























X: Modern Days Political Battlefield



Penelitian PVRI: Pemilu 2024 Banyak Kecurangan dan Berpotensi Picu Konflik



@LagiBeTeAii

Soal videotron dari Kpopres diturunkan, Rocky Gerung: "Masa menghalangi, itukan norak namanya"

Tp kalo gak salah mereka jg punya Videotron bang tp gak menarik kek baliho mereka yg digambar gemoy² tp aslinya 🔎 Makanya Pak Lurah Panik

Abah Anies deserve better #AMINAjaDulu 🔞 🌑

SabdoPalon @Notorosoati · Jan 29

Perempuan menjadi sosok penting dalam kehidupan sehari-hari.

Maka dari itu, Ganjar-Mahfud mengeluarkan program "Perempuan Harus Berani Maju dan Berdaya"

#SahabatGanjar #GanjarUntukSemua #GanjarMahfud2024 #PastiLebihBaik #SatSetLebihBaik



Dyah Adila 3 @dyhadila · Jan 7

Kalo prabowo kalah saya bikin open live lecture ttg Transformers architecture, skalian di coding in di notebook pake zoom 😊

> https://nasional.kompas.com/read/2024/03/20/09250191/penelitian-pvri-pemilu-2024-banyak-kecurangan-dan-berpotensi-picu-konflik https://www.antaranews.com/foto/3823323/penetapan-nomor-urut-pasangan-capres-dan-cawapres-pemilu-2024













X: Modern Days Political Battlefield

Jumlah User Sosial Media*

User Sosial Media vs. Total Populasi*



139.0 Juta



49.9%

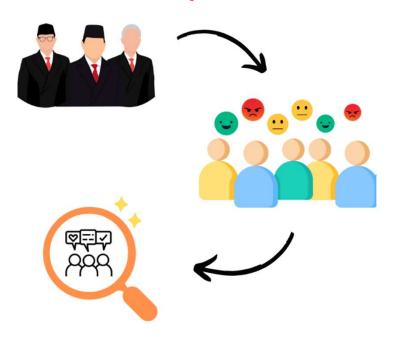
Total User X *



79 Juta

*Per Januari 2024, Indonesia

Analisis Opini Publik















Analisis Opini Publik

SENTIMENT ANALYSIS



TOPIC MODELLING



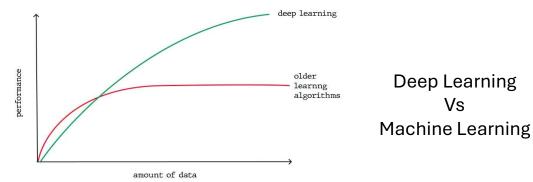
Bagaimana Perkembangan Topik dan Sentimen Untuk Masing-Masing Kandidat Pilpers?

Wisnu dkk. (2020): Pemilihan Gubernur Jawa Tengah

SVM (Sentiment Analysis)



LDA (Topic Modelling)



Penelitian ini menggunakan:

BERT (Sentiment Analysis)



LDA (Topic Modelling)





