

ANALISIS SENTIMEN TERHADAP PENGGUNAAN APLIKASI SIREKAP PADA PEMILU 2024 MENGGUNAKAN SVM DAN INDOBERTWEET

AULIA AZZAHRA SYAFITRI



**PROGRAM SARJANA ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**

@Hak cipta milik IPB University

IPB University



- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Analisis Sentimen Terhadap Penggunaan Aplikasi Sirekap pada Pemilu 2024 Menggunakan SVM dan IndoBERTweet” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Agustus 2025

Aulia Azzahra Syafitri
G6401211036

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

ABSTRAK

AULIA AZZAHRA SYAFITRI. Analisis Sentimen Terhadap Penggunaan Aplikasi Sirekap pada Pemilu 2024 Menggunakan SVM dan IndoBERTweet. Dibimbing oleh MUHAMMAD ASYHAR AGMALARO.

Pemilu 2024 di Indonesia memperkenalkan penggunaan aplikasi SIREKAP (Sistem Informasi Rekapitulasi) untuk meningkatkan efisiensi dan akurasi dalam perhitungan suara. Penggunaan SIREKAP ini tentunya akan menghasilkan berbagai respons dari publik, salah satunya melalui platform media sosial seperti X (Twitter). Penelitian ini bertujuan untuk melakukan analisis sentimen terhadap opini publik mengenai penggunaan aplikasi SIREKAP pada Pemilu 2024, berdasarkan data dari X. Penelitian ini juga membandingkan kinerja dua model klasifikasi sentimen: *Support Vector Machine* (SVM), algoritma *machine learning* berbasis statistika, dan IndoBERTweet, model BERT yang dilatih khusus untuk *tweet* berbahasa Indonesia. Data yang diperoleh merupakan data *imbalance* sehingga akan dilakukan SMOTE. Hasil penelitian menunjukkan bahwa IndoBERTweet secara keseluruhan memiliki kinerja yang unggul dibandingkan SVM, dengan nilai *accuracy* 69% dan *f-1 score* 68% pada model dengan data *imbalance* sedangkan pada model dengan SMOTE *accuracy* 71% dan *f-1 score* 70%.

Kata kunci: analisis sentimen, IndoBERTweet, SIREKAP, SVM

ABSTRACT

AULIA AZZAHRA SYAFITRI. Sentiment Analysis on the Use of the Sirekap Application in the 2024 Indonesian General Election Using SVM and IndoBERTweet. Supervised by MUHAMMAD ASYHAR AGMALARO.

The 2024 General Election in Indonesia introduced the use of the SIREKAP application (Sistem Informasi Rekapitulasi or Recapitulation Information System) to improve efficiency and accuracy in vote counting. The use of SIREKAP has naturally sparked various public responses, including those expressed on social media platforms such as X (formerly Twitter). This study aims to conduct sentiment analysis on public opinions regarding the use of the SIREKAP application in the 2024 General Election, based on data collected from X. The research also compares the performance of two sentiment classification models: Support Vector Machine (SVM), a statistical based machine learning, and IndoBERTweet, a BERT model specifically trained for Indonesian social media contexts. The dataset used is imbalanced, therefore SMOTE (Synthetic Minority Oversampling Technique) is applied. The results show that IndoBERTweet outperforms SVM overall, achieving an accuracy of 69% and an F1 score of 68% on the imbalanced data model, and an accuracy of 71% and an F1 score of 70% on the SMOTE-applied model.

Keywords: IndoBERTweet, sentiment analysis, SIREKAP, SVM

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2025
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.

**ANALISIS SENTIMEN TERHADAP PENGGUNAAN APLIKASI
SIREKAP PADA PEMILU 2024 MENGGUNAKAN
SVM DAN INDOBERTWEET**

AULIA AZZAHRA SYAFITRI

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Ilmu Komputer

**PROGRAM SARJANA ILMU KOMPUTER
SEKOLAH SAINS DATA, MATEMATIKA, DAN INFORMATIKA
INSTITUT PERTANIAN BOGOR
BOGOR
2025**



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:

- 1 Dr. Aziz Kustiyo, S.Si, M.Kom
- 2 Hari Agung Adrianto, S.Kom., M.Si., Ph.D.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Judul Skripsi : Analisis Sentimen Terhadap Penggunaan Aplikasi Sirekap pada Pemilu 2024 Menggunakan SVM dan IndoBERTweet
Nama : Aulia Azzahra Syafitri
NIM : G6401211036

Disetujui oleh

Pembimbing 1:
Muhammad Asyhar Agmalaro S.Si., M.Kom.

Diketahui oleh

Ketua Program Sarjana Ilmu Komputer:
Dr. Sony Hartono Wijaya, S.Kom., M.Kom
19810809 2008121002

Tanggal Ujian:
13 Agustus 2025

Tanggal Lulus:

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Penelitian yang dilaksanakan sejak bulan Desember 2024 sampai bulan Agustus 2025 berjudul "Analisis Sentimen Terhadap Penggunaan Aplikasi Sirekap pada Pemilu 2024 Menggunakan SVM dan IndoBERTweet".

