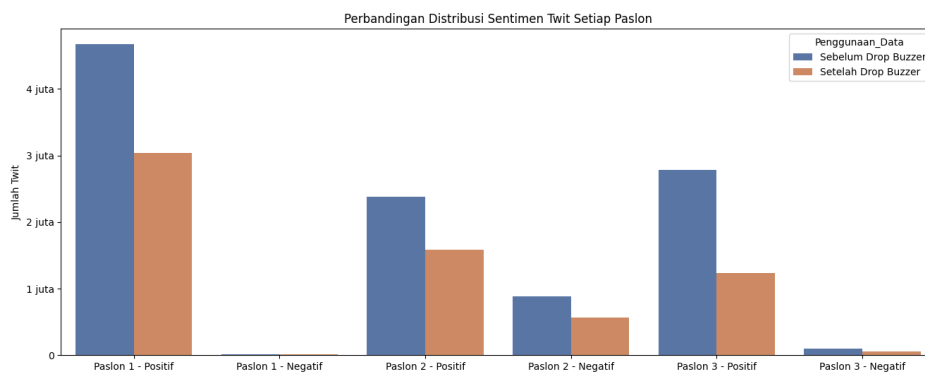
Berdasarkan hasil yang didapatkan, *insight* terkait perbandingan jumlah data *buzzer* terhadap setiap paslon dapat diketahui dan bisa dilihat pada Gambar 11. Berdasarkan hasil tersebut, dapat diamati bahwa persentase *buzzer* terhadap paslon 1 dan paslon 2 memiliki

nilai yang mirip yaitu di angka kisaran 35%. Namun hal menarik dapat diamati pada nilai persentase *buzzer* paslon 3 yang relatif tinggi yaitu di angka 54.9%. Hal ini mengindikasikan tren kuat penggunaan *buzzer* yang dilakukan terhadap paslon 3 pada platform medsos X.



Gambar 9 Distribusi Twit Buzzer pada Setiap Paslon 1, 2, dan 3 secara berurutan dari kiri ke kanan. Twit Buzzer (Merah) dan Non-Buzzer (Biru)

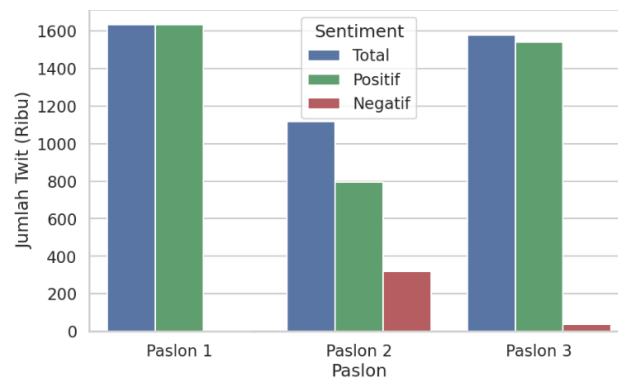
Untuk mendapatkan data twit organik, proses drop buzzer perlu dilakukan terhadap data twit. Hasil tahap analisis sentimen dapat digunakan untuk melihat efek penggunaan *buzzer* terhadap sentimen tiap paslon. Gambar 12 menampilkan distribusi sentimen twit terhadap setiap paslon sebelum dan sesudah proses *drop buzzer*. Berdasarkan hasil tersebut, ditemukan bahwa twit *buzzer* banyak digunakan untuk mempengaruhi sentimen positif setiap paslon.



Gambar 10 Perbandingan Distribusi Sentimen Twit Setiap Paslon Sebelum dan Sesudah Drop Buzzer

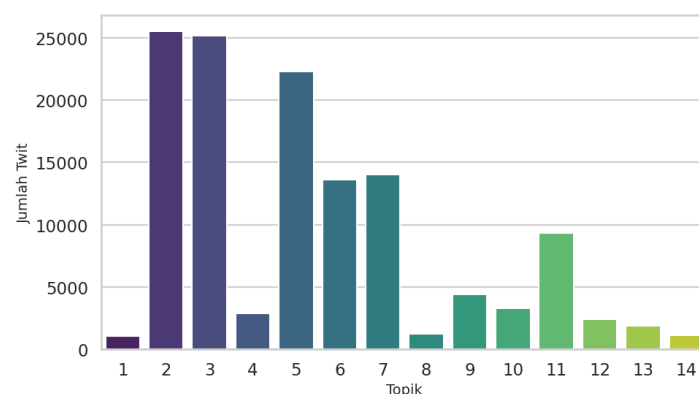
Pada Gambar 12 dapat juga diamati adanya penggunaan twit *buzzer* untuk memberikan sentimen negatif terhadap paslon. Gambar 13 menampilkan distribusi sentimen pada data twit *buzzer* setiap paslon. Berdasarkan hasil tersebut, dapat diamati bahwa penggunaan *buzzer* terhadap paslon 1 dan 3 banyak dipakai untuk meningkatkan sentimen positif paslon. Sedangkan terhadap paslon 2, ditemukan penggunaan *buzzer* cukup banyak dipakai untuk menyerang dan memberikan sentimen negatif terhadap

paslon. Temuan ini dapat didukung dengan mengingat gejala kampanye pada saat itu dimana terdapat banyak serangan dari berbagai pihak kepada paslon 2.