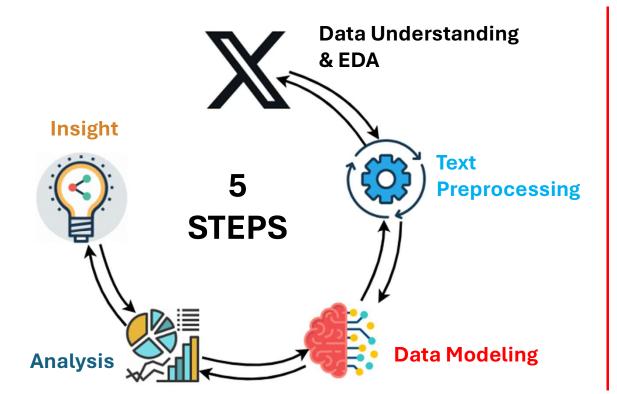


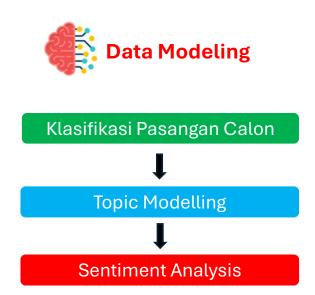






Data Workflow

















Data Understanding & Exploratory Data Analysis

Periode: 31 Desember - 31 Januari 2024

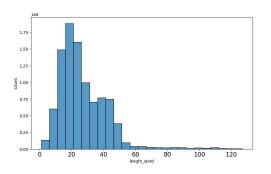
Jumlah Data: 9.817.355

Duplikat Tweet: 8.038.817

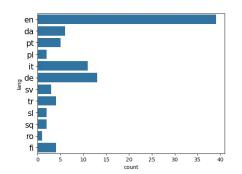
Jumlah User: 9.815.739

created_at	username	tcode	num_retweets	type	frn_cnt	flw_cnt	sts_cnt	loc	lst_cnt	content	lang
04/01/2024	user_1	rt	1248	twit	266	107	9687	NaN	0	RT K-Pope	id
04/01/2024	user_2	rt	195	twit	564	303	12461	indonesia	2	RT Bapak	id

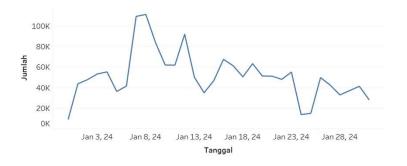
Distribusi Jumlah Kata



Distribusi Bahasa Asing



Trend Tweet Tiap Waktu















Text Preprocessing

Mention, Link, Retweet, Hashtag

RT K-Popers Berencana Kirim Food Truck untuk Anies Baswedan di Acara Desak Anies https://t.co/xp14zfSsTw #TempoSeleb [RE tempodotco]



K-Popers Berencana Kirim Food Truck untuk Anies Baswedan di Acara Desak Anies

Bahasa gaul dan singkatan

Kl gw mau jujur, gw bukan fans berat dr Anis. Bahkan, gw masih skeptis tentang memilih dia



Kalau saya mau jujur, saya bukan penggemar berat dari Anies. Bahkan, saya masih skeptis tentang memilih dia

Bahasa asing

gibran lecehin gelar akademik mahfud md, meski keliru; nganggep profesor a walking encyclopedia.



gibran lecehin gelar akademik mahfud md, meski keliru; nganggep profesor adalah ensiklopedia berjalan.













Klasifikasi Pasangan Calon

Klasifikasi berdasarkan keyword terbanyak dari masing masing pasangan calon

Paslon 01



anies, baswedan, cak, imin, muhaimin, park ahn nice, 01

3.861.574 Tweet

Paslon 02



prabowo, owo, subianto, gibran, raka, rakabuming, 02

2.262.601 Tweet

Paslon 03



ganjar, pranowo, mahfud, md, 03

2.448.668 Tweet





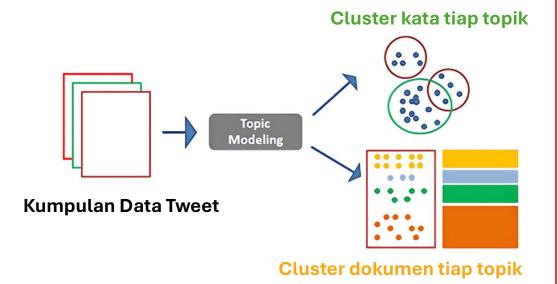






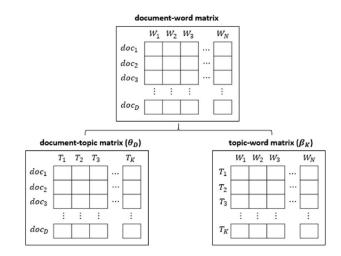


Topic Modelling



Latent Dirichlet Allocation (LDA)

LDA didasarkan pada asumsi setiap topik terdiri dari kata-kata yang mirip, dan setiap dokumen dapat memuat beberapa topik (Blei, 2003)







