

**PERBANDINGAN KLASIFIKASI SPAM SMS DAN SPAM
TELEGRAM MENGGUNAKAN ALGORITMA NAÏVE BAYES
BERBASIS MACHINE LEARNING**

TUGAS AKHIR



Disusun Oleh :

Ahmad Khobirul Jannah

8020210200

Untuk Persyaratan Penelitian dan Penulisan Tugas Akhir

Sebagai Akhir Proses Studi Strata 1

PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS DINAMIKA BANGSA

JAMBI

2025

LEMBAR PENGESAHAN PEMBIMBING

Judul Tugas Akhir : PERBANDINGAN KLASIFIKASI *SPAM SMS* DAN *SPAM TELEGRAM* MENGGUNAKAN ALGORITMA *NAÏVE BAYES* BERBASIS *MACHINE LEARNING*

Nama : Ahmad Khobirul Jannah

NIM : 8020210200

Tanggal Sidang :

Program Studi : Teknik Informatika

Mengesahkan

Pembimbing I

Pembimbing II

(Eko Arip Winanto, S.Kom, M.Phil)

(Pareza Alam Jusia, S.Kom, M.Kom)

NIK : YDB.22.94.190

NIK : YDB.11.90.085

Ka. Prodi Teknik Informatika

(Marrylinteri Istoningtyas, ST, M.Kom)

NIK : YDB.16.82.108

LEMBAR PERNYATAAN KEASLIAN TUGAS AKHIR

Saya yang bertanda tangan dibawah ini :

NAMA : AHMAD KHOBIRUL JANNAH

NIM : 8020210200

**JUDUL TUGAS AKHIR : PERBANDINGAN KLASIFIKASI *SPAM SMS*
DAN *SPAM TELEGRAM* MENGGUNAKAN
ALGORITMA *NAÏVE BAYES* BERBASIS
*MACHINE LEARNING***

Menyatakan dengan sesungguhnya bahwa Tugas Akhir ini merupakan hasil penelitian, pemikiran dan pemaparan asli saya sendiri. Saya tidak mencantumkan tanpa pengakuan bahan - bahan yang telah dipublikasikan sebelumnya atau ditulis oleh orang lain, atau sebagai bahan yang pernah diajukan untuk gelar atau ijazah pada Universitas Dinamika Bangsa.

Apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik sesuai dengan peraturan yang berlaku di Universitas Dinamika Bangsa.

Demikian pernyataan ini saya buat.

Jambi, Januari 2025

Yang membuat pernyataan,

Ahmad Khobirul Jannah

ABSTRAK

Ahmad Khobirul Jannah, 8020210200

PERBANDINGAN KLASIFIKASI *SPAM SMS* DAN *SPAM TELEGRAM* MENGGUNAKAN ALGORITMA *NAÏVE BAYES* BERBASIS *MACHINE LEARNING*

Kata Kunci: *Spam SMS, Spam Telegram, Machine Learning, Naïve Bayes, Klasifikasi, Deteksi Spam*

(xi+72+Lampiran)

Dalam era digital saat ini, keamanan komunikasi menjadi tantangan yang penting untuk diperhatikan, terutama dengan meningkatnya penggunaan layanan pesan singkat (SMS). Salah satu ancaman utama adalah pesan spam yang dapat mengganggu pengguna dan berpotensi merugikan. Penelitian ini merekomendasikan penggunaan algoritma Naïve Bayes berbasis machine learning karena kemampuannya dalam mengklasifikasikan data teks secara efisien. Sebelum data diproses oleh Naïve Bayes, dilakukan preprocessing data, termasuk pembersihan teks dan ekstraksi fitur menggunakan TF-IDF, untuk meningkatkan efisiensi pemrosesan dan kinerja model. Struktur Naïve Bayes dirancang untuk mendeteksi pola probabilitas yang relevan dengan klasifikasi spam SMS. Hasil evaluasi menunjukkan bahwa metode ini mampu mengklasifikasikan pesan spam dengan akurasi sebesar 98% pada dataset SMS dan 92% pada dataset Telegram, yang menunjukkan efektivitasnya dalam mendeteksi pesan spam secara andal dan efisien.

