

**IMPLEMENTASI ALGORITMA *NAÏVE BAYES* DAN TF-IDF
DALAM MENDETEKSI PLAGIARISME ARTIKEL JURNAL**



TUGAS AKHIR

**Disusun untuk Memenuhi Syarat Menyelesaikan Pendidikan
Sarjana Terapan pada Jurusan Teknik Elektro
Program Studi Teknik Telekomunikasi
Politeknik Negeri Sriwijaya**

OLEH:

LADYSA AZZAHRAH

0620 4035 2253

POLITEKNIK NEGERI SRIWIJAYA

PALEMBANG

2024

TUGAS AKHIR

IMPLEMENTASI ALGORITMA *NAÏVE BAYES* DAN TF-IDF DALAM MENDETEKSI PLAGIARISME ARTIKEL JURNAL



**Disusun untuk Memenuhi Syarat Menyelesaikan Pendidikan
Sarjana Terapan pada Jurusan Teknik Elektro
Program Studi Teknik Telekomunikasi
Politeknik Negeri Sriwijaya**

Oleh:

Nama : Ladysa Azzahrah (0620 4035 2253)
Dosen Pembimbing I : Lindawati, S.T., M.T.I.
Dosen Pembimbing II : Sholihin, S.T., M.T.

**POLITEKNIK NEGERI SRIWIJAYA
PALEMBANG
2024**

**IMPLEMENTASI ALGORITMA NAIVE BAYES DAN TF-IDF
DALAM MENDETEKSI PLAGIARISME ARTIKEL JURNAL**



TUGAS AKHIR

**Disusun untuk Memenuhi Syarat Menyelesaikan Pendidikan
Sarjana Terapan pada Jurusan Teknik Elektro
Program Studi Teknik Telekomunikasi
Politeknik Negeri Sriwijaya**

OLEH:

LADYSA AZZAHRAH

0620 4035 2253

Palembang, September 2024

Pembimbing I,

Lindawati, S.T., M.T.I.
NIP. 197105282006042001

Pembimbing II,

Sholihin, S.T., M.T.
NIP. 197404252001121001

Mengetahui,

Ketua Jurusan Teknik Elektro

Ir. Iskandar Lutfi, M.T.
NIP. 195301291991031002

**Koordinator Program Studi
Sarjana Terapan Teknik Telekomunikasi**

Lindawati, S.T., M.T.I.
NIP. 197105282006042001

SURAT PERNYATAAN

Saya yang bertanda tangan dibawah ini menyatakan:

Nama : Ladysa Azzahrah
Jenis Kelamin : Perempuan
Tempat, Tanggal Lahir : Lahat, 12 Desember 2002
Alamat : Perum. Griya Palem Kencana No.26
NPM : 062040352253
Program Studi : Sarjana Terapan Teknik Telekomunikasi
Jurusan : Teknik Elektro
Judul Skripsi/Laporan : Implementasi Algoritma *Naïve Bayes* dan TF-IDF
Tugas Akhir : Dalam Mendeteksi Plagiarisme Artikel Jurnal

Menyatakan dengan sesungguhnya bahwa :

1. Skripsi/Laporan Tugas Akhir ini adalah hasil karya saya sendiri serta bebas dari tindakan plagiasi dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar.
2. Dapat menyelesaikan segala urusan terkait pengumpulan revisi Skripsi/Laporan Tugas Akhir yang sudah disetujui oleh dewan penguji paling lama 1 bulan setelah ujian Skripsi/Laporan Tugas Akhir.
3. Dapat menyelesaikan segala urusan peminjaman/penggantian alat/buku dan lainnya paling lama 1 bulan setelah ujian Skripsi/Laporan Tugas Akhir.

Apabila dikemudian hari diketahui ada pernyataan yang terbukti tidak benar dan tidak dapat dipenuhi, maka saya siap bertanggung jawab dan menerima sanksi tidak diikutsertakan dalam prosesi wisuda serta dimpasukan dalam daftar hitam oleh jurusan Teknik Elektro sehingga berdampak tertundanya pengambilan Ijazah & Transkrip (ASLI & COPY). Demikian surat pernyataan ini dibuat dengan sebenar-benarnya dan dalam keadaan sadar tanpa paksaan.