Gambar 11 Distribusi Sentimen Twit Buzzer Setiap Paslon

Dari nilai sentimen negatif pada data twit buzzer terhadap paslon 2, dapat dianalisis topik atau isu yang menjadi incaran *buzzer* bersentimen negatif tersebut dengan menggunakan hasil dari *topic modeling*. Gambar 14 menampilkan analisis topik dari twit *buzzer* negatif terhadap paslon 2. Berdasarkan hasil tersebut, didapatkan topik yang sering muncul adalah topik 2, 3, dan 5 yang merupakan ‘Reaksi dan Emosi Masyarakat’, ‘Pertahanan Negara dan Prestasi’, dan ‘Persaingan Politik dan Media Sosial’ secara berurutan. Mengingat latar belakang dari paslon 2, topik dan isu yang ditemukan pada penelitian ini mencerminkan latar belakang paslon dan mengindikasikan bahwa pihak lain condong menyerang terhadap topik tersebut.



Gambar 12 Distribusi Topik Twit Buzzer Paslon 2 dengan Sentimen Negatif

IV. PENUTUP

Dari tahapan penelitian yang telah dilakukan, didapatkan bahwa pendeteksian *buzzer* pada level twit dapat dilakukan dengan menggunakan fitur teks, non-teks, dan graf. Fitur teks dapat dimanfaatkan dengan mengembangkan model IndoBERTweet,

sedangkan fitur non-teks dan graf dapat dimanfaatkan dengan menggunakan model XGBoost. Hasil deteksi kedua model dapat digabungkan dengan metode *soft vote* untuk memperoleh hasil prediksi akhir yang terbaik dengan nilai akurasi 0.9436. Penelitian ini mendapatkan persentase twit *buzzer* dari keseluruhan twit sebesar 41.30%. Dengan penghapusan twit *buzzer* ini, bisa didapatkan nilai sentimen yang lebih representatif karena mendapatkan twit yang organik. Terdapat perubahan nilai sentimen pada setiap paslon, yakni berupa penurunan jumlah sentimen positif pada setiap paslon secara signifikan. Hal ini menunjukkan bahwa keberadaan *buzzer* terbukti memiliki pengaruh terhadap opini masyarakat yang didapatkan dari medsos X. Kombinasi deteksi *buzzer* dan metode LDA sebagai *topic modeling* bisa digunakan untuk mendapatkan fokus topik atau isu yang digunakan untuk melakukan serangan maupun menjadi kekuatan bagi masing-masing paslon.

Berdasarkan hasil penelitian, terdapat persentase twit *buzzer* yang sangat banyak pada medsos X, dalam hal ini seputar Pemilu Presiden 2024. Oleh sebab itu, masyarakat diharapkan untuk dapat lebih waspada dalam mencerna informasi yang ada pada data twit. Twit organik yang didapatkan dari penghapusan *buzzer* dapat digunakan oleh paslon dan tim sukses untuk menentukan strategi yang diambil. Aspek topik dan isu yang diangkat bisa menjadi *insight* tambahan untuk mendukung pembentukan strategi. Pendekatan yang digunakan dalam penelitian ini, terbukti dapat mendeteksi twit *buzzer* dengan akurasi yang sangat baik, sehingga bisa dipertimbangkan untuk diimplementasi oleh lembaga survei dalam menganalisis elektabilitas setiap paslon pada medsos X. Penelitian ini juga mendapatkan temuan baru, utamanya dalam *task* pendeteksian *buzzer* pada level twit yang masih sedikit diangkat pada penelitian terdahulu. Penggunaan fitur graf juga menjadi temuan baru yang terbukti dapat meningkatkan performa deteksi *buzzer*. Penelitian ini dapat dikembangkan dengan menambah fitur yang digunakan, utamanya dalam jenis fitur graf.

V. REFERENSI

- Aliyah Salsabila, N., Ardhito Winatmoko, Y., Akbar Septiandri, A., & Jamal, A. (2018). Colloquial Indonesian Lexicon. *Proceedings of the 2018 International Conference on Asian Language Processing, IALP 2018*, 226–229. <https://doi.org/10.1109/IALP.2018.8629151>
- Ibrahim, M., Abdillah, O., Wicaksono, A. F., & Adriani, M. (2016). Buzzer Detection

- and Sentiment Analysis for Predicting Presidential Election Results in a Twitter Nation. *Proceedings - 15th IEEE International Conference on Data Mining Workshop, ICDMW 2015*, 1348–1353. <https://doi.org/10.1109/ICDMW.2015.113>
- Juzar, M. T., & Akbar, S. (2018). Buzzer Detection on Twitter Using Modified Eigenvector Centrality. *Proceedings of 2018 5th International Conference on Data and Software Engineering, ICoDSE 2018*, 1–5. <https://doi.org/10.1109/ICODSE.2018.8705788>
- Kemp, S. (2024). *DIGITAL 2024: INDONESIA*. DataReportal. <https://datareportal.com/reports/digital-2024-indonesia>
- Koto, F., Lau, J. H., & Baldwin, T. (2021). IndoBERTweet: A Pretrained Language Model for Indonesian Twitter with Effective Domain-Specific Vocabulary Initialization. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (EMNLP 2021)*.
- Mustofa, A., Humaira, F. M., Ermawati, M., Natasari, P. S., Kurdianto, A. A., Prasetyo, A. A., & Faisal, A. L. F. (2023). Twitter Buzzer Detection System Using Tweet Similarity Feature and Support Vector Machine. *NJCA (Nusantara Journal of Computers and Its Applications)*, 8(1), 7. <https://doi.org/10.36564/njca.v8i1.306>
- Panatra, A. J., Chandra, F. B., Darmawan, W., Warnars, H. L. H. S., Utomo, W. H., & Matsuo, T. (2019). Buzzer Detection to Maintain Information Neutrality in 2019 Indonesia Presidential Election. *Proceedings - 2019 8th International Congress on Advanced Applied Informatics, IIAI-AAI 2019*, 873–876. <https://doi.org/10.1109/IIAI-AAI.2019.00177>
- Suciati, A., Wibisono, A., & Mursanto, P. (2019). Twitter Buzzer Detection for Indonesian Presidential Election. *ICICOS 2019 - 3rd International Conference on Informatics and Computational Sciences: Accelerating Informatics and Computational Research for Smarter Society in The Era of Industry 4.0, Proceedings*, 0–4. <https://doi.org/10.1109/ICICoS48119.2019.8982529>
- Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., & Purwarianti, A. (2020). IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding. *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural*

Language Processing.

