

**PENGEMBANGAN APLIKASI DETEKSI DAN PENYARINGAN KOMENTAR
JUDI ONLINE PADA YOUTUBE DENGAN
SUPPORT VECTOR MACHINE**

SKRIPSI

Oleh:

**CALVIN WIJAYA
NIM. 221112275
CALVIN WINATA
NIM. 221111448**



**PROGRAM STUDI S-1 TEKNIK INFORMATIKA
FAKULTAS INFORMATIKA
UNIVERSITAS MIKROSKIL
MEDAN
2026**

DEVELOPMENT OF AN APPLICATION FOR DETECTING AND FILTERING ONLINE GAMBLING COMMENTS ON YOUTUBE USING SUPPORT VECTOR MACHINE

UNDERGRADUATE THESIS

By:

**CALVIN WIJAYA
ID NUMBER. 221112275
CALVIN WINATA
ID NUMBER. 221111448**



**UNDERGRADUATE PROGRAM OF INFORMATICS ENGINEERING
FACULTY OF INFORMATICS
UNIVERSITAS MIKROSKIL
MEDAN
2026**

LEMBARAN PENGESAHAN

PENGEMBANGAN APLIKASI DETEKSI DAN PENYARINGAN
KOMENTAR JUDI ONLINE PADA YOUTUBE DENGAN
SUPPORT VECTOR MACHINE

SKRIPSI

Diajukan untuk Melengkapi Persyaratan Guna
Mendapatkan Gelar Sarjana
Program Studi S-1 Teknik Informatika

Oleh:

CALVIN WIJAYA
NIM. 221112275
CALVIN WINATA
NIM. 221111448

Disetujui Oleh:

Dosen Pembimbing I,


Ng Poi Wong, S.Kom., M.T.I

Dosen Pembimbing II,


Mirza Ihami, S.Kom., M.T.I

Medan, 30 Juli 2026
Diketahui dan Disahkan Oleh:

Ketua Program Studi
S-1 Teknik Informatika


Carles Juliandy, S.Kom., M.Kom.

HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Calvin Wijaya
NIM : 221112275

Saya telah melaksanakan penelitian dan penulisan Skripsi dengan judul dan tempat penelitian sebagai berikut:

Judul Skripsi : Pengembangan Aplikasi Deteksi dan Penyaringan Komentar Judi Online pada Youtube dengan Support Vector Machine
Tempat Penelitian : -
Alamat Tempat Penelitian : -
No. Telp. Tempat Penelitian : -

Sehubungan dengan Skripsi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Skripsi saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Skripsi saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 25 Juni 2026

Saya yang membuat pernyataan,



HALAMAN PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Calvin Winata

NIM : 221111448

Saya telah melaksanakan penelitian dan penulisan Skripsi dengan judul dan tempat penelitian sebagai berikut:

Judul Skripsi : Pengembangan Aplikasi Deteksi dan Penyaringan Komentar Judi Online pada Youtube dengan Support Vector Machine

Tempat Penelitian : -

Alamat Tempat Penelitian : -

No. Telp. Tempat Penelitian : -

Sehubungan dengan Skripsi tersebut, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Skripsi tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh Universitas Mikroskil Medan, yakni pencabutan ijazah yang telah saya terima dan ijazah tersebut dinyatakan tidak sah.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada Universitas Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Skripsi saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, Universitas Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Skripsi saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Skripsi saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak/sistem informasi.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 25 Juni 2026

Saya yang membuat pernyataan,



PENGEMBANGAN APLIKASI DETEKSI DAN PENYARINGAN KOMENTAR JUDI ONLINE PADA YOUTUBE DENGAN SUPPORT VECTOR MACHINE

Abstrak

Maraknya komentar yang mengandung promosi judi online pada YouTube menunjukkan perlunya sistem yang mampu melakukan deteksi secara otomatis. Penelitian ini bertujuan mengembangkan aplikasi deteksi dan penyaringan komentar judi online menggunakan IndoBERT sebagai ekstraksi fitur dan Support Vector Machine (SVM) sebagai metode klasifikasi. Tahapan penelitian meliputi pengumpulan dan pelabelan data, prapemrosesan teks, ekstraksi fitur menggunakan token [CLS] dari IndoBERT, serta klasifikasi menggunakan SVM yang dioptimalkan melalui Grid Search. Aplikasi dikembangkan menggunakan Streamlit dan terintegrasi dengan YouTube Data API v3 untuk mengambil, mendeteksi, dan mengelola komentar pada video YouTube. Hasil pengujian menunjukkan bahwa model mampu mengklasifikasikan komentar judi online dan non-judi dengan baik berdasarkan nilai accuracy sebesar 96,13%, dengan precision, recall, dan F1-score (rata-rata macro) masing-masing sebesar 95,76%, 93,19%, dan 94,39%. Selain itu, aplikasi yang dikembangkan menyediakan fitur penyaringan dan penghapusan komentar yang teridentifikasi sebagai promosi judi online sehingga dapat mendukung proses moderasi secara lebih efisien. Dengan demikian, sistem yang dihasilkan mampu membantu kreator maupun pengelola kanal YouTube dalam mengurangi penyebaran komentar bermuatan judi online secara lebih efektif, cepat, dan konsisten.