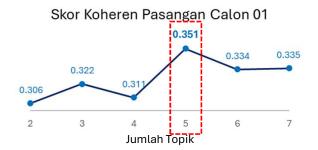






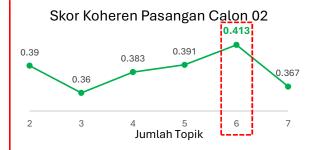


Hasil Topic Modelling Tiap Paslon



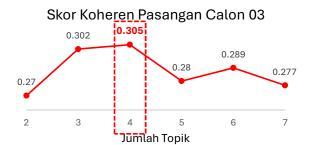


Membahas Terkait Debat, Kampanye, Dukungan, dan Konflik





Membahas Terkait Debat, Kampanye, Dukungan, dan Program Kerja





Membahas Terkait Program Kerja dan Dukungan





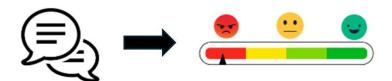




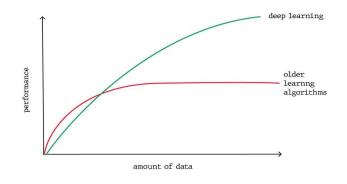


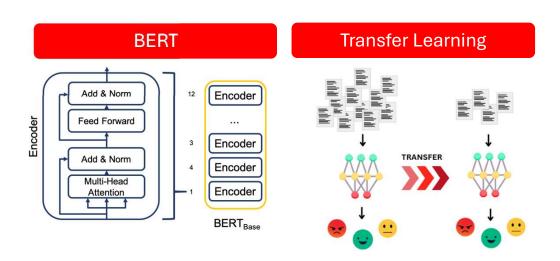


Sentiment Analysis



Machine Learning VS Deep Learning





Model yang digunakan: IndoBERT, IndoRoBERTa

Devlin, J., Chang, M., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv (Cornell University). https://doi.org/10.48550/arxiv.1810.04805













Hasil Evaluasi Pemodelan Sentiment Analysis



2000 data digunakan untuk proses training dan evaluasi menggunakan crossvalidation sebanyak 4 fold

Model akhir yang dipilih untuk proses analisis sentimen adalah model RoBERTa Fine-Tuned karena memiliki evaluasi metrik F1-Score tertinggi sebesar 90.5%







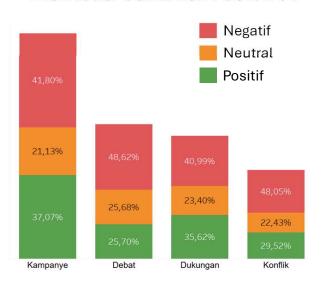






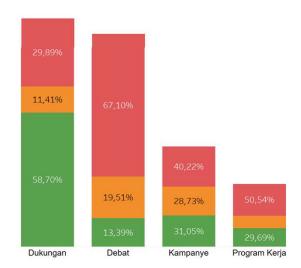
Distribusi Sentimen tiap Topik

Distribusi Sentimen Paslon 01



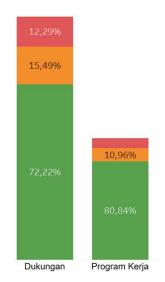


Distribusi Sentimen Paslon 02





Distribusi Sentimen Paslon 03









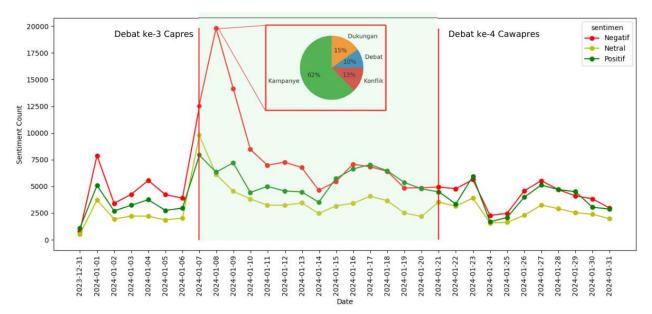








Hasil Analisis Tweet Pada Paslon 1





Anies lupa bahwa demokrasi bukan hanya kata-kata kosong. #UntukKepentinganPribadi

OmonOmon Beretorika



Umpan lambung 03 ke 01 menciptakan gol epic..

Retweetnya dong buat penampilan abah Anies & pak Ganjar semalam!!





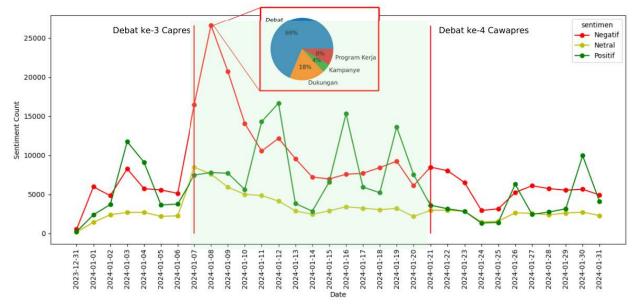








Hasil Analisis Tweet Pada Paslon 2





According to Prabowo, Gaza is opressed because they don't have a sufficient military power. This is such a tone-deaf statement coming from a Minister of Defense.