Daftar Pustaka (2014 - 2024)

ABSTRACT

Ahmad Khobirul Jannah, 8020210200

COMPARISON OF SMS SPAM AND TELEGRAM SPAM CLASSIFICATION USING NAÏVE BAYES ALGORITHM BASED ON MACHINE LEARNING

Keywords: SMS Spam, Telegram Spam, Machine Learning, Naïve Bayes, Classification, Spam Detection

(xi+72+Attachment)

In the current digital era, communication security is an important challenge to pay attention to, especially with the increasing use of short message services (SMS). One of the main threats is spam messages which can annoy users and have the potential to cause harm. This research recommends using the machine learning-based Naïve Bayes algorithm because of its ability to classify text data efficiently. Before the data is processed by Naïve Bayes, data preprocessing is carried out, including text cleaning and feature extraction using TF-IDF, to improve processing efficiency and model performance. The Naïve Bayes structure is designed to detect probability patterns relevant to SMS spam classification. The evaluation results show that this method is able to classify spam messages with an accuracy of 98% on the SMS dataset and 92% on the Telegram dataset, which shows its effectiveness in detecting spam messages reliably and efficiently.

Bibliography (2014 -2024)

KATA PENGANTAR

Puji syukur kehadiran Allah SWT, atas limpahan rahmat dan karunia-Nya, sehingga tugas akhir yang berjudul “PERBANDINGAN KLASIFIKASI *SPAM SMS* DAN *SPAM TELEGRAM* MENGGUNAKAN ALGORITMA *NAÏVE BAYES* BERBASIS *MACHINE LEARNING*” ini dapat diselesaikan dengan baik sebagai persyaratan memperoleh gelar sarjana komputer pada Universitas Dinamika Bangsa.

Penulis juga mengucapkan terimakasih yang sebesar-sebesarnya kepada pihak yang telah membantu dalam penulisan laporan ini, antara lain :

1. Bapak Dr. Ir. Eddy Anthony, SH, MM selaku Ketua Yayasan Dinamika Bangsa Jambi.
2. Bapak Dr. Ir. Herry Mulyono, MM selaku Pembina Yayasan Dimanika Bangsa Jambi.
3. Bapak Prof. Setiawan Assegaff, ST, MMSI, Ph.D selaku Rektor Universitas Dinamika Bangsa Jambi.
4. Bapak Beny, S.Kom, M.Sc selaku Dekan Fakultas Ilmu Komputer Universitas Dinamika Bangsa Jambi.
5. Ibu Marrylinteri Istoningtyas, ST, M.Kom selaku Ketua Program Studi Teknik Informatika Universitas Dinamika Bangsa Jambi.
6. Bapak Eko Arip Winanto, S.Kom, M.Phil selaku Pembimbing 1 yang telah memberikan bantuan dan pengarahan pada penelitian ini.
7. Bapak Pareza Alam Jusia, S.Kom, M.Kom selaku Pembimbing 2 yang telah memberikan bantuan dan pengarahan pada penelitian ini.
8. Kepada ibu dan ayah serta keluarga yang telah membantu, mendukung dan memberikan do'a agar penulis dapat menyelesaikan Tugas Akhir ini dengan baik dan tepat waktu.
9. Bapak dan ibu Dosen Universitas Dinamika Bangsa Jambi yang telah memberikan bekal ilmu pengetahuan dilingkungan kampus.
10. Kepada teman-teman dan sahabat yang telah membantu dalam suka maupun duka dalam proses pembuatan laporan Tugas Akhir ini, serta seluruh pihak yang tidak disebutkan satu persatu yang telah membantu dalam menyelesaikan Tugas Akhir ini.

Semoga dapat dijadikan amal jariyah, yang akan dibalas oleh Allah SWT. Akhir kata penulis mengharapkan semoga laporan Tugas Akhir ini dapat bermanfaat bagi kita semua dan pihak-pihak yang membutuhkan.

