



**ANALISIS SENTIMEN MASYARAKAT TWITTER TERHADAP JUDI
ONLINE DI INDONESIA MENGGUNAKAN METODE INDOROBERTA
DAN SVM TF-IDF**

LAPORAN TUGAS AKHIR

MUHAMMAD ALVIN SYAHRIN

41521010137

**UNIVERSITAS
MERCU BUANA**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**



**ANALISIS SENTIMEN MASYARAKAT TWITTER TERHADAP JUDI
ONLINE DI INDONESIA MENGGUNAKAN METODE INDOROBERTA
DAN SVM TF-IDF**

LAPORAN TUGAS AKHIR

MUHAMMAD ALVIN SYAHRIN

41521010137

**UNIVERSITAS
MERCU BUANA**

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Alvin Syahrin
NIM : 41521010137
Program Studi : Teknik Informatika
Judul Proposal Penelitian : Analisis Sentimen Masyarakat Twitter Terhadap Judi Online di Indonesia Menggunakan Metode IndoRoBERTa dan SVM TF-IDF

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 19 Juli 2025



Muhammad Alvin Syahrin

UNIVERSITAS
MERCU BUANA

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Muhammad Alvin Syahrin
NIM : 41521010137
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Sentimen Masyarakat Twitter
Terhadap Judi Online di Indonesia
Menggunakan Metode IndoRoBERTa dan
SVM TF-IDF

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Ida Farida, S.T., M.Kom.

NIDN : 0324018301



Ketua Penguji : Saruni Dwiasnati, ST., MM, M.Kom

NIDN : 0325128802



Penguji 1 : Misni, S.Kom., M.Kom.

NIDN : 0413046802



Penguji 2 : Muhaimin Hasanudin, S.T., M.Kom.

NIDN : 0420027508



Jakarta, 19 Juli 2025

Mengetahui,

Dekan

Ketua Program Studi



Dr. Bambang Jokonowo, S.Si., MTI
NIDN : 0320037002



Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701

KATA PENGANTAR

Puji syukur kehadiran Tuhan yang Maha Esa, atas segala rahmat dan ridha-Nya sehingga penulis dapat menyelesaikan proposal penelitian yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

Penulis menyadari bahwa proposal penelitian ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Ibu Ida Farida, ST,M.Kom. selaku dosen pembimbing TA yang telah memberikan pengarahan, motivasi, menyediakan waktu, tenaga, dan pikiran sehingga selama pembuatan proposal penelitian ini terjadwal dengan baik.
5. Kedua Orang Tua saya yang selalu mensuport dan mendukung saya selama menjalani masa studi sebagai mahasiswa Universitas Mercubuana.
6. Semua teman kuliah yang selalu berbagi informasi dan memberikan dukungan dalam bentuk yang berbeda-beda.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 19 Juli 2025



Muhammad Alvin Syahrin

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Muhammad Alvin Syahrin
NIM : 41521010137
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Sentimen Masyarakat Twitter Terhadap Judi Online
di Indonesia Menggunakan Metode IndoRoBERTa dan SVM
TF-IDF

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

UNIVERSITAS
MERCU BUANA

Jakarta, 19 Juli 2025

Yang menyatakan,



Muhammad Alvin Syahrin

ABSTRAK

Nama : Muhammad Alvin Syahrin
NIM : 41521010137
Program Studi : Teknik Informatika
Analisis Sentimen Masyarakat Twitter Terhadap
Judul Proposal Penelitian : Judi Online di Indonesia Menggunakan Metode
IndoRoBERTa dan SVM TF-IDF
Dosen Pembimbing : Ida Farida, ST, M.Kom