Paslon presiden nomor urut 2, Prabowo-Gibran, secara tegas menegaskan komitmennya untuk memperjuangkan perdamaian dunia dalam forum-forum internasional.



#PakBOWOTegasMumpuni

Membangun NKRIkuat





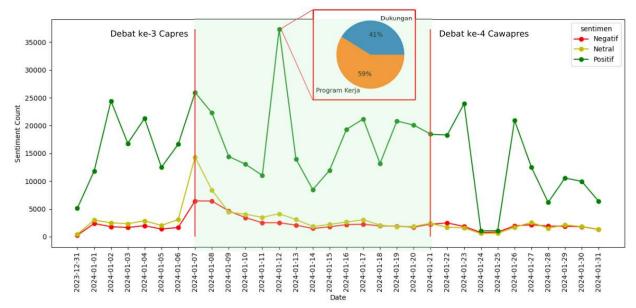








Hasil Analisis Tweet Pada Paslon 3





Pasangan capres-cawapres no urut 3 Ganjar Pranowo & Mahfud MD berkomitmen menghadirkan 17 juta lapangan kerja baru, jika memenangkan Pilpres 2024.

#SahabatGanjar #GanjarPranowo #GenerasiGotongRoyong #GanjarPresiden #GanjarUntukSemua #GanjarMahfud2024 #SatSet #SatSetLebihBaik



Banyak wong cilik yang terbantu sekolah gratis di SMKN Jateng yang didirikan Ganjar Pranowo Sebwah pidio yg mengandung semangat..Untuk menuju Indonesia Unggul













Pengaruh Debat Capres & Anomali pada Paslon 3



H+1 debat

Cara

Kampanye



5.3x lipat dari pra-debat



Anies lupa bahwa demokrasi bukan hanya kata-kata kosong. #UntukKepentinganPribadi

OmonOmon Beretorika



H+1 debat
Performa
Debat



3.4x lipat dari pra-debat



According to Prabowo, Gaza is opressed because they don't have a sufficient military power. This is such a tone-deaf statement coming from a Minister of Defense.



12 Januari **Program Kerja**



2.1x lipat dari rata-rata sentiment positif



Banyak wong cilik yang terbantu sekolah gratis di SMKN Jateng yang didirikan Ganjar Pranowo Sebwah pidio yg mengandung semangat..Untuk menuju Indonesia Unggul











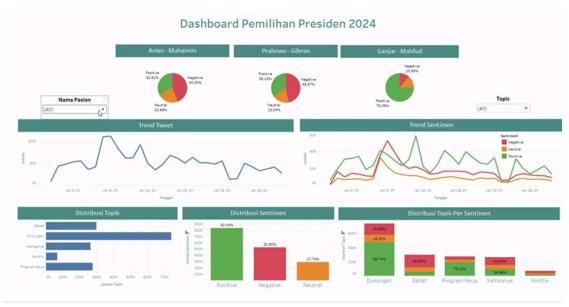


Saran Pengembangan

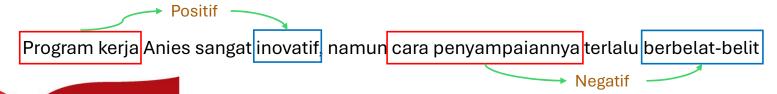
Live Dashboard

Real-time Analysis

Dashboard ABSA Pilpres 2024



Pemodelan Aspect-Based Sentiment Analysis (ABSA)













TERIMA KASIH













Appendix List

- 1. X Introduction
- 2. Aspect Based Sentiment Analysis (ABSA)
- 3. Data Workflow
- 4. Data Understanding & Exploratory Data Analysis
- 5. Text Preprocessing
- 6. Klasifikasi Paslon
- 7. Topic Modelling
- 8. Hasil Topik Modelling Tiap Paslon
- 9. Sentiment Analysis
- 10. Evaluasi Model Sentiment Analysis
- 11. Distribusi Sentimen Tiap Topik
- 12. Pembahasan Paslon 1
- 13. Pembahasan Paslon 2
- 14. Pembahasan Paslon 3
- 15. Pengaruh Debat Capres
- 16. Saran Pengembangan