Palembang, September 2024

Yang Menyatakan



(Ladysa Azzahrah)

Mengetahui,

Pembimbing I : Lindawati, S.T., M.T.I.

Pembimbing II : Sholihin, S.T., M.T.

.....
.....

MOTTO DAN PERSEMBAHAN

“Karena sesungguhnya sesudah kesulitan itu ada kemudahan.” Q.S. Al-Insyirah: 6

“Kesulitan yang besar bisa diatasi dengan sabar dan doa, dan Allah akan memberikan kemudahan dari arah yang tidak disangka-sangka.” - Abu Bakr As-Siddiq

“Kunci untuk menghadapi kesulitan adalah D.U.I.T: Doa yang tak pernah putus, Usaha yang tak pernah henti, Ikhtiar yang tak kenal lelah, dan Tawakkal yang penuh keikhlasan.” - Ladysa Azzahrah

Tugas Akhir ini kupersembahkan kepada:

- Allah SWT yang telah memberikan kekuatan, kesehatan, dan semua yang dibutuhkan. Tiada henti untuk bersyukur kepadaMu ya Allah.
- Ayah, Bunda, Adek Tia, Adek Aulia, dan Adek Medina terima kasih yang sebesar-besarnya atas kasih sayang, dukungan, serta doa yang tiada hentinya.
- Ibu Lindawati dan Pak Sholihin terima kasih atas bimbingan, ilmu, dan nasihat yang sangat berharga selama proses penyusunan Tugas Akhir ini.
- Teman-teman terdekat yang selalu ada untuk memberikan dukungan, motivasi, dan bantuan selama masa studi.
- Teman seperjuangan angkatan tahun 2020.
- Semua pihak yang terlibat, baik secara langsung maupun tidak langsung, dalam membantu dan mendukung penyusunan Tugas Akhir ini.
- Almamater “Politeknik Negeri Sriwijaya”.

ABSTRAK

IMPLEMENTASI ALGORITMA *NAÏVE BAYES* DAN TF-IDF DALAM MENDETEKSI PLAGIARISME ARTIKEL JURNAL

(2024:xv + 70 halaman + 21 gambar + 10 tabel + 4 persamaan + 10 lampiran)

LADYSA AZZAHRAH

062040352253

JURUSAN TEKNIK ELEKTRO

PROGRAM STUDI SARJANA TERAPAN TEKNIK TELEKOMUNIKASI

POLITEKNIK NEGERI SRIWIJAYA

Abstrak - Plagiarisme merupakan ancaman serius bagi integritas akademik di era kemudahan akses informasi. Penelitian ini bertujuan untuk mengimplementasikan kombinasi algoritma *Naïve Bayes* dan TF-IDF dalam mendeteksi plagiarisme pada abstrak artikel jurnal berbahasa Indonesia tentang *Internet of Things* (IoT). Studi ini menggunakan dataset yang terdiri dari seratus abstrak artikel jurnal online teratas dari *software* Mendeley. Metode yang digunakan menggabungkan kelebihan *Naïve Bayes* dalam menghitung probabilitas teks sebagai plagiat berdasarkan kemunculan kata-kata, dengan TF-IDF yang menghitung bobot kata berdasarkan frekuensi kemunculan dan kejarangannya dalam koleksi dokumen. *Cosine similarity* ditambahkan untuk mengukur kemiripan antar dokumen. Implementasi dilakukan menggunakan bahasa pemrograman Python dan dipresentasikan melalui sebuah *website* untuk memudahkan pengguna. Penelitian ini bertujuan untuk menjelaskan secara detail implementasi algoritma tersebut dan mengukur efektivitasnya dalam mendeteksi plagiarisme. Pada penelitian ini evaluasi model atau hasil penelitian menggunakan beberapa macam uji skenario. Hasil menunjukkan bahwa model algoritma pada penelitian ini efektif untuk mendeteksi plagiarisme.