Kata kunci: YouTube, Judi Online, IndoBERT, Support Vector Machine, Grid Search

Abstract

The increasing number of online gambling promotional comments on YouTube highlights the need for an automated detection system. This study aims to develop an application for detecting and filtering online gambling comments using IndoBERT for feature extraction and Support Vector Machine (SVM) for classification. The research process includes data collection and labeling, text preprocessing, feature extraction using the [CLS] token from IndoBERT, and classification using an SVM model optimized through Grid Search. The application was developed using Streamlit and integrated with the YouTube Data API v3 to retrieve, detect, and manage comments on YouTube videos. The evaluation results indicate that the proposed model successfully classifies gambling and non-gambling comments, achieving an accuracy of 96.13%, with macro-average precision, recall, and F1-score of 95.76%, 93.19%, and 94.39%, respectively. In addition, the application provides features for filtering and deleting comments identified as online gambling promotions, thereby supporting a more efficient moderation process. Therefore, the developed system can assist YouTube content creators and channel administrators in reducing the spread of online gambling promotional comments more effectively, quickly, and consistently.

Keywords: YouTube, Online Gambling, IndoBERT, Support Vector Machine, Grid Search

KATA PENGANTAR

Puji Syukur kepada Tuhan Yang Maha Esa berkat rahmat dan karunia-Nya penulis dapat menyelesaikan penulisan tugas akhir ini dengan baik dan selesai tepat waktu yang berjudul “Pengembangan Aplikasi Deteksi dan Penyaringan Komentar Judi Online pada YouTube dengan Support Vector Machine”.

Penulisan tugas akhir ini dilakukan untuk memenuhi persyaratan dari akademik untuk menyelesaikan Program Sarjana Strata Satu (S1) Program Studi Teknik Informatika pada Fakultas Informatika di Universitas Mikroskil Medan.