- A. LDA
- B. BERT
- C. BERT vs RoBERTa
- D. Hasil Topik Modelling Paslon 1
- E. Hasil Topik Modelling Paslon 2
- F. Hasil Topik Modelling Paslon 3
- G. Evaluasi Topik Modelling
- H. Contoh Interpretasi Topik
- I. Active Learning
- J. Kesimpulan Model





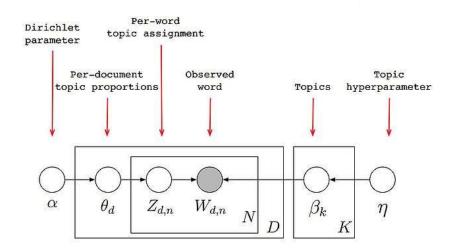


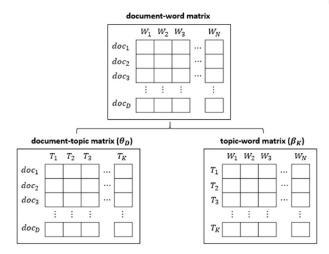




A











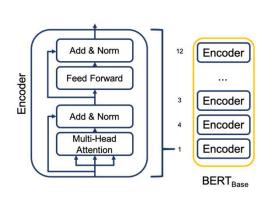


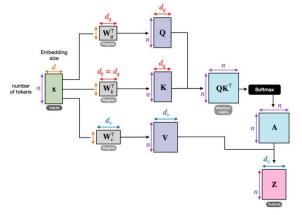


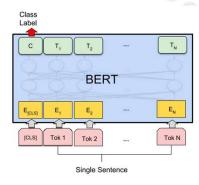


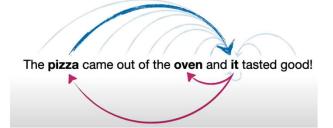
B

BERT













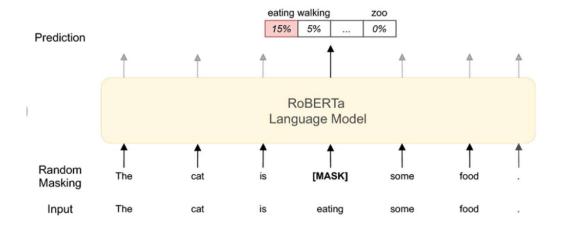






\mathbb{C}

BERT VS RoBERTa



Bert Menggunakan Static Masking, sedangkan Roberta menggunakan Dynamic Masking











Hasil Topic Modelling Paslon 1

Tabel 5. Representasi topik paslon 01

No	Kata kunci	Representasi	Kategori	
1	debat, data, bicara, kinerja	Debat ke-3 terkait data hankam,	Datas	
		kinerja pertahanan keamanan	Debat	
2	desak, partai, pendukung,	Organisasi di belakang anies dan	77	
	masyarakat	desak anies	Kampanye	
3	presiden, kalah, menang,	dukungan terhadap anies-	Dulana	
	putaran	muhaimin sebagai capres.	Dukungan	
4	kampanye, videotron, deserve,	Kampanye anies dan tantangan		
	better	yang dihadapinya	Kampanye	
5	bubble, wan, angin, abud	Konflik antara komunitas	Konflik	
		pendukung vs haters		











Hasil Topic Modelling Paslon 2

Tabel 7. Representasi topik paslon 02

No	Kata kunci	Representasi	Kategori	
1	debat, untuk02, putaran,	Tanggapan Debat ketiga	Debat	
	menang			
2	dukung, partai, kampanye,	Kampanye Partai pengusung	Kampanye	
	putaran	Prabowo.		
3	rakyat, 02menang,	Dukungan rakyat terhadap	Dukungan	
	rakyatsenang, prabowomasgib	paslon.		
4	perubahan, positif, saudara,	Harapan dari masyarakat agar	Dukungan	
	keren	paslon dapat membawa		
		perubahan positif.		
5	nilai, negara, program, korupsi,	Program unggulan Prabowo	Program	
		(pemberantasan korupsi, makan	Kerja	
		siang gratis, kesejahteraan tni)		
6	presiden, dukungan, jabar,	Dukungan rakyat terhadap	Dukungan	
	untukmasyarakat	paslon.		











