

**IMPLEMENTASI APLIKASI DETEKSI KOMENTAR SPAM IKLAN
JUDI ONLINE PADA YOUTUBE BERBAHASA INDONESIA
MENGUNAKAN INDOBERT**

SKRIPSI

Oleh:

**ROHID ZAIDAN NASUTION
NIM. 221111415**



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

**APPLICATION IMPLEMENTATION FOR DETECTING
INDONESIAN-LANGUAGE ONLINE GAMBLING SPAM
COMMENTS ON YOUTUBE WITH INDOBERT**

FINAL RESEARCH

By:

**ROHID ZAIDAN NASUTION
NIM. 221111415**



**MAJOR OF S-1 INFORMATICS ENGINEERING
FACULTY OF INFORMATICS
UNIVERSITAS MIKROSKIL
MEDAN
2026**

LEMBARAN PENGESAHAN

**IMPLEMENTASI APLIKASI DETEKSI KOMENTAR SPAM IKLAN
JUDI ONLINE PADA YOUTUBE BERBAHASA INDONESIA
MENGUNAKAN INDOBERT**

SKRIPSI

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

Oleh:

ROHID ZAIDAN NASUTION
NIM. 221111415

Disetujui Oleh:

Dosen Pembimbing I,



Syanti Irviantina, S.Kom., M.Kom.

Dosen Pembimbing II,



Erlina Halim, S.Kom., M.Kom.

Medan, 13 Ferbuari 2026

Diketahui dan Disahkan Oleh:

Ketua Program Studi
Teknik Informatika



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

HALAMAN PERNYATAAN

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

Nama : Rohid Zaidan Nasution

NIM : 221111415

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

Judul Tugas Akhir : Implementasi Aplikasi Deteksi Komentar Spam Iklan Judi Online Pada Youtube Berbahasa Indonesia Menggunakan Indobert

Tempat Penelitian : Universitas Mikroskil

Alamat Tempat Penelitian : Jl. M.H Thamrin No.140, Pusat Ps., Kec. Medan Kota, Kota Medan, Sumatra Utara 20212

No. Telp. Tempat Penelitian : (061) 4573767

Sehubungan dengan Tugas Akhir tersebut, dengan ini saya menyatakan dengan sebenarnya bahwa penelitian dan penulisan Tugas Akhir 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 Tugas Akhir 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 Tugas Akhir 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 Tugas Akhir 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, 29 Januari 2026

Saya yang membuat pernyataan,



Rohid Zaidan Nasution

IMPLEMENTASI APLIKASI DETEKSI KOMENTAR SPAM IKLAN JUDI ONLINE PADA YOUTUBE BERBAHASA INDONESIA MENGGUNAKAN INDOBERT

Abstrak

Platform YouTube saat ini menghadapi tantangan serius terkait maraknya komentar spam yang mempromosikan situs judi online. Komentar spam ini sering kali menggunakan teknik penyamaran dan bahasa gaul (slang) yang sulit dideteksi oleh filter otomatis berbasis kata kunci. Penelitian ini bertujuan untuk mengembangkan sistem deteksi komentar spam iklan judi online otomatis pada Youtube menggunakan model Deep Learning IndoBERT serta mengimplementasikannya dalam bentuk aplikasi Chrome Extension. Model IndoBERT dilatih menggunakan teknik fine-tuning dengan learning rate $2e-5$, batch size 16, dan epoch 3 pada dataset komentar berbahasa Indonesia. Hasil pengujian menunjukkan bahwa model mampu menangani ketidakseimbangan kelas (imbalanced dataset) dengan sangat baik, mencapai nilai Akurasi 95,05% dan F1-Score 90% pada kelas spam. Tingginya nilai Presisi sebesar 91% menunjukkan bahwa model memiliki tingkat kesalahan False Positive yang rendah, sehingga aman digunakan pengguna. Sistem berhasil diimplementasikan menggunakan arsitektur Client-Server, di mana ekstensi peramban berfungsi sebagai antarmuka pengguna yang melakukan penyensoran (blurring), filtrasi (hide), serta pemberian label peringatan (warning) komentar spam secara real-time berdasarkan prediksi dari API server. Penelitian ini menyimpulkan bahwa integrasi model IndoBERT ke dalam Chrome Extension merupakan solusi teknis yang layak dan efektif untuk membantu pengguna menciptakan lingkungan digital yang bersih dari gangguan promosi judi online.

Kata kunci: Deteksi spam, Judi Online, IndoBERT, Chrome Extension, Deep Learning.