Yang, S. H., Kolcz, A., Schlaikjer, A., & Gupta, P. (2014). Large-scale high-precision topic modeling on twitter. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1907–1916. <https://doi.org/10.1145/2623330.2623336>

LAMPIRAN

Lampiran 1 Contoh Hasil *Preprocessing* Teks Twit

<i>Preprocess</i>	<i>Hasil</i>
Twit asli	RT pemilu bukan hanya tentang memilih, tapi juga tentang memilih yang terbaik bersama jnk & Ganjar Pranowo Mahfud MD. Â jn19v5 #jnk #ganjarmahfudrebound #ganjarpranowopilihanumat @0FdbcMXvQjETOi9bqnSH4e8mcVM6salFYRhhtnQ0xQ= https://t.co/b2InuP5Dym [RE rebekahsau64515]
<i>Case folding</i>	rt pemilu bukan hanya tentang memilih, tapi juga tentang memilih yang terbaik bersama jnk & ganjar pranowo mahfud md. â jn19v5 #jnk #ganjarmahfudrebound #ganjarpranowopilihanumat @0fdbcmxvqjeto9bqnsh4e8mcm6salffyrhhtnq0xq= https://t.co/b2inup5dym [re rebekahsau64515]
Penghapusan <i>mention, re dan rt, serta link</i>	pemilu bukan hanya tentang memilih, tapi juga tentang memilih yang terbaik bersama jnk & ganjar pranowo mahfud md. â jn19v5 #jnk #ganjarmahfudrebound #ganjarpranowopilihanumat [RE rebekahsau64515]
<i>Unescape HTML</i>	pemilu bukan hanya tentang memilih, tapi juga tentang memilih yang terbaik bersama jnk & ganjar pranowo mahfud md. â jn19v5 #jnk #ganjarmahfudrebound #ganjarpranowopilihanumat [RE rebekahsau64515]
Penghapusan <i>unicode/emoji</i>	pemilu bukan hanya tentang memilih, tapi juga tentang memilih yang terbaik bersama jnk & ganjar pranowo mahfud md. jn19v5 #jnk #ganjarmahfudrebound #ganjarpranowopilihanumat
<i>Replace kamus alay</i>	pemilu bukan hanya tentang memilih, tapi juga tentang memilih yang terbaik bersama jnk & ganjar pranowo mahfud md. jn19v5 #jnk #ganjarmahfudrebound #ganjarpranowopilihanumat

<i>Preprocess</i>	<i>Hasil</i>
Penghapusan tanda baca, digit, dan hashtag	pemilu bukan hanya tentang memilih tapi juga tentang memilih yang terbaik bersama jnk ganjar pranowo mahfud md jnv
Penghapusan <i>stop words</i>	pemilu memilih memilih jnk jnv

Lampiran 2 Daftar *Keyterms* Masing-Masing Paslon

<i>Paslon</i>	<i>Keyterms</i>
1	anies, baswedan, muhaimin, iskandar, beswedan, nies, ahnies, abahswedan, amin, imin, chainmin, paslon 1, 01
2	prabowo, subianto, gibran, rakabuming, raka, prab, pragi, ps, pbowogbran, prbwo, wowo, bowo, pragib, gbran, pbowogbran, paslon 2, 02
3	ganjar, pranowo, mahfud, gajar, md, mahmud, mahdud, gm, paslon 3, 03

Lampiran 3 Daftar *Hashtag* Sentimen Paslon

<i>Paslon</i>	<i>Sentimen</i>	<i>Hashtag</i>
1	Positif	#AMINajaDulu, #AniesMuhaimin2024, #DesakAnies, #HaveAniesDay, #PahitManiesAlwaysWithAnies, #ANIES, #AniesBaswedan, #NakesDesakAnies, #AMINajaDulu, #SpotAbahAnies, #2024AniesPresiden, #Anies2ndStage, #Anies, #AniesDeservesBetter, #TakeAniesLookatJakarta, #HwaitingAbahAnies, #aniesbubble, #AMINkanIndonesia, #aniesbaswedan, #HUMANIES, #AniesPresiden2024, #anies, #BandungLautanAmin, #AMINIndonesiaLebihBaik, #AminMenang1Putaran, #amin, #AniesBestBoy, #UmpanManies, #cakimin, #IMINRocksDebate, #AMINSatukanIndonesia, #AbahAnies, #HumaniesSayangAbah, #WeAreHumanies, #AminPalingSiapUntukIndonesia, #AniesPresidenRI2024,