Tabel 9. Representasi topik paslon 03

No	Kata kunci	Representasi	Kategori	
1	rakyat, program,	program kebijakan unttuk	Program	
	kesejahteraan, kebijakan	menciptakan kesejahteraan dan	Keria	
		kemanan ekonomi		
2	pilpres, jnk, anak, dukung,	dukungan terhadap paslon 3	Dukungan	
	rekam			
3	terbaik, bangsa, hebat,	harapan masyarakat untuk pasangan	Dukungan	
	perubahan, kemajuan	ganjar mahfud		
4	prajurit, hapuskan, kredit,	Program kerja ganjar untuk hapuskan	Program	
	polri.	kredit dan hutang nelayan & petani,	Kerja	
		serta dukungan prajurit.		











Evaluasi Topic Modelling

N Topic	LDA	NMF	SVD
2	0.322	0.346	0.284
3	0.344	0.324	0.275
4	0.344	0.34	0.313
5	0.35	0.33	0.263
6	0.348	0.354	0.341
7	0.372	0.358	0.336
8	0.357	0.364	0.295
9	0.354	0.362	0.309





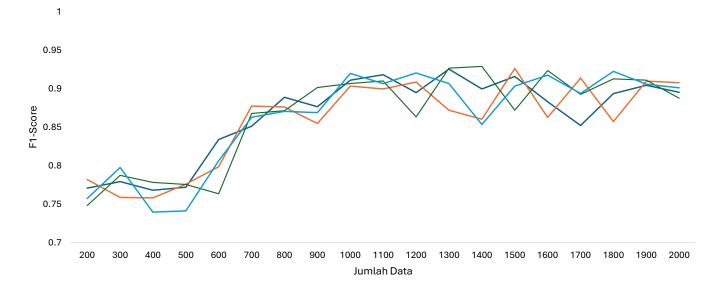






Active Learning

Active Learning Cross Validation









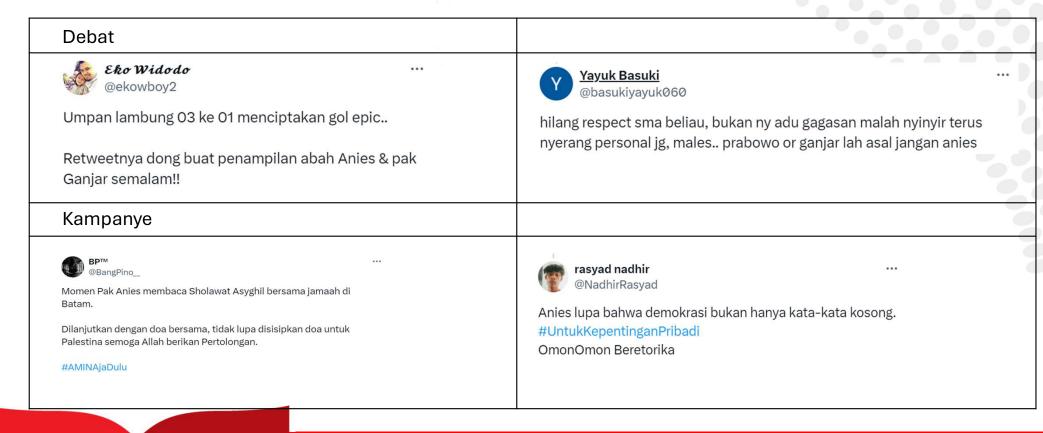








Contoh Interpretasi Topik Anies







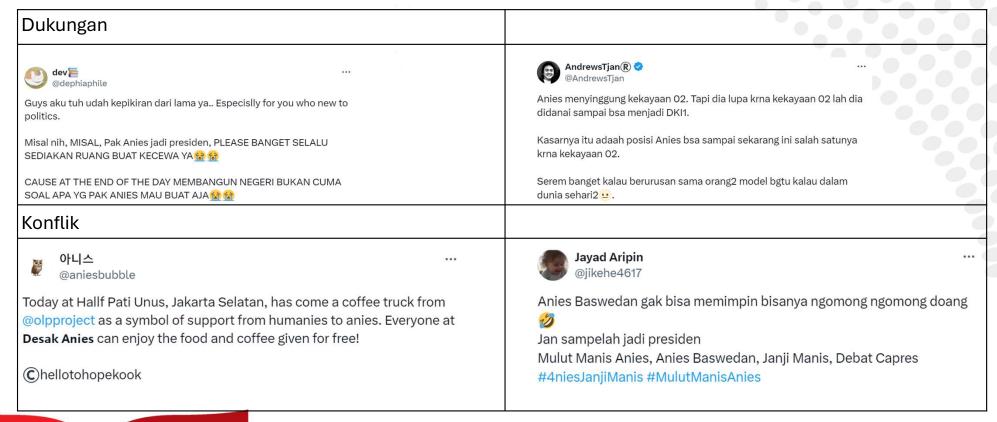






) C







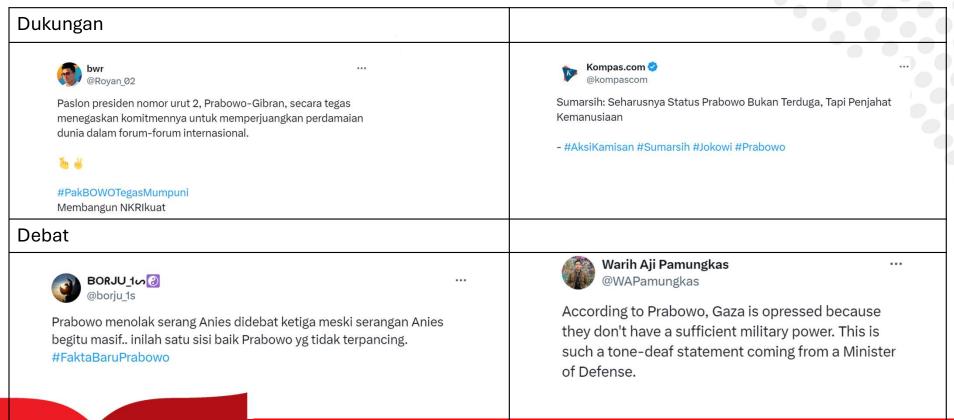








Contoh Interpretasi Topik Prabowo













K



Contoh Interpretasi Topik Prabowo













Contoh Interpretasi Topik Ganjar