Abstract

YouTube is currently facing significant challenges regarding the proliferation of spam comments promoting online gambling sites. These spam comments often employ obfuscation techniques and slang, making them difficult to detect using keyword-based automatic filters. This study aims to develop an automatic online gambling spam detection system on YouTube using the IndoBERT Deep Learning model and implement it as a Chrome Extension application. The IndoBERT model was trained using fine-tuning techniques with a learning rate of $2e-5$, a batch size of 16, and 3 epochs on an Indonesian comment dataset. The test results demonstrate that the model effectively handles class imbalance, achieving an Accuracy of 95.05% and an F1-Score of 90% for the spam class. A high Precision score of 91% indicates a low False Positive rate, ensuring the system is safe for user use. The system was successfully implemented using a Client-Server architecture, where the browser extension serves as a user interface that performs blurring, hiding, and warning labeling on spam comments in real-time based on predictions from the API server. This study concludes that the integration of the IndoBERT model into a Chrome Extension is a feasible and effective technical solution to assist users in creating a digital environment free from online gambling promotions.

Keywords: Spam Detection, Online Gambling, IndoBERT, Chrome Extension, Deep Learning.

KATA PENGANTAR

Ucapan syukur yang tulus saya panjatkan kepada Tuhan yang Maha Esa atas segala limpahan Rahmat, petunjuk, dan karunia-Nya yang telah menyertai perjalanan saya dalam menyelesaikan skripsi ini. Dengan penuh rasa Syukur, saya mengakui bahwa tanpa pertolongan-Nya, saya tidak akan mampu mengatasi berbagai tantangan dan hambatan yang saya hadapi selama proses penyusunan skripsi ini.

Penyusunan skripsi ini merupakan bagian integral dari tahapan akhir dalam menyelesaikan pendidikan saya di Program Studi Teknik Informatika, Fakultas Informatika Universitas Mikroskil Medan.

Saya ingin menyampaikan ucapan terima kasih yang sebesar besarnya kepada semua pihak yang telah memberikan dukungan, bimbingan, dan motivasi dalam proses penyusunan skripsi ini.

1. Ibu Syanti Irviantina, S.Kom., M.Kom., selaku Dosen Pembimbing I.
2. Ibu Erlina Halim, S.Kom., M.Kom., selaku Dosen Pembimbing II.
3. Bapak Hardy, S.kom., M.Sc., Ph.D, selaku Rektor Universitas Mikroskil Medan.
4. Bapak Sunaryo Winardi, S.kom., M.T., selaku Dekan Fakultas Informatika Universitas Mikroskil Medan.
5. Bapak Carles Juliandy, S.kom., M.kom., selaku Ketua Program Studi S-1 Teknik Informatika, Fakultas Informatika Universitas Mikroskil Medan.
6. Orang tua, keluarga dan teman-teman

Saya menyadari bahwa skripsi ini jauh dari kesempurnaan, oleh karena itu, kritik dan saran yang membangun dari pembaca sangat saya harapkan untuk perbaikan di masa mendatang. Demikianlah ucapan terima kasih ini saya sampaikan, semoga kebaikan dan berkah senantiasa menyertai kita semua.

Medan, 29 Januari 2026

Penulis,

Rohid Zaidan Nasution

DAFTAR ISI

| | |
|--|------------|
| Abstrak | i |
| KATA PENGANTAR | ii |
| DAFTAR ISI | iii |
| DAFTAR GAMBAR | v |
| DAFTAR TABEL | vi |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 2 |
| 1.3 Tujuan | 2 |
| 1.4 Manfaat | 3 |
| 1.5 Ruang Lingkup | 3 |
| BAB II KAJIAN LITERATUR | 5 |
| 2.1 Deteksi Komentar Spam | 5 |
| 2.2 <i>Natural Language Processing</i> (NLP) | 6 |
| 2.3 Model <i>Machine Learning</i> untuk Deteksi Spam | 7 |
| 2.4 <i>Transformer</i> dan BERT | 9 |
| 2.5 IndoBERT | 10 |
| 2.5.1 Dasar Arsitektur IndoBERT | 10 |
| 2.5.2 Proses Embedding | 11 |
| 2.5.3 Proses Pretraining | 11 |
| 2.5.4 Proses Fine-Tuning | 12 |
| 2.5.5 Keunggulan IndoBERT | 13 |
| 2.5.6 Keterbatasan IndoBERT | 13 |
| 2.6 <i>Confusion Matrix</i> | 14 |
| 2.7 <i>Chrome Extension</i> | 15 |