Jambi, Januari 2025

Penulis

DAFTAR ISI

	Halaman
LEMBAR PENGESAHAN PEMBIMBING	ii
LEMBAR PERNYATAAN KEASLIAN TUGAS AKHIR	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR	vi
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	xi
BAB I PENDAHULUAN	Error! Bookmark not defined.
1.1 LATAR BELAKANG MASALAH	Error! Bookmark not defined.
1.2 RUMUSAN MASALAH	Error! Bookmark not defined.
1.3 BATASAN MASALAH	Error! Bookmark not defined.
1.4 TUJUAN DAN MANFAAT PENELITIAN	Error! Bookmark not defined.
1.4.1 TUJUAN PENELITIAN	Error! Bookmark not defined.
1.4.2 MANFAAT PENELITIAN	Error! Bookmark not defined.
1.5 SISTEMATIKA PENULISAN	Error! Bookmark not defined.
BAB II LANDASAN TEORI	Error! Bookmark not defined.
2.1 <i>SPAM SMS</i>	Error! Bookmark not defined.
2.2 <i>SPAM TELEGRAM</i>	Error! Bookmark not defined.
2.3 <i>DATA MINING</i>	Error! Bookmark not defined.
2.4 CRISP-DM	Error! Bookmark not defined.
2.5 <i>MACHINE LEARNING</i>	Error! Bookmark not defined.
2.5.1 <i>Toxonomy Machine Learning</i>	Error! Bookmark not defined.
2.5.2 Tugas Utama <i>Machine Learning</i>	Error! Bookmark not defined.
2.6 KLASIFIKASI	Error! Bookmark not defined.
2.7 <i>NAÏVE BAYES</i>	Error! Bookmark not defined.
2.8 <i>GOOGLE COLABORATORY</i>	Error! Bookmark not defined.
2.9 PENELITIAN SEJENIS	Error! Bookmark not defined.
BAB III METODOLOGI PENELITIAN	Error! Bookmark not defined.
3.1 KERANGKA PENELITIAN	Error! Bookmark not defined.

3.2 ALAT BANTU PENELITIAN	Error! Bookmark not defined.
3.3 ALUR EKSPERIMEN	Error! Bookmark not defined.
BAB IV HASIL DAN PEMBAHASAN	Error! Bookmark not defined.
4.1 DATA EKSPERIMEN.....	Error! Bookmark not defined.
4.1.1 Profil Dataset	Error! Bookmark not defined.
4.1.2 <i>Preprocessing Data</i>	Error! Bookmark not defined.
4.2 <i>SPLIT DATA</i>	Error! Bookmark not defined.
4.3 Fitur TF-IDF (<i>Term Frequency - Inverse Document Frequency</i>).....	Error! Bookmark not defined.
Bookmark not defined.	
4.4 IMPLEMENTASI ALGORITMA <i>NAÏVE BAYES</i>	Error! Bookmark not defined.
defined.	
4.4.1 Pelatihan <i>Naïve Bayes</i> Pada <i>Spam SMS Dan Telegram</i>	Error! Bookmark not defined.
not defined.	
4.4.2 Evaluasi <i>Naïve Bayes</i> Pada <i>Spam SMS Dan Telegram</i>	Error! Bookmark not defined.
defined.	
4.4.3 Laporan Klasifikasi <i>Naïve Bayes</i> Pada <i>Spam SMS Dan Telegram</i>	Error! Bookmark not defined.
Bookmark not defined.	
4.4.4 <i>Confusion Matrix Naïve Bayes</i> Pada <i>Spam SMS Dan Telegram</i>	Error! Bookmark not defined.
Bookmark not defined.	
4.4.5 Kurva ROC/AUC <i>Naïve Bayes</i> Pada <i>Spam SMS Dan Telegram</i>	Error! Bookmark not defined.
Bookmark not defined.	
4.4.6 Pengujian <i>5-Fold Cross Validation Naïve Bayes</i> Pada <i>Spam SMS Dan Telegram</i>	Error! Bookmark not defined.
Error! Bookmark not defined.	
4.5 VISUALISASI HASIL.....	Error! Bookmark not defined.
4.5.1 Grafik Perbandingan <i>Accuracy, Precision, Recall, F1-Score</i> Pada <i>Dataset Spam SMS Dan Telegram</i>	Error! Bookmark not defined.
Error! Bookmark not defined.	
4.5.2 Grafik Perbandingan ROC/AUC Pada <i>Dataset Spam SMS Dan Telegram</i> ...	Error! Bookmark not defined.
Error! Bookmark not defined.	
BAB V KESIMPULAN DAN SARAN.....	Error! Bookmark not defined.
5.1 KESIMPULAN	Error! Bookmark not defined.
5.2 SARAN	Error! Bookmark not defined.
DAFTAR PUSTAKA	Error! Bookmark not defined.
LAMPIRAN	Error! Bookmark not defined.