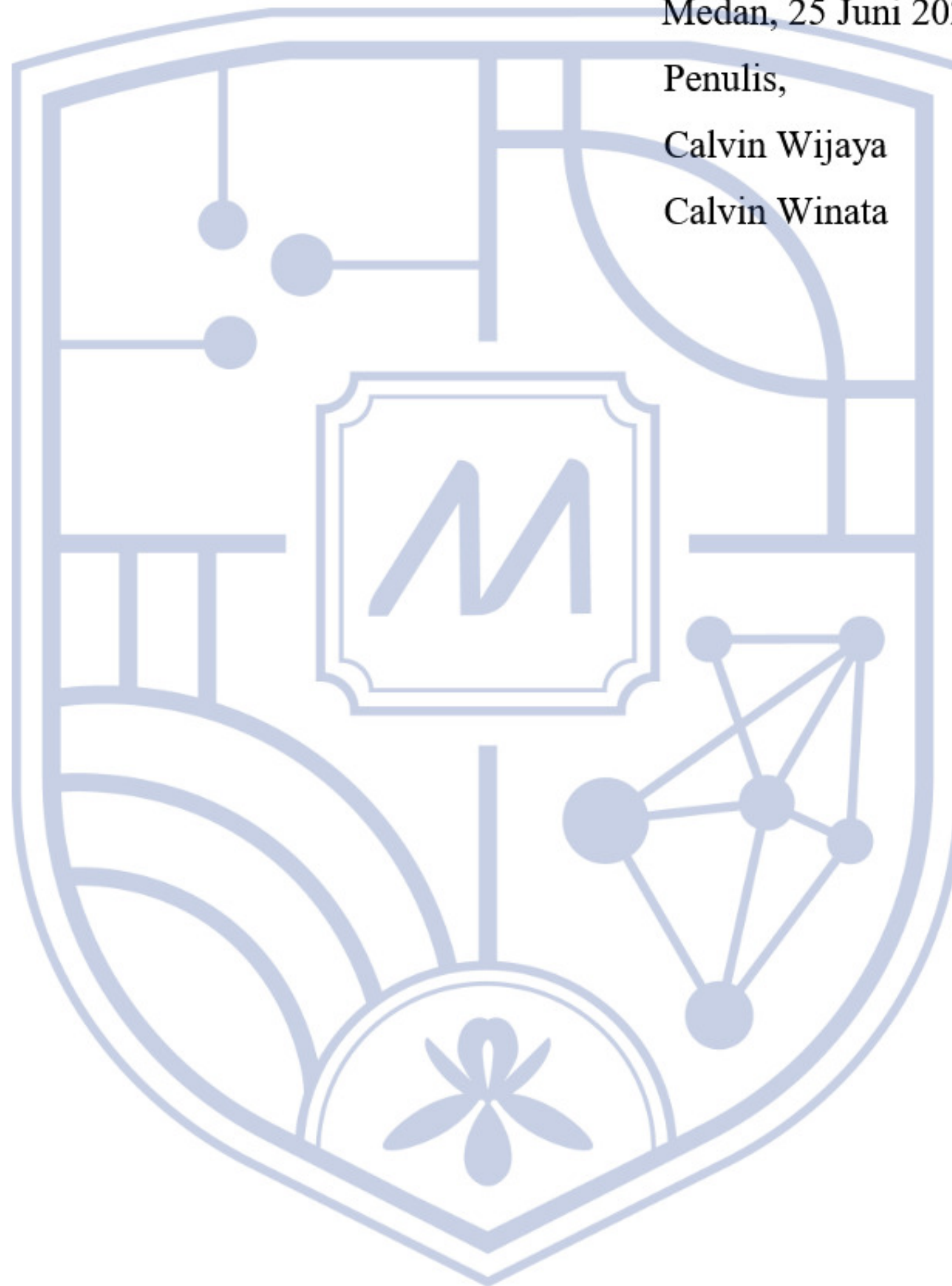
Dalam penulisan tugas akhir ini, penulis menyadari banyaknya rintangan dan hambatan yang terjadi selama proses pengerjaan tugas akhir, namun berkat adanya dukungan, motivasi, doa, bimbingan, dan bantuan dari berbagai pihak, akhirnya penulis dapat menyelesaikan tugas akhir ini. Dalam kesempatan ini, penulis ingin mengucapkan terima kasih kepada:

1. Bapak Ng Poi Wong, S.Kom.,M.T.I., selaku Dosen Pembimbing I yang telah bersedia meluangkan waktu dan memberikan bantuan berupa bimbingan, pengetahuan dan arahan selama proses penulisan Tugas Akhir ini.
2. Bapak Mirza Ilhami, S.Kom.,M.TI., selaku Dosen Pembimbing II yang telah bersedia meluangkan waktu dan memberikan bantuan berupa bimbingan, pengetahuan dan arahan selama proses penulisan Tugas Akhir ini.
3. Bapak Hardy, S.Kom., M.Sc., Ph.D., selaku Rektor Universitas Mikroskil Medan.
4. Bapak Djoni, S.Kom., M.T.I., selaku Wakil Rektor I Universitas Mikroskil Medan.
5. Bapak Sunaryo Winardi, S.Kom., M.T., selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
6. Bapak Carles Juliandy, S.Kom., M.Kom. , selaku Ketua Program Studi S-1 Teknik Informatika Universitas Mikroskil Medan.
7. Ibu Hernawati Gohzali, S.Kom., M.TI., selaku Dosen Wali Teknik Informatika C Sore.
8. Seluruh dosen Universitas Mikroskil yang telah banyak memberikan pelajaran serta ilmu yang berharga kepada penulis selama duduk di bangku kuliah.
9. Orang tua dan keluarga penulis yang telah memberikan bimbingan, semangat, doa dan dukungan kepada penulis selama proses penulisan Tugas Akhir ini.
10. Teman-teman mahasiswa dan seluruh pihak yang telah memberikan dukungan dalam menyelesaikan Tugas Akhir ini yang tidak dapat disebutkan satu persatu.

Penulis menyadari bahwa dalam penulisan tugas akhir ini masih banyak kekurangan dan keterbatasan. Oleh karena itu, penulis mengharapkan kritik dan saran dari pembaca yang membangun demi kesempurnaan tugas akhir ini. Semoga tugas akhir ini dapat memberikan manfaat dan kontribusi bagi semua pihak serta bagi pengembangan ilmu pengetahuan, khususnya di bidang Teknik Informatika.