Terima kasih penulis ucapkan kepada:

- Kedua orang tua penulis, Bapak Ganto Dahlan dan Ibu Sri Purnamalia, adik penulis, M. Fadhil Syaputra, serta keluarga kecil penulis yang senantiasa memberikan kepercayaan, dukungan moral, material, dan doa kepada penulis.
- Bapak Muhammad Asyhar Agmalaro, S.Si., M.Kom. selaku dosen pembimbing yang telah memberikan dukungan, membimbing, dan banyak memberi saran dalam penelitian dan penulisan skripsi ini.
- Bapak Dr. Aziz Kustiyo, S.Si, M.Kom. dan Bapak Hari Agung Adrianto, S.Kom., M.Si., Ph.D. selaku dosen penguji yang telah memberikan masukan, kritik, dan saran yang konstruktif untuk penyempurnaan penulisan ini.
- Ucapan terima kasih juga disampaikan kepada pembimbing akademik, moderator seminar, dan penguji luar komisi pembimbing.
- Rekan-rekan seperjuangan, Assembly, Ilmu Komputer angkatan 58, kakak tingkat angkatan 57, serta teman-teman *Otw Gacoan* yang senantiasa menyebarkan semangat selama berkuliah.
- Teman-teman SMP dan rekan sepermainan yang selalu menjadi pemantik motivasi penulis untuk mengerjakan penelitian dan penulisan skripsi ini.
- Semua pihak yang tidak dapat disebutkan satu per satu yang telah membantu dalam penyelesaian skripsi ini.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan.

Bogor, Agustus 2025

Aulia Azzahra Syafitri

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	x
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Manfaat	2
1.5 Ruang Lingkup	3
II TINJAUAN PUSTAKA	4
2.1 Analisis Sentimen	4
2.2 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	4
2.3 <i>Support Vector Machine</i>	5
2.4 IndoBERTweet	5
2.5 <i>Cohen's Kappa</i>	7
2.6 <i>Confusion Matrix</i>	8
III METODE	10
3.1 Tahapan Penelitian	10
3.2 Lingkungan Penelitian	13
IV HASIL DAN PEMBAHASAN	14
4.1 Pengambilan Data Twitter	14
4.2 Praproses Data	14
4.3 Pelabelan Data	17
4.4 Pembobotan Kata TF-IDF dan BERT <i>Tokenization</i>	19
4.5 Pembagian Data	20
4.6 Pemodelan	20
4.7 Evaluasi	21
SIMPULAN DAN SARAN	26
4.8 Simpulan	26
4.9 Saran	26
DAFTAR PUSTAKA	27
LAMPIRAN	1
RIWAYAT HIDUP	4



Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

DAFTAR TABEL

1	Rumus <i>Cohen's Kappa</i>	8
2	Kriteria pelabelan per kategori	12
3	Contoh data yang dieliminasi	14
4	Hasil <i>case folding</i>	15
5	Hasil <i>data cleaning</i>	15
6	Hasil <i>tokenizing</i>	16
7	Contoh daftar kata <i>slang</i> dan kata baku pada <i>kamuskatabaku</i> dan <i>kamusalay</i>	16
8	Hasil normalisasi bahasa	16
9	Hasil <i>filtering</i>	17
10	Hasil <i>stemming</i>	17
11	Hasil perhitungan <i>Cohen's Kappa</i>	18
12	Kombinasi <i>hyperparameter tuning</i> pada model SVM	20
13	Hasil parameter optimal masing-masing kernel pada model SVM	21
14	Kombinasi <i>hyperparameter tuning</i> pada model IndoBERTweet	21
15	Hasil evaluasi data uji pada model SVM	22
16	Hasil evaluasi data uji pada model IndoBERTweet	22

DAFTAR GAMBAR

1	Perhitungan TF-IDF (Almarashy <i>et al.</i> 2023)	4
2	Ilustrasi <i>hyperplane</i> dan <i>margin</i> maksimal (Fremmuzar dan Baita 2023)	5
3	Ilustrasi <i>pre-training</i> dan <i>fine-tuning</i> BERT (Devlin <i>et al.</i> 2019)	
4	Ilustrasi tokenisasi BERT (Nugroho <i>et al.</i> 2021)	7
5	<i>Confusion matrix</i> anotator (Sabharwal 2021)	8
6	Rentang interpretasi nilai kappa (Sabharwal 2022)	8
7	Ilustrasi <i>confusion matrix</i> pada: a) <i>binary classification</i> dan b) <i>multiclass classification</i> (Markoulidakis <i>et al.</i> 2021)	8
8	Diagram tahapan penelitian	10
9	Diagram sebaran jumlah <i>tweet</i> bulan Februari 2024	14
10	Distribusi pelabelan per kategori antara: a) peneliti dengan AK; b) peneliti dengan MYDF; dan c) MYDF dengan AK	18
11	Persentase hasil pelabelan data	19
12	<i>Wordcloud</i> dari pembobotan kata TF-IDF: a) kategori positif, b) kategori netral, dan c) kategori negatif	19
13	Tokenisasi BERT	20
14	Diagram <i>confusion matrix</i> data uji pada model SVM dengan: a) data <i>imbalance</i> dan b) menerapkan SMOTE	23
15	Diagram <i>confusion matrix</i> data uji pada model IndoBERTweet dengan: a) data <i>imbalance</i> dan b) menerapkan SMOTE	24