Kata Kunci: Plagiarisme, *Naïve Bayes*, TF-IDF, *Cosine Similarity*, Deteksi Plagiarisme

ABSTRACT

IMPLEMENTATION OF *NAÏVE BAYES* AND TF-IDF ALGORITHMS IN DETECTING PLAGIARISM OF JOURNAL ARTICLES

(2024:xv + 70 pages + 21 images + 10 tables + 4 equations + 10 attachments)

LADYSA AZZAHRAH

062040352253

DEPARTMENT OF ELECTRICAL ENGINEERING

BACHELOR OF APPLIED SCIENCE IN TELECOMMUNICATIONS

ENGINEERING STUDY PROGRAM

SRIWIJAYA STATE POLYTECHNIC

Abstract - Plagiarism is a serious threat to academic integrity in the era of easy access to information. This study aims to implement a combination of *Naïve Bayes* and TF-IDF algorithms in detecting plagiarism in abstract Indonesia-language journal articles about the Internet of Things (IoT). The study used a dataset consisting of a hundred abstracts of top online journal articles from Mendeley software. The method used combines *Naïve Bayes*' advantages in calculating the probability of text plagiarism based on the occurrence of words, with TF-IDF which calculates the weight of words based on the frequency of occurrence and their rarity in the collection of documents. Cosine similarity is added to measure similarity between documents. The implementation is carried out using the Python programming language and presented through a website to make it easier for users. This study aims to explain in detail the implementation of the algorithm and measure its effectiveness in detecting plagiarism. In this study, the evaluation of the model or research results uses several types of scenario tests. The results show that the algorithm model in this study is effective in detecting plagiarism.

Keywords: Plagiarism, *Naïve Bayes*, TF-IDF, Cosine Similarity, Plagiarism Detection

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Allah SWT atas berkah rahmat dan ridhonya semua urusan menjadi lancar. Tidak lupa juga penulis bershalawat dan salam kepada Nabi Muhammad SAW sebagai suri tauladan umat yang selalu semangat dalam menebar kebaikan. Dari keteladanan tersebut membuat penulis juga ikut bersemangat sehingga dapat menyelesaikan tugas akhir dengan judul **“Implementasi Algoritma *Naïve Bayes* dan TF-IDF dalam Mendeteksi Plagiarisme Artikel Jurnal”** yang disusun untuk memenuhi syarat menyelesaikan tugas akhir Pendidikan Sarjana Terapan pada Jurusan Teknik Elektro Program Studi Teknik Telekomunikasi Politeknik Negeri Sriwijaya Palembang.

Dengan terselesaikannya tugas akhir ini penulis mengucapkan terima kasih yang sebesar-besarnya kepada **Ibu Lindawati, S.T., M.T.I.** dan **Bapak Sholihin, S.T., M.T.** selaku dosen pembimbing yang telah memberikan nasihat, saran, dan waktunya kepada penulis dalam penyusunan dan pengerjaan tugas Akhir ini. Adapun penulis juga mengucapkan terimakasih kepada:

1. Orang tua dan saudara-saudariku yang selalu memberikan do’a, motivasi, dan dukungan moral maupun material.
2. Bapak Ir. Irawan Rusnadi, M.T. selaku Direktur Politeknik Negeri Sriwijaya.
3. Bapak Ir. Iskandar Lutfi, M.T. selaku Ketua Jurusan Teknik Elektro Politeknik Negeri Sriwijaya.
4. Ibu Lindawati, S.T., M.T.I. selaku Ketua Program Studi Sarjana Terapan Teknik Telekomunikasi Politeknik Negeri Sriwijaya.
5. Bapak dan Ibu Dosen Program Studi Teknik Telekomunikasi.
6. Teman-teman seperjuangan semester 8 Politeknik Negeri Sriwijaya tahun 2024, khususnya Suci Lutfia Nisa, Nadiyah Salsabila Fauzi, Cantika Tri Inayah, serta teman kelas 8 TEM.
7. Semua pihak lain yang telah membantu dalam menyelesaikan tugas akhir ini yang tidak dapat penulis sebutkan satu persatu.