#AMINTolakIKN, #AMINajadullu, #CakImin,
#MulutManisAnies, #DesakSlepetAMIN,
#AniesIkonGenZMilenial, #14Feb2024AniesPresidenRI,
#EkspedisiAMIN, #HotlineParisAMIN, #SobatAnies, #AMIN,
#DesakAniesSamarinda, #SlepetImin, #UmatIslamPilihAMIN,
#AniesAtKadin, #AMIN2024MenangSATUPutaran,
#PKSmenangAMIN, #aminajadulu, #abahanies, #JejakAnies,
#DesakAniesAmbon, #AminMiskinkanKoruptor,
#KPUBerjamaahJegalAnies, #JanganGengsiDukungAMIN,
#BandungLautanAMIN, #MuhaiminIskandar,
#AMINPeduliGuru, #SorryYeAniesTerbaik, #AniesUnggul,
#IndustriKreatifAMIN, #AniesTakTerbendung,
#AniesMuhaimin, #AniesUnstoppable, #AMIN2024,
#AMINajadulu, #PahitManisAlwaysWithAnies,
#aniesmuhaimin2024, #AminAjaDulu, #NelayanPilihAMIN,
#AMINfor2024, #AniesStyle, #AniesFashion,
#muhaiminiskandar, #PerpustakaanAMIN,
#AMINBangunStadion, #Resolusi2024AniesPresiden,
#AMINMiskinkanKoruptor, #AMINPilihanRakyat,
#AniesDeserveBetter, #ProgramAirBersihAMIN, #MANIES,
#TimnasAMIN, #JalurKeretaAMIN, #AniesBubble, #Amin,
#AMINKanJawaTengah, #DukungPrabowoCoblosAnies,
#JanganGengsiDukungAmin, #HumaniesStayWithAnies,
#desakanies, #Humanies, #RumahMurahAMIN,
#IminRocksDebate, #JakLingkoAnies, #AMINAjaDuluâ,
#Manies, #KampanyeAkbarAMIN2024, #ProtectAnies,
#14Feb2024AniesPresidenRI, #SumselAminAja,
#4niesJanjiManis, #DanaPesantrenAMIN, #AminTolakIKN,
#JanjiManisAnies, #AminBerantasJudol, #AniesDeservesbetter,
#AMINPeduliPekerja, #AminkanIndonesia, #TEBARAMIN,
#CakImin1stWin, #AniesTerpelajar, #AniesMuaimin2024,
#2024AniesPresidenâ, #JaklingkoAnies, #AniesJanjiManis,

		<p>#aniesbubble, #fanartanies, #FokusAMIN, #JanjiAnies, #aminkanindonesia, #AniesTolakIKN, #AniesRocksDebatCapres, #AniesPengaruhPBB, #TeraniesAnies, #Aminajadulu, #AniesBuble, #ANIESPRESIDENKU, #Anies2024, #Anies2ndDebat, #UmatIslamPilihAmin, #CakIminBerkelas, #aniesimin, #TakeAniesLookAtJaka, #DesakAniesambon, #AMINMenang14Februari2024, #Aniesfanart, #HotLineParisAMIN, #AMINPilihanRakyatAkakSehat, #AniesMuhaimin2024, #DesakAniesDihambat, #AniesBaswedanProRakyat, #AniesMuhamin, #ApologizeToAnies, #aniesmuhaimin, #5JanjiAnies, #muhaimin, #AMINmenangpilpres2024, #panganmurahamin, #FaktaAnies, #pksmenanganiespresiden, #Muhaimin, #pksmenangamin, #BuruhDesakAnies, #gusimin, #AMINKanIndonedia, #AMINSajaDulu, #UMKMPilihAMIN, #AniesMenangDebat, #AMINmenangPKSmenang, #RelawanAMIN, #AMINMenangJudolHilang, #haveaniesday, #WELOVEYOUANIES, #DesakAniesJogja, #gimananies, #desakaniesgorontalo, #jawabnies, #imin, #APOLOGIZETOANIES, #humanies, #AnisBaswedan, #FPIBersamaAnies, #DesakAMIN, #FPIBersama4nies, #videotronaniesbubble, #DukungGibranCoblosImin, #AnieMuhaimin2024, #Anies2ndstage, #AminMenangJudolHilang, #AMINSATUkanIndonesia</p>
	Negatif	<p>#TerorisBomBaliDukung4nies, #CakIminDislepetKhofifah, #4niesOmonPalsu, #kocak, #0mon0monManis4nies, #TerorisPilih4nies, #4niesJ4njiM4nis, #AniesJanjiM4nis, #mulutmanisanies, #IndonesiaAMANTanpaAMIN</p>
2	Positif	<p>#PrabowoGemoy, #prabowo, #JabarTetapPakBowo, #BersamaPrabowoGibran, #Prabowo, #PrabowoGibranIstimewa, #prabowogibran,</p>