Program Kerja Negatif Program Kerja Positif Mbak Dhayu @Dhavu260975 Ganta Pasangan capres-cawapres no urut 3 Ganjar Pranowo & Mahfud MD berkomitmen menghadirkan 17 juta lapangan kerja baru, jika Bro siapapun yg pernah ke Wadas tau seberapa destruktif tangan Ganjar memenangkan Pilpres 2024. terhadap desa indah ini. Betul bahwa ini proyek nasional, tp jangan dikira Ganjar diem aja & ga ngapa-ngapain. Siang-Malem dibombardir oleh #SahabatGanjar #GanjarPranowo #GenerasiGotongRoyong buzzer Ganjar. Ga ad tuh satu pun sikap Ganjar yg berpihak Wadas #GanjarPresiden #GanjarUntukSemua #GanjarMahfud2024 #SatSet #SatSetLebihBaik **Dukungan Positif Dukungan Negatif** Habieb Selow 📀 bantoro @WagimanDeep212 @Boediantar4 Banyak wong cilik yang terbantu sekolah gratis di Ternyata banjir bukan cuma di Semarang, Kudus, Demak, Pekalongan, SMKN Jateng yang didirikan Ganjar Pranowo dan Pati pun banjir. Wes den Ganjar, sementara kurangi ngonten dan Sebwah pidio yg mengandung semangat...Untuk bagi-bagi sembakonya, fokus urusi masyarakat yg kebanjiran... menuju Indonesia Unggul Translate post

N



Kesimpulan Model

Sentiment Analysis

Model terbaik yang digunakan untuk **Analisis Sentimen** adalah model **RoBERTa Fine-Tuned**

Accuracy : 0.922 F1-score : 0.914

precision : 0.916

Recall : 0.913

Jumlah Topik: 5 Skor Koheren: 0.351

Topic Modeling



Jumlah Topik: 6 Skor Koheren: 0.413



Jumlah Topik: 4 Skor Koheren: 0.305