| | |
|---|-----------|
| BAB III TAHAPAN PELAKSANAAN | 17 |
| 3.1 Kerangka Tahapan Pelaksanaan | 17 |
| 3.2 Analisis Proses..... | 18 |
| 3.3 Pembangunan Model IndoBERT | 22 |
| 3.4 Evaluasi Model..... | 23 |
| 3.5 Perancangan Chrome Extension..... | 25 |
| BAB IV HASIL DAN PEMBAHASAN..... | 27 |
| 4.1 Hasil..... | 27 |
| 4.1.1 Lingkungan Implementasi | 27 |
| 4.1.2 Pengumpulan Data..... | 27 |
| 4.1.3 Preprocessing Data | 30 |
| 4.1.4 Pelabelan Data | 34 |
| 4.1.5 Hasil Pengujian Skenario Pembagian Data | 35 |
| 4.1.6 Hasil Eksperimen Hyperparameter Tuning | 36 |
| 4.1.7 Implementasi dan Pelatihan Model IndoBERT..... | 38 |
| 4.1.8 Evaluasi Model..... | 40 |
| 4.1.9 Analisis Kesalahan | 42 |
| 4.1.10 Visualisasi Fitur (Word Cloud) | 44 |
| 4.1.11 Implementasi Chrome Extension | 44 |
| 4.1.12 Pengujian Fungsionalitas Chrome Extension dan Integrasi API..... | 46 |
| 4.2 Pembahasan | 48 |
| 4.2.1 Kinerja Deteksi Spam pada Data Tidak Seimbang | 48 |
| 4.2.2 Implementasi Sistem Deteksi Berbasis Browser | 49 |
| BAB V PENUTUP | 50 |
| DAFTAR PUSTAKA | 53 |
| DAFTAR RIWAYAT HIDUP | 58 |

DAFTAR GAMBAR

| | |
|---|----|
| Gambar 3.1 Flowchart Analisis Proses | 18 |
| Gambar 3.2 Contoh <i>Confusion Matrix</i> | 24 |
| Gambar 3.3 Contoh <i>Classification Report</i> Model | 24 |
| Gambar 3.4 Perancangan tampilan <i>Chrome Extension</i> | 26 |
| Gambar 3.5 Implementasi aplikasi deteksi komentar judi online pada youtube | 26 |
| Gambar 4.1 Grafik Distribusi Kelas pada Dataset Final | 35 |
| Gambar 4.2 Grafik Peningkatan Akurasi dan F1-Score | 39 |
| Gambar 4.3 <i>Confusion Matrix</i> Model IndoBERT pada Data Uji..... | 40 |
| Gambar 4.4 <i>Word Cloud</i> Dataset Komentar YouTube..... | 44 |
| Gambar 4.5 Tampilan Menu Pengaturan (<i>Popup</i>) pada <i>Chrome Extension</i> | 45 |
| Gambar 4.6 Implementasi Deteksi Real-time dengan Mode <i>Blur</i> | 46 |
| Gambar 4.7 Tampilan Hasil Pengujian Mode <i>Blur</i> | 46 |
| Gambar 4.8 Tampilan Hasil Pengujian Interaksi <i>Hover</i> | 47 |
| Gambar 4.9 Tampilan Hasil Pengujian Mode <i>Hide</i> | 47 |
| Gambar 4.10 Hasil Pengujian Integrasi API menggunakan Postman | 48 |

DAFTAR TABEL

| | |
|---|----|
| Tabel 2.1 <i>Confusion Matrix</i> | 14 |
| Tabel 3.1 Contoh hasil dari tahap <i>Text Cleaning</i> | 20 |
| Tabel 3.2 Contoh hasil dari tahap <i>Case Folding</i> | 20 |
| Tabel 3.3 Contoh hasil dari tahap <i>Normalization</i> | 21 |
| Tabel 4.1 Contoh Data Mentah Hasil <i>Crawling</i> | 28 |
| Tabel 4.2 Hasil Tahap <i>Text Cleaning</i> | 30 |
| Tabel 4.3 Hasil Tahap <i>Case Folding</i> | 31 |
| Tabel 4.4 Hasil Tahap <i>Normalization</i> | 33 |
| Tabel 4.5 Perbandingan Hasil Skenario Pembagian Data | 36 |
| Tabel 4.6 Hasil Eksperimen <i>Hyperparameter Tuning</i> | 37 |
| Tabel 4.7 <i>Classification Report</i> Model IndoBERT..... | 41 |
| Tabel 4.8 Sampel Analisis <i>False Positive</i> | 42 |
| Table 4.9 Sampel Analisis <i>False Negative</i> | 43 |