#PrabowoGibran2024, #gibran, #mendingprabowo,
#PrabowoGibran, #PrabowoSubianto, #WaktunyaPrabowo,
#MunculFenomenaPAKbowo, #prabowopresiden, #lipsus,
#FaktaBaruPrabowo, #GueEluKitaGIBRANER,
#PakbowoGakAsal2an, #PrabowoGibranInspiratif,
#gibranrakabuming, #Gibran, #prabowosubianto,
#prabowo2024, #HatiKitaMakinPAKBowo,
#DesakAniesDesakPrabowo, #prabowogibran2024,
#PrabowoGibranMakinAsik, #PrabowoGibranBanjirDukungan,
#TKNPrabowoGibran, #prabowosangperekat, #prabowosatria,
#prabowomenang2024, #prabowopresidenpetanisejahtera,
#petanidukungprabowo, #prabowodanpetani,
#UntungPakBOWObijak, #prabowopresiden2024,
#patriotprabowo, #PrabowoGibran2024Menang,
#gibranrakabumingraka, #PrabowoLawanKorupsi,
#prabowogibranmenangsatuputaran,
#PrabowoSelamatkanWilfrida, #GibranRakabuming,
#prabowomenang, #PrabowoPancasilais, #PrabowoPresiden,
#DukungPrabowoCoblosAnies, #allinprabowo,
#GibranRakabumingRaka, #PakBOWOTegasMumpuni,
#AllinPrabowoGibran, #prabowogibrankalsel,
#PrabowoMenangKoperasiJaya, #WaktunyaGibran,
#prabowoganjar, #MenangkanPrabowo,
#temanprabowosubiantoð, #PrabowoTerpercayaLawanKorupsi,
#prabowogibranmenang, #golkarprabowo, #Gibran2024,
#PrabowoPresiden2024, #AsliHanyaPrabowo,
#gibranmendengar, #MendingPrabowo, #prabowogemoy,
#LoginPrabowoGibran, #Prab0woMainB0la,
#tknprabowogibran, #PrabowoUntukSemua,
#PrabowoGibranUntukOjol,
#PrabowoGibranMenangSatuPutaran,
#PrabowoGibranSekaliPutaran, #gibran2024, #Patriotprabowo,

#Prabowo2024, #AsalBukanPSI, #TerusMajuBersamaPrabowo,
#GibranMendengar, #GibranPalingOK, #PBowoSatuPutaran,
#Prabowoâ, #YakinSamaGibran, #JatengAllinPrabowo,
#MasGibranInspirasiku, #GibranPenerusJokowi,
#DukungPrabowoGibran, #allinprabowogibran, #PrabowoAja,
#prabowogibran02, #prabowomenhan,
#pedagangsatukomandobersamaprabowo,
#prabowogibransulsel, #PRABOWOSELALUDIHATI,
#kampanyegibran, #DiplomasiProgresifPrabowo,
#GibranDukungDesa, #PRABOWOUNTUKRAKYAT,
#MenhanPrabowo, #AyoAklamasiPrabowoGibran,
#relawanprabowo, #GibranLanjutkanJokowi, #PSI,
#pakbowogakasal2an, #IkutJokowiPilihPSI,
#RelawanMasGibran, #MasGibranUntukIndonesia,
#NyataPAKBOWONaikTrus, #AllInPrabowo,
#PrabowoTheLivingLegend, #PrabowoAdalahIndonesia,
#PSIMenang, #AkuKamuPilihPrabowoGibran,
#AklamasiPrabowoGibran, #PrabowoGibranMenang,
#Prabowogibran, #temanprabowosubianto,
#PilihPrabowoGibran, #PrabowoGibranSatuPutaran,
#PrabowoGibranForChange, #prabowokalsel,
#PRABOWOBERSAMARAKYAT, #wayaheprabowogibran,
#GibranMenyapa, #Prabowogibrankalsel,
#DemokratBersamaPrabowo, #Saatnyaprabowo,
#PrabowoKasihBuktiBukanJanji, #GibranWakilPresidenRI24,
#PrabowoPresidenRI24, #prestasiPrabowo, #PrabowoWae,
#PRABOWOGIBRAN, #2024PrabowoGibran,
#JawaTimurPilihPrabowoGibran, #genziaininprabowo,
#gibrankalsel, #GibranLanjutkanJKW, #PrabowoGibran4Ojol,
#PrabowoGibranMerangkul, #PrabowoGibran4Change,
#AslinyaHanyaPrabowo, #prabowopedia, #gibranmasukgolkar

Negatif #AsalBukanPrabowo, #asalbukanprabowo, #terusmajubersamaprabowo, #GanjarSikatKorupsi, #DugaanKorupsiKemenhan, #AsalBukanPrabowoGibran, #GibranBocilSongong, #jikaprabowopresiden, #MenhanPrabowoGagal, #GerakanPerubahan, #asalbukanprabowogibran, #IndustriPertahananBelaPrabowo, #JanganPilihPrabowoGibran, #TamatkanPrabowoGibran, #asalBukanPrabowo, #MasGibranAnakHaramMK, #ASALBUKANWOWO, #AsalJanganPrabowo, #StopPrabowo, #JanganPilihPrabowo, #AsalkanBukanPrabowo, #TangkapPRABOWO, #Asalbukanprabowo, #AsalBukanGibran, #KontesNangisPrabowo, #asalBukanGibran, #PrabowoBukanJokowi, #JanganPilihPrabowoGibranð, #prabowoemosi, #GibranCawapresCacatHukum, #PrabowoGibranCurang, #PrabowoPelanggarHAM, #DukungGibranCoblosImin, #GibranNaikanPajak, #SoloBukanGibran, #GibranBukanKami, #PrabowoBabakBelur, #BocilGibran, #AyoPrabowoGamparAnies, #prabwojumatanidimana

3 Positif #GanjarMahfudRebound, #GanjarPranowoPilihanUmat, #GanjarMahfud2024, #DuluJokowiSekarangGanjar, #GanjarMahfudM3nangDebat, #MahfudLebihBaik3, #GanjarMahfudM3nang, #GanjarPresidenRakyat, #GanjarSandiSeliweRUN, #JNKBersamaGanjarMahfud, #3MandatGanjarMahfud, #GanjarPranowoPranowoPilihanUmat, #SahabatGanjar, #GanjarPresiden, #GanjarPranowo, #GanjarUntukSemua, #TemanAlamGanjar, #Ganjar, #JNKBarengGanjar, #GPMMD3, #MENangkanGAnjar, #JNKGanjarRI1, #GanjarMahfud, #2024JokowiDigantiGanjar, #JNKGanjar2024, #MahfudMD, #ganjarpranowo,