Medan, 25 Juni 2026

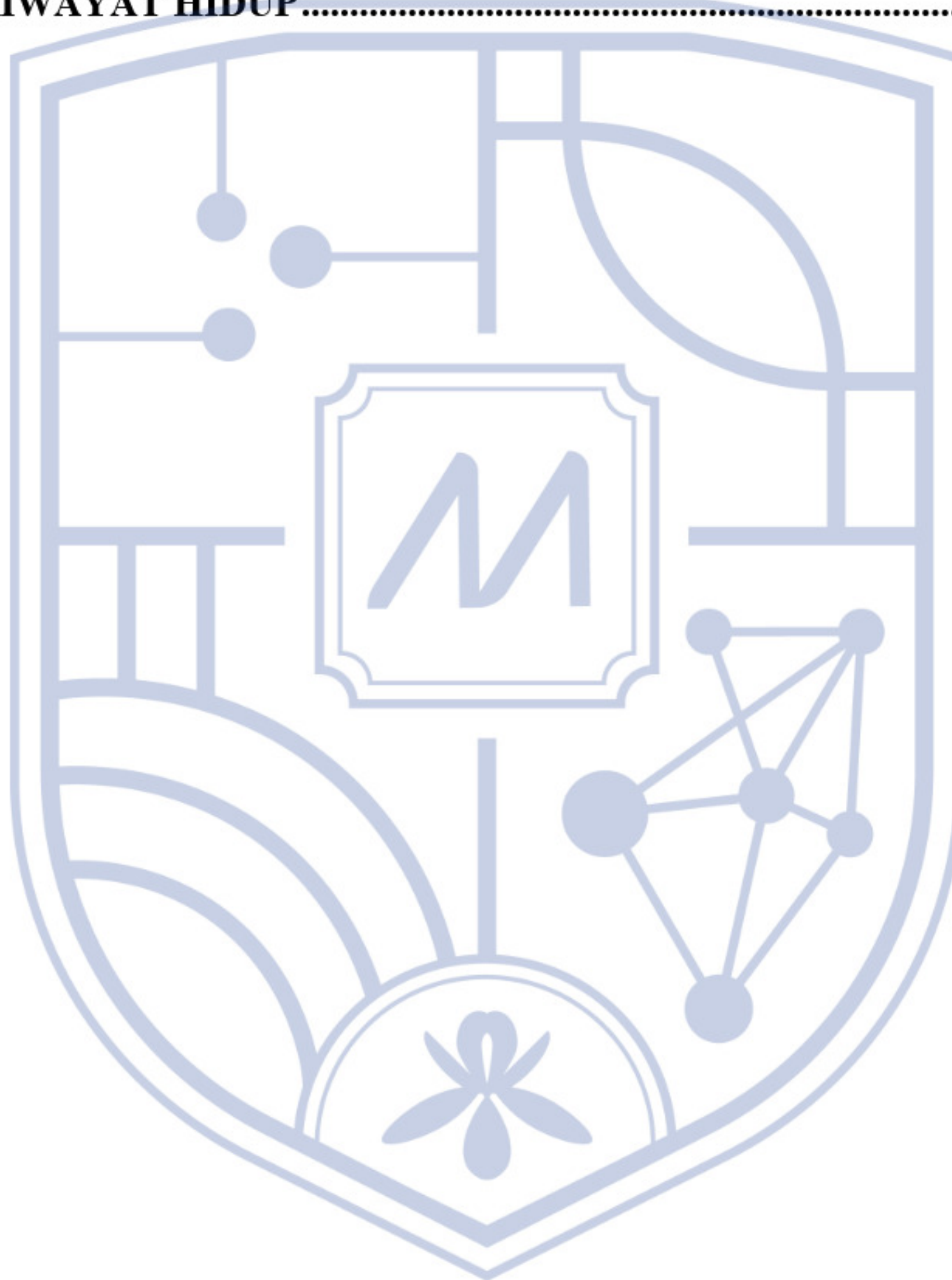
Penulis,
Calvin Wijaya
Calvin Winata



DAFTAR ISI

ABSTRAK.....	i
KATA PENGANTAR	ii
DAFTAR ISI	iv
DAFTAR GAMBAR	vi
DAFTAR TABEL.....	vii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan	3
1.4 Manfaat	4
1.5 Ruang Lingkup.....	4
BAB II KAJIAN LITERATUR.....	5
2.1 Artificial Intelligence (AI)	5
2.1.1 Machine Learning.....	5
2.1.2 Text Mining.....	6
2.1.3 Natural Language Processing (NLP).....	7
2.2 Support Vector Machine (SVM).....	7
2.3 IndoBERT	14
2.4 Confusion Matrix	16
2.5 Media Sosial.....	18
2.6 Judi Online	19
BAB III TAHAPAN PELAKSANAAN	21
3.1 Kerangka Tahapan Pelaksanaan.....	21
3.2 Analisis.....	22
3.2.1 Analisis Proses	23
3.2.2 Analisis Kebutuhan Fungsional	43
3.2.3 Analisis Kebutuhan Non-Fungsional	52
3.3 Perancangan	53
BAB IV HASIL DAN PEMBAHASAN	56
4.1 Hasil	56
4.2 Pembahasan.....	58
4.2.1 Persiapan dan Prapemrosesan Dataset	58