Perkembangan teknologi informasi dan komunikasi telah mendorong maraknya aktivitas media sosial sebagai wadah ekspresi masyarakat, termasuk dalam menyuarakan opini terkait isu-isu sosial seperti judi online. Penelitian ini bertujuan untuk menganalisis sentimen masyarakat Indonesia di platform Twitter terhadap fenomena judi online dengan menggunakan dua pendekatan algoritma: IndoRoBERTa dan SVM TF-IDF. Dataset yang digunakan terdiri dari 16.899 tweet berbahasa Indonesia yang dikumpulkan selama periode dua bulan. Data tersebut diklasifikasikan ke dalam tiga kategori sentimen, yaitu positif, netral, dan negatif. Penelitian ini mengimplementasikan model IndoRoBERTa sebagai representasi dari pendekatan deep learning berbasis transformer, serta SVM dengan representasi fitur TF-IDF sebagai pendekatan machine learning konvensional. Hasil evaluasi menunjukkan bahwa model IndoRoBERTa mencapai akurasi sebesar 81,27%, sedangkan SVM TF-IDF memperoleh akurasi sebesar 79,74%. Meskipun terdapat ketidakseimbangan kelas pada dataset dengan dominasi sentimen negatif sebesar 55,6% model IndoRoBERTa mampu mengklasifikasikan sentimen secara lebih akurat dan stabil. Temuan penelitian mengungkapkan bahwa mayoritas masyarakat Indonesia memiliki persepsi negatif terhadap judi online. Penelitian ini juga merekomendasikan penggunaan teknik penyeimbangan data, perluasan sumber dataset, serta eksplorasi pendekatan model hibrida dalam penelitian selanjutnya.

Kata kunci: Judi Online, Analisis Sentimen, Twitter, NLP, IndoRoBERTa, SVM, TF-IDF

ABSTRACT

Nama : Muhammad Alvin Syahrin
NIM : 41521010137
Program Studi : Teknik Informatika
Analisis Sentimen Masyarakat Twitter Terhadap
Judul Proposal Penelitian : Judi Online di Indonesia Menggunakan Metode
IndoRoBERTa dan SVM TF-IDF
Dosen Pembimbing : Ida Farida, ST, M.Kom

The rapid advancement of information and communication technology has fueled the widespread use of social media as a platform for public expression, including opinions on social issues such as online gambling. This study aims to analyze public sentiment in Indonesia on Twitter regarding the phenomenon of online gambling using two algorithmic approaches: IndoRoBERTa and SVM with TF-IDF. The dataset consists of 11,841 Indonesian-language tweets collected over a two-month period. The tweets were categorized into three sentiment classes: positive, neutral, and negative. This research implements the IndoRoBERTa model as a deep learning approach based on the transformer architecture, and SVM with TF-IDF as a conventional machine learning method. Evaluation results show that the IndoRoBERTa model achieved an accuracy of 81.27%, while the SVM TF-IDF model achieved 79.74%. Despite class imbalance in the dataset where negative sentiment dominates at 55.6% IndoRoBERTa demonstrated more accurate and stable performance in sentiment classification. The findings indicate that the majority of Indonesian users express negative perceptions toward online gambling. The study also recommends the application of data balancing techniques, expansion of data sources, and exploration of hybrid model approaches for future research.

Kata kunci: Online Gambling, Sentiment Analysis, Twitter, NLP, IndoRoBERTa, SVM, TF-IDF

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	v
ABSTRAK	vi
ABSTRACT	vii
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMPIRAN	xiii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Manfaat Penelitian	3
1.5 Batasan Masalah	4
BAB II TINJAUAN PUSTAKA	5
2.1 Penelitian Terdahulu	5
2.2 Teori Pendukung.....	19
2.2.1 Analisis Sentimen	19
2.2.2 Media Sosial X (Twitter).....	19
2.2.3 Judi Online di Indonesia	20
2.2.4 Natural Language Processing (NLP).....	21
2.2.5 Text Mining.....	21
2.2.6 Pre-processing Text	22
2.2.7 BERT (Bidirectional Encoder Representations Transformers)	24