#SlankGanjarMahfud, #GanjarSikatKorupsi, #DialogGanjar,
#GanjarPilihanRakyat, #Ganjarpranowo, #MahfudGanjar2024,
#PilihGanjarPranowoMahfudMD, #GanjarSatSet, #Mahfud,
#GanjarMahfudPilihanRakyat, #MahfudTasTes, #mahfudmd,
#RembuGanjar, #GanjarHobiBlusukan, #TPNGanjarMahfud,
#GanjarMahfud2024 , #DuluJokowiSekarangGanjar ,
#GanjarUpToDat3, #GasssGanjarMahfud,
#GanjarTanpaDrama, #DeepTalkGanjar, #ganjarmahfud2024,
#TabrakProfMahfud, #ganjar, #ganjarlebihbaik,
#GanjarMahfudSatset, #GanjarL3bihBaik,
#GanjarMahfudSatSet, #ganjarmahfudm3nang,
#GanjarPranowo2024, #GanjarForPresident,
#GanjarGratisKuliah, #GPMMD, #GanjarPilihanku,
#M3nangkanGanjarMahfud, #GanjarMahfudPilihanRealistis,
#03GPMMDGASPOL, #sipalinganjar, #JnkBarengGanjar,
#BantuAlamGanjar, #GPPMD, #ganjarmahfud,
#MahfudPembelaRakyat, #GanjarPresidenku,
#MahfudLebihBaik, #GerakCepatGanjarMahfud,
#ganjaruntuksemua, #ganjarMahfudM3nangdebat,
#ganjarpresiden2024, #GanjarPelayanRakyat,
#CapresGanjarPranowoBersamaRakyat, #mafudmd,
#GanjarTaatEtika, #GanjarBikinSenang,
#GanjarMahfudCanvasingDay3, #SiapKetuaGanjar,
#HitamPutihGanjar, #GanjarMudaBeliRumah,
#GanjarBikinMimpiNyata, #prabowoganjar, #GanjarBeraksi,
#MenangkanGanjar, #GanjarBukanJokowi,
#SantriSuksesBarengGanjar, #GanjarPeduli, #ganjarpresidenku,
#JNKGanjarRI1GanjarRI1, #3GPMMDUNGGUL,
#GanjarTeladan, #ganjarsikatkorupsi, #GanjarPeduliRakyat,
#santridukungganjar, #GanjarBikinBahagia,
#HariRayaGanjarianSpartan, #mahfudmdÂ,
#UnggulBersamaGanjar, #SalamMetalGanjarMahfud,

#GanjarBantuImpian, #GanjarMemajukanNKRI, #GPMD3,
#MahfudGanjarPaduSuara, #GanjarBikinMudah,
#GanjarMewujudkanImpian, #L3bihBaikGanjarMahfud,
#GanjarInspirasi, #GanjarMahfudTaatiKonstitusi,
#GanjarBukanJKW, #GanjarPilihanKita, #MahfudMd,
#Ganjar2024, #GanjarPahlawanRakyat,
#ProgramUnggulanGanjar, #GanjarDengarAspirasi,
#GanjarCapres2024, #bantualamganjarjadianakpresiden2024,
#GanjarMembangunEkonomi, #GanjarPeduliPekerjaMigran,
#GanjarInklusif, #GanjarMahfudMenang,
#PilihanCerdasGanjar, #GanjarUntukBangsaNKRI,
#tetappilihganjar, #GanjarBikinImpianNyata,
#GanjarMahfud24, #GanjarSebagaiTeladan,
#PilihGanjar2024NKRIMaju, #ganjarrebound,
#3GPMDTASTES, #GPMMD2024, #GanjarPembela,
#GanjarPresidenKU, #GanjarBikinCerita,
#AtikohGanjarPranowo, #GanjarkeunJabar,
#GanjarPemimpinTerpercaya, #GanjarPekerjaMigran,
#SantriBerkualitasBarengGanjar, #GanjarSolusiTerbaikNKRI,
#GanjarMahfudTaatiHukum, #ganjarpresidenrakyat,
#GanjarMembangunMasaDepan, #GanjarMerawatPancasila,
#GanjarBikinGampang, #MAHFUDMDTASTES,
#SejahteraBersamaGanjar, #CumaGanjarYangTerbaik,
#PilihGanjarMajukanRI, #HanyaGanjarMahfud,
#PilihGanjar2024, #GanjarMahfudMenang2024,
#GanjarSolusiRakyatNKRI, #sahabatganjar,
#ganjarmahfudm3nangdebat, #GanjarIntegritas,
#GanjarUpayakanNKRIMaju,
#ArekSoroboyoPilihGanjarPranowo, #mahfud,
#GanjarPerlindungan, #JNKGanjarUntukNegeri,
#GanjarEnergiPositif, #AyoPilihGanjar2024,
#GanjarLebihBaik, #GanjarRumahImpian,

#GanjarMasaDepanKita, #MahfudMDMenangDebat,
 #GanjarDoaRakyat, #TepatPilihGanjar2024, #GanjarBukti,
 #GanjarBersamaBisa, #RumahGampangBarengGanjar,
 #GanjarMahfudM3nangTotal, #BeliRumahBarengGanjar,
 #GanjarSatuPilihan, #PilihGanjarDemiKemajuanRI,
 #RumahCepatBersamaGanjar, #GanjarPagiBaru,
 #GanjarBergerakBersama, #GanjarMahfud2024Â,
 #GanjarSinarkanHarapan, #GanjarSiapPimpinNKRI,
 #GanjarArsitekPerubahan, #GanjarMudaInovatif,
 #GasPollGanjarMahfud, #GanjarPuzzleSukses,
 #usaforganjarmahfud, #PetaniMajuBersamaGanjar,
 #ganjarin_aja, #GanjarForRakyat, #SuaraGanjarSukses,
 #GanjarianSpartan