4.2.2 Ekstraksi Fitur menggunakan IndoBERT	61
4.2.3 Pelatihan dan Hyperparameter Tuning SVM	62
4.2.4 Confusion Matrix	63
BAB V PENUTUP	67
5.1 Kesimpulan	67
5.2 Saran.....	67
DAFTAR PUSTAKA	69
DAFTAR RIWAYAT HIDUP.....	75



DAFTAR GAMBAR

Gambar 2.1 Support Vector Machine.....	8
Gambar 2.2 SVM Kernel Linear	9
Gambar 2.3 SVM Kernel Non-Linear	10
Gambar 2.4 Arsitektur Model IndoBERT	15
Gambar 2.5 Confusion Matrix.....	17
Gambar 3.1 Flowchart (a) training (b) testing.....	23
Gambar 3.2 Flowchart Prapemrosesan Data	26
Gambar 3.3 Flowchart Training SVM.....	37
Gambar 3.4 Use Case Diagram Kebutuhan Fungsional	44
Gambar 3.5 Halaman Utama	54
Gambar 3.6 Halaman Hasil	54
Gambar 3.7 Statistik Deteksi	55
Gambar 4.1 Tampilan Halaman Beranda	56
Gambar 4.2 Hasil Penyaringan Komentar Youtube	57
Gambar 4.3 Tampilan Statistik Komentar Youtube	58
Gambar 4.4 Dataset Awal.....	59
Gambar 4.5 Prapemrosesan Dataset	60
Gambar 4.6 Distribusi Label dan Penghapusan data kosong dan duplikasi	60
Gambar 4.7 Pembagian Dataset Training dan Testing	61
Gambar 4.8 Hasil Tokenisasi Menggunakan IndoBERT	62
Gambar 4.9 Hasil Ekstraksi Fitur CLS Embedding Menggunakan IndoBERT	62
Gambar 4.10 Hasil Pelatihan dan Hyperparameter Tuning SVM	63
Gambar 4.11 Visualisasi Confusion Matrix Dataset Testing	64

DAFTAR TABEL

Tabel 2.1 Perbandingan SVM dengan Metode Lainnya.....	13
Tabel 3.1 Data Mentah	24
Tabel 3.2 Data yang telah dilabeli	25
Tabel 3.3 Tahapan Normalisasi Unicode	27
Tabel 3.4 Tahapan Lowercasing.....	27
Tabel 3.5 Tahapan Normalisasi angka	28
Tabel 3.6 Tahapan Remove URL, Emoji, Special Character, Multiple space	29
Tabel 3.7 Tahapan Normalisasi Pemisahan Kata tidak wajar	29
Tabel 3.8 Tahapan Normalisasi Huruf Berulang.....	30
Tabel 3.9 Hasil Prapemrosesan	31
Tabel 3.10 Proses Tokenisasi	31
Tabel 3.11 Konversi Token ke ID	32
Tabel 3.12 Tahapan Padding	32
Tabel 3.13 Tahapan Word Embedding.....	33
Tabel 3.14 Proses Transformer.....	34
Tabel 3.15 Tahapan Hasil Ekstraksi Fitur	35
Tabel 3.16 Data yang digunakan untuk Training	36
Tabel 3.17 Hasil Klasifikasi Data.....	42
Tabel 3.18 Data tabel untuk evaluasi model	42
Tabel 3.19 Hasil Confusion Matrix	43
Tabel 3.20 Deskripsi Aktor	44
Tabel 3.21 Use Case Input Link YouTube Diagram	44
Tabel 3.22 Use Case Autentikasi Youtube API	45
Tabel 3.23 Use Case Proses Deteksi Komentar	46
Tabel 3.24 Use Case Menampilkan Komentar Judi	48
Tabel 3.25 Use Case Melihat Statistik Deteksi	49
Tabel 3.26 Use Case Export Hasil Deteksi	50
Tabel 3.27 Use Case Hapus Komentar	50
Tabel 3.28 Use Case Biarkan Komentar	51
Tabel 3.29 Analisis Kebutuhan Non-Fungsional	53
Tabel 4.1 Komponen Confusion Matrix per Kelas	64