2.2.8 IndoRoBERTa (Indonesian Robustly Optimized BERT Pretraining Approach)	25
2.2.9 SVM (Support Vector Machine)	27
2.2.10 TF-IDF (Term Frequency - Inverse Document Frequency)	28
2.2.11 Confusion Matrix	29
2.3 Gap Penelitian	30
BAB III METODE PENELITIAN	33
3.1 Jenis Penelitian.....	33
3.2 Tahapan Penelitian.....	35
BAB IV PEMBAHASAN	42
4.1 Data Understanding	42
4.3 Pre-Processing Text	45
4.4 Pelabelan Data	49
4.5 Split Data	51
4.6 Modelling.....	52
4.6.1 IndoRoBERTa	52
4.6.2 SVM + TF-IDF.....	58
4.7 Evaluasi Model	63
4.7.1 IndoRoBERTa	64
4.7.2 SVM + TF-IDF.....	68
4.8 Visualisasi Hasil.....	70
BAB V KESIMPULAN DAN SARAN	76
5.1 Kesimpulan	76
5.2 Saran	77
DAFTAR PUSTAKA.....	78
LAMPIRAN.....	84

DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait	5
Tabel 2. 2 Perbandingan Penelitian Terdahulu	18
Tabel 4. 1 Hasil Cleaning Data	46
Tabel 4. 2 Hasil Case Folding.....	46
Tabel 4. 3 Hasil Normalisasi Teks	47
Tabel 4. 4 Hasil Tokenizing	48
Tabel 4. 5 Hasil Stopword Removal	48
Tabel 4. 6 Hasil Stemming.....	49
Tabel 4. 7 Perbandingan Hasil Labeling.....	51
Tabel 4. 8 Hasil Pelabelan Menggunakan Lexicon Based.....	51
Tabel 4. 9 Metrik Akurasi dan Loss Hasil Pelatihan	65



DAFTAR GAMBAR

Gambar 2. 1 Perkembangan Judi Online Di Indonesia.....	20
Gambar 2. 2 Arsitektur Transformers.....	25
Gambar 2. 3 Arsitektur Model RoBERTa.....	26
Gambar 2. 4 Algoritma Support Vector Machine.....	28
Gambar 2. 5 Confusion Matrix.....	29
Gambar 3. 1 Alur Tahapan Penelitian.....	36
Gambar 3. 1 Alur Tahapan Penelitian.....	36
Gambar 4. 1 Crawling Data.....	42
Gambar 4. 1 Crawling Data.....	42
Gambar 4. 2 Dataset Hasil Crawling.....	43
Gambar 4. 2 Dataset Hasil Crawling.....	43
Gambar 4. 3 Informasi Dataset.....	43
Gambar 4. 3 Informasi Dataset.....	43
Gambar 4. 4 Dataset Setelah Penghapusan Kolom.....	44
Gambar 4. 4 Dataset Setelah Penghapusan Kolom.....	44
Gambar 4. 5 Contoh Komentar Tweet Promosi Judi Online.....	44
Gambar 4. 5 Contoh Komentar Tweet Promosi Judi Online.....	44
Gambar 4. 6 Informasi Data Duplikat.....	45
Gambar 4. 6 Informasi Data Duplikat.....	45
Gambar 4. 7 Informasi Data Kosong (null).....	45
Gambar 4. 7 Informasi Data Kosong (null).....	45
Gambar 4. 8 Kamus Slang Words.....	47
Gambar 4. 8 Kamus Slang Words.....	47
Gambar 4. 9 Kamus Lexicon InSet.....	50
Gambar 4. 9 Kamus Lexicon InSet.....	50
Gambar 4. 10 Kode Split Data.....	51
Gambar 4. 11 Kode Mengubah Label Teks Menjadi Angka.....	52
Gambar 4. 11 Kode Mengubah Label Teks Menjadi Angka.....	52
Gambar 4. 12 Kode Memuat Tokenizer dan Model.....	52
Gambar 4. 12 Kode Memuat Tokenizer dan Model.....	52
Gambar 4. 13 Kode Mengubah Kalimat Teks Menjadi Format Numerik.....	53
Gambar 4. 13 Kode Mengubah Kalimat Teks Menjadi Format Numerik.....	53
Gambar 4. 14 Kode Menampilkan Hasil Tokenisasi.....	54
Gambar 4. 14 Kode Menampilkan Hasil Tokenisasi.....	54
Gambar 4. 15 Kode Encode Data.....	55
Gambar 4. 15 Kode Encode Data.....	55
Gambar 4. 16 Proses Encoding (Tokenisasi + Formatting).....	56
Gambar 4. 16 Proses Encoding (Tokenisasi + Formatting).....	56
Gambar 4. 17 Memuat Model IndoRoBERTa.....	56
Gambar 4. 17 Memuat Model IndoRoBERTa.....	56
Gambar 4. 18 Compile Model.....	57
Gambar 4. 18 Compile Model.....	57
Gambar 4. 19 Pelatihan Model IndoRoBERTa.....	57
Gambar 4. 19 Pelatihan Model IndoRoBERTa.....	57