Negatif #GanjarSokJadiJKW, #ganjarblunder, #GanjarBlunder,
 #ganjargagal, #GanjarGagal, #UbanSokJadiJWK,
 #GanjarBukanJWK, #UbanBukanPakde, #GanjarGakJelas,
 #ganjargakjelas, #gamausamagamud, #bikingamud,
 #ganjarbaper, #GanjarBaper, #ganjarantikritik,
 #GanjarJKWPalsu

Lampiran 4 *Keyterms* Hasil Topic Modeling

Topik	<i>Keyterms</i>
1	'0.044*"anak" + 0.027*"muda" + 0.022*"positif" + 0.021*"visi" + ' '0.017*"memiliki" + 0.017*"kesejahteraan" + 0.016*"menciptakan" + ' '0.014*"kebijakan" + 0.014*"pendidikan" + 0.014*"bangsa" + 0.013*"langkah" + ' '0.012*"kemajuan" + 0.011*"ekonomi" + 0.011*"membangun" + 0.011*"membawa" + ' '0.010*"generasi" + 0.010*"program" + 0.009*"misi" + 0.009*"pembangunan" + ' '0.009*"mendukung"
2	'0.019*"lihat" + 0.013*"suka" + 0.012*"sehat" + 0.010*"video" + ' '0.009*"langsung" + 0.008*"bilang" + 0.008*"gama" + 0.007*"biar" + ' '0.007*"masuk" + 0.006*"uang" + 0.006*"fitnah" + 0.006*"tua" +

Topik	Keyterms
1	'0.044*"anak" + 0.027*"muda" + 0.022*"positif" + 0.021*"visi" + ' '0.017*"memiliki" + 0.017*"kesejahteraan" + 0.016*"menciptakan" + ' '0.014*"kebijakan" + 0.014*"pendidikan" + 0.014*"bangsa" + 0.013*"langkah" + ' '0.012*"kemajuan" + 0.011*"ekonomi" + 0.011*"membangun" + 0.011*"membawa" + ' '0.010*"generasi" + 0.010*"program" + 0.009*"misi" + 0.009*"pembangunan" + ' '0.009*"mendukung"
	0.006*"ku" + ' '0.006*"bikin" + 0.005*"pakai" + 0.005*"menangis" + 0.005*"cari" + ' '0.005*"takut" + 0.005*"coba" + 0.005*"berani"
3	'0.024*"salah" + 0.019*"menhan" + 0.018*"data" + 0.015*"omon" + ' '0.014*"negara" + 0.013*"nilai" + 0.012*"pertahanan" + 0.012*"bilang" + ' '0.009*"tanah" + 0.009*"rahasia" + 0.009*"goblok" + 0.009*"jejak" + ' '0.007*"mengomong" + 0.007*"rekam" + 0.007*"tni" + 0.007*"kinerja" + ' '0.006*"anggaran" + 0.006*"prestasi" + 0.006*"etika" + 0.006*"ham"
4	'0.021*"allah" + 0.021*"warga" + 0.018*"kampanye" + 0.014*"relawan" + ' '0.013*"selamat" + 0.009*"acara" + 0.009*"sehat" + 0.008*"slank" + ' '0.008*"tim" + 0.008*"ketua" + 0.007*"papua" + 0.007*"salam" + 0.006*"jawa" ' ' + 0.006*"santri" + 0.006*"ulang" + 0.006*"hadir" + 0.006*"kota" + ' '0.005*"pagi" + 0.005*"doa" + 0.005*"atikoh"
5	'0.041*"putaran" + 0.022*"kalah" + 0.019*"partai" + 0.013*"pdip" + ' '0.011*"kali" + 0.011*"tiktok" + 0.009*"pilpres" + 0.006*"kubu" + ' '0.006*"pks" + 0.006*"gerindra" + 0.006*"live" + 0.006*"suara" + ' '0.005*"demokrat" + 0.005*"dki" + 0.005*"ahok" + 0.005*"nyapres" + 0.005*"p" ' ' + 0.005*"kau" + 0.005*"koalisi" + 0.005*"x"
6	'0.021*"bikin" + 0.015*"pakai" + 0.014*"petani" + 0.012*"bawa" + ' '0.011*"nelayan" + 0.011*"kampanye" + 0.010*"lihat" + 0.008*"bangga" + ' '0.006*"lucu" + 0.006*"benaran" + 0.006*"senang" + 0.006*"hutang" + ' '0.006*"foto" + 0.006*"biar" + 0.005*"videotron" + 0.005*"akun" + ' '0.005*"pintar" + 0.005*"buzzer" + 0.005*"kali" + 0.005*"tom"
7	'0.080*"jokowi" + 0.023*"korupsi" + 0.014*"hukum" + 0.012*"demokrasi" + ' '0.010*"negara" + 0.009*"cita" + 0.007*"menteri" +