DAFTAR TABEL

Tabel 2. 1 Tabel Masalah Deteksi <i>Spam SMS</i>	Error! Bookmark not defined.
Tabel 2. 2 Tabel Penelitian Sejenis.....	Error! Bookmark not defined.
Tabel 3. 1 Dataset Yang Digunakan.....	Error! Bookmark not defined.
Tabel 3. 2 Tabel Fitur <i>Spam SMS</i>	Error! Bookmark not defined.
Tabel 3. 3 Tabel Fitur <i>Spam Telegram</i>	Error! Bookmark not defined.
Tabel 3. 4 Tabel Label <i>Spam SMS</i>	Error! Bookmark not defined.
Tabel 3. 5 Tabel Label <i>Spam Telegram</i>	Error! Bookmark not defined.
Tabel 3. 6 <i>Pseudocode NB</i>	Error! Bookmark not defined.
Tabel 4. 1 Profil <i>Dataset</i>	Error! Bookmark not defined.

DAFTAR GAMBAR

Gambar 2. 1 <i>Phase</i> CRISP-DM	Error! Bookmark not defined.
Gambar 3. 1 Kerangka Kerja Penelitian	Error! Bookmark not defined.
Gambar 3. 2 Alur Eksperimen	Error! Bookmark not defined.
Gambar 4. 1 Hasil Data Yang Sudah Di <i>Cleaning</i>	Error! Bookmark not defined.
Gambar 4. 2 Visualisasi <i>WordCloud</i> Yang Sudah Di <i>Cleaning</i>	Error! Bookmark not defined.
Gambar 4. 3 <i>Split</i> Data <i>Spam SMS</i> dan <i>Spam Telegram</i>	Error! Bookmark not defined.
Gambar 4. 4 Fitur TF-IDF	Error! Bookmark not defined.
Gambar 4. 5 Kode Pelatihan <i>Naïve Bayes</i> Pada <i>Spam SMS</i> Dan <i>Telegram</i>	Error! Bookmark not defined.
Gambar 4. 6 Evaluasi <i>Naïve Bayes</i> Pada <i>Spam SMS</i> Dan <i>Telegram</i> ...	Error! Bookmark not defined.
Gambar 4. 7 Hasil Laporan Klasifikasi <i>Naïve Bayes</i> Pada <i>Spam SMS</i> Dan <i>Telegram</i>	Error! Bookmark not defined.
Gambar 4. 8 Visualisasi <i>Confusion Matrix</i> Pada <i>Spam SMS</i>	Error! Bookmark not defined.
Gambar 4. 9 Visualisasi <i>Confusion Matrix</i> Pada <i>Spam Telegram</i>	Error! Bookmark not defined.
Gambar 4. 10 Visualisasi Kurva ROC/AUC Pada <i>Spam SMS</i>	Error! Bookmark not defined.
Gambar 4. 11 Visualisasi Kurva ROC/AUC Pada <i>Spam Telegram</i>	Error! Bookmark not defined.

Gambar 4. 12 Hasil *5-Fold Cross Validation* Pada *Dataset Spam SMS*.... **Error! Bookmark not defined.**

Gambar 4. 13 Hasil *5-Fold Cross Validation* Pada *Dataset Spam Telegram* **Error! Bookmark not defined.**

Gambar 4. 14 Visualisasi Perbandingan *Accuracy, Precision, Recall, F1-Score* **Error! Bookmark not defined.**

Gambar 4. 15 Visualisasi Perbandingan ROC/AUC Pada *Spam SMS Dan Telegram*. **Error! Bookmark not defined.**

DAFTAR LAMPIRAN

Lampiran A Daftar Riwayat Hidup.....	L1
Lampiran B Kartu Bimbingan.....	L2
Lampiran C Dataset.....	L4
Lampiran D Hasil Plagiasi	L8
Lampiran E Listing Program	L10