Gambar 4. 20 Import Library	58
Gambar 4. 20 Import Library	58
Gambar 4. 21 Persiapan Data.....	59
Gambar 4. 21 Persiapan Data.....	59
Gambar 4. 22 Pemisahan Fitur dan Konversi Label Menjadi Numerik.....	60
Gambar 4. 22 Pemisahan Fitur dan Konversi Label Menjadi Numerik.....	60
Gambar 4. 23 Membuat Mapping Label.....	60
Gambar 4. 23 Membuat Mapping Label.....	60
Gambar 4. 24 Menerapkan dan Menyimpan Mapping Label	61
Gambar 4. 24 Menerapkan dan Menyimpan Mapping Label	61
Gambar 4. 25 Kode Split Data	62
Gambar 4. 25 Kode Split Data	62
Gambar 4. 26 Kode Feature Extraction TF-IDF	62
Gambar 4. 26 Kode Feature Extraction TF-IDF	62
Gambar 4. 27 Hyperparameter Tuning Grid Search.....	63
Gambar 4. 27 Hyperparameter Tuning Grid Search.....	63
Gambar 4. 28 Grafik Hasil Training IndoRoBERTa	64
Gambar 4. 28 Grafik Hasil Training IndoRoBERTa	64
Gambar 4. 29 Hasil Akurasi Model IndoRoBERTa	66
Gambar 4. 29 Hasil Akurasi Model IndoRoBERTa	66
Gambar 4. 30 Laporan Klasifikasi IndoRoBERTa	67
Gambar 4. 30 Laporan Klasifikasi IndoRoBERTa	67
Gambar 4. 31 Confusion Matrix IndoRoBERTa	68
Gambar 4. 31 Confusion Matrix IndoRoBERTa	68
Gambar 4. 32 Laporan Klasifikasi SVM	68
Gambar 4. 32 Laporan Klasifikasi SVM	68
Gambar 4. 33 Confusion Matrix SVM	69
Gambar 4. 33 Confusion Matrix SVM	69
Gambar 4. 34 Grafik Distribusi Klasifikasi Sentimen	70
Gambar 4. 34 Grafik Distribusi Klasifikasi Sentimen	70
Gambar 4. 35 Wordcloud Positif	71
Gambar 4. 35 Wordcloud Positif	71
Gambar 4. 36 Wordcloud Negatif.....	72
Gambar 4. 36 Wordcloud Negatif.....	72
Gambar 4. 37 Wordcloud Netral.....	73
Gambar 4. 37 Wordcloud Netral.....	73
Gambar 4. 38 Grafik Perbandingan Performa IndoRoBERTa dan SVM.....	73
Gambar 4. 38 Grafik Perbandingan Performa IndoRoBERTa dan SVM.....	73

DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi	84
Lampiran 2 Curriculum Vitae (CV)	85
Lampiran 3 Surat Pernyataan HAKI.....	86
Lampiran 4 Sertifikat BNSP	88
Lampiran 5 Form Revisi Dosen Penguji.....	89
Lampiran 6 Hasil Cek Turnitin	91
Lampiran 7 Halaman Persetujuan	92