Topik	Keyterms
1	'0.044*"anak" + 0.027*"muda" + 0.022*"positif" + 0.021*"visi" + ' '0.017*"memiliki" + 0.017*"kesejahteraan" + 0.016*"menciptakan" + ' '0.014*"kebijakan" + 0.014*"pendidikan" + 0.014*"bangsa" + 0.013*"langkah" + ' '0.012*"kemajuan" + 0.011*"ekonomi" + 0.011*"membangun" + 0.011*"membawa" + ' '0.010*"generasi" + 0.010*"program" + 0.009*"misi" + 0.009*"pembangunan" + ' '0.009*"mendukung"
	0.007*"pemerintahan" + ' '0.007*"mundur" + 0.006*"uu" + 0.006*"kekuasaan" + 0.006*"kpk" + ' '0.006*"bawaslu" + 0.006*"konstitusi" + 0.005*"koruptor" + 0.005*"bebas" + ' '0.005*"anti" + 0.005*"aturan" + 0.004*"pejabat" + 0.004*"pemilu"
8	'0.044*"negara" + 0.039*"maju" + 0.029*"membawa" + 0.024*"memimpin" + ' '0.020*"sejahtera" + 0.019*"negeri" + 0.017*"setuju" + 0.016*"kepemimpinan" ' ' + 0.012*"pertahanan" + 0.012*"dunia" + 0.012*"adil" + 0.011*"keamanan" + ' '0.010*"bangsa" + 0.009*"pokoknya" + 0.009*"sepakat" + 0.009*"menghadapi" + ' '0.008*"cerdas" + 0.008*"percaya" + 0.008*"menjaga" + 0.008*"kuat"
9	'0.055*"program" + 0.016*"makan" + 0.013*"rumah" + 0.012*"jalan" + ' '0.012*"jakarta" + 0.011*"gratis" + 0.009*"gubernur" + 0.009*"air" + ' '0.008*"joget" + 0.008*"siang" + 0.007*"sorry" + 0.007*"bantuan" + ' '0.007*"salah" + 0.007*"gagal" + 0.007*"sekolah" + 0.006*"miskin" + ' '0.006*"harga" + 0.006*"jateng" + 0.006*"dki" + 0.006*"food"
10	'0.054*"kerja" + 0.032*"nyata" + 0.024*"janji" + 0.017*"guru" + ' '0.017*"bukti" + 0.017*"optimis" + 0.015*"sukses" + 0.015*"bangsa" + ' '0.015*"duo" + 0.015*"internet" + 0.015*"ktp" + 0.014*"program" + ' '0.012*"umkm" + 0.011*"angin" + 0.011*"sakti" + 0.010*"juta" + ' '0.010*"tindakan" + 0.010*"bubble" + 0.010*"lapangan" + 0.009*"wan"
11	'0.104*"debat" + 0.037*"bersih" + 0.030*"kuatkankpk" + 0.020*"menyerang" + ' '0.018*"tokoh" + 0.012*"gagasan" + 0.011*"personal" + 0.010*"malam" + ' '0.009*"gp" + 0.008*"tenang" + 0.008*"serangan" + 0.008*"belajar" + ' '0.007*"ide" + 0.007*"serang" +

Topik	Keyterms
1	'0.044*"anak" + 0.027*"muda" + 0.022*"positif" + 0.021*"visi" + ' '0.017*"memiliki" + 0.017*"kesejahteraan" + 0.016*"menciptakan" + ' '0.014*"kebijakan" + 0.014*"pendidikan" + 0.014*"bangsa" + 0.013*"langkah" + ' '0.012*"kemajuan" + 0.011*"ekonomi" + 0.011*"membangun" + 0.011*"membawa" + ' '0.010*"generasi" + 0.010*"program" + 0.009*"misi" + 0.009*"pembangunan" + ' '0.009*"mendukung"
	0.007*"emosi" + 0.006*"marah" + ' '0.006*"menonton" + 0.006*"setuju" + 0.005*"habis" + 0.005*"emak"
12	'0.072*"politik" + 0.030*"jakarta" + 0.027*"jawa" + 0.023*"survei" + ' '0.019*"warga" + 0.016*"jabar" + 0.014*"hasil" + 0.013*"srimulat" + ' '0.013*"main" + 0.013*"survey" + 0.012*"barat" + 0.011*"pemilih" + ' '0.011*"buruh" + 0.010*"kemenangan" + 0.010*"persen" + 0.010*"timur" + ' '0.009*"suara" + 0.008*"kelas" + 0.007*"jatim" + 0.007*"keceriaan"
13	'0.061*"pilpres" + 0.055*"memilih" + 0.052*"pasangan" + 0.046*"cawapres" + ' '0.043*"calon" + 0.038*"pilihan" + 0.027*"mendukung" + 0.025*"suara" + ' '0.021*"jnk" + 0.021*"pemilu" + 0.015*"teman" + 0.013*"wakil" + 0.012*"ri" + ' '0.011*"memenangkan" + 0.011*"positif" + 0.010*"yes" + 0.010*"bangsa" + ' '0.009*"unggul" + 0.009*"pemilihan" + 0.008*"ajak"
14	'0.054*"kerja" + 0.032*"nyata" + 0.024*"janji" + 0.017*"guru" + ' '0.017*"bukti" + 0.017*"optimis" + 0.015*"sukses" + 0.015*"bangsa" + ' '0.015*"duo" + 0.015*"internet" + 0.015*"ktp" + 0.014*"program" + ' '0.012*"umkm" + 0.011*"angin" + 0.011*"sakti" + 0.010*"juta" + ' '0.010*"tindakan" + 0.010*"bubble" + 0.010*"lapangan" + 0.009*"wan"

Lampiran 5 Contoh Hasil Deteksi Buzzer

nama_ re	sts_ cnt	flw_ cnt	...	content	prediksi
-------------	-------------	-------------	-----	---------	----------

Lampiran 7 Pergerakan Nilai Koherensi c_v dan u_mass Terhadap Jumlah Topik