Dalam penyusunan tugas akhir ini penulis mengharapkan kritik dan saran dari pembaca yang dapat membangun dan membantu menyempurnakan tugas akhir ini. Hal tersebut dikarenakan penulis menyadari bahwa masih terdapat banyak kekurangan dan kekeliruan baik dalam penyajian tulisan maupun isi dari tugas akhir ini. Akhir kata penulis berharap semoga tugas akhir ini dapat dilanjutkan menjadi tugas akhir yang bermanfaat bagi semua pembaca khususnya penulis serta semua mahasiswa Politeknik Negeri Sriwijaya Jurusan Teknik Elektro Program Studi Teknik Telekomunikasi.

Palembang, Juli 2024

Penulis

Ladysa Azzahrah

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN JUDUL	ii
HALAMAN PENGESAHAN	iii
SURAT PERNYATAAN	iv
MOTTO DAN PERSEMBAHAN	v
ABSTRAK	vi
ABSTRACT	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiii
DAFTAR PERSAMAAN MATEMATIS	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah	3
1.4 Tujuan	3
1.5 Manfaat	4
1.6 Metode Penulisan	4
1.7 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Plagiarisme	6
2.2 <i>Machine Learning</i>	7
2.3 Algoritma	10
2.3.1 <i>Naïve Bayes</i>	11
2.4 TF-IDF (<i>Term Frequency-Inverse Document Frequency</i>).....	13
2.5 <i>Cosine Similarity</i>	14
2.6 <i>Website</i>	15
2.6.1 Ngrok	16
2.7 Python.....	16
2.7.1 VS Code (<i>Visual Studio Code</i>).....	17
2.7.2 Flask.....	18
2.8 Penelitian – Penelitian Sebelumnya	19
BAB III METODOLOGI PENELITIAN	25
3.1 Kerangka Penelitian	25
3.2 Blok Diagram Sistem	26
3.3 Perancangan Penelitian Model <i>Machine Learning</i>	28
3.3.1 Dataset Penelitian	29
3.3.2 Tahap <i>Preprocessing</i> Data.....	30
3.3.3 Implementasi Algoritma <i>Naïve Bayes</i> dan TF-IDF	31
3.4 Perancangan dan Cara Kerja <i>Website</i>	32
3.5 Uji Skenario Sistem.....	34

BAB IV HASIL DAN PEMBAHASAN.....	36
4.1 Hasil Penelitian	36
4.1.1 Hasil Pengumpulan Dataset	36
4.1.2 Hasil <i>Preprocessing</i> Data.....	37
4.1.3 Hasil Implementasi TF-IDF	40
4.1.4 Hasil Implementasi Algoritma <i>Naïve Bayes</i>	42
4.1.5 Hasil <i>Cosine Similarity</i> dari TF-IDF	45
4.1.6 Hasil Implementasi ke <i>Website</i>	46
4.2 Hasil Pengujian Sistem.....	48
4.2.1 Pengujian Fungsional <i>Website</i>	48
4.2.2 Hasil Uji Skenario Sistem.....	55
4.3 Pembahasan	62
4.3.1 Analisis Hasil Penelitian dan Pengujian	62
4.3.2 Keterbatasan Penelitian	63
BAB V PENUTUP.....	65
5.1 Kesimpulan.....	65
5.2 Saran.....	66
DAFTAR PUSTAKA	67
LAMPIRAN	

DAFTAR GAMBAR

Gambar 2.1 <i>Machine learning</i> bekerja seperti manusia [10].....	7
Gambar 2.2 Tiga metode atau tipe <i>machine learning</i> [13].....	8
Gambar 2.3 Algoritma <i>Naïve Bayes</i> [18].....	11
Gambar 2.4 Dokumen serupa ditempatkan pada ruang multidimensi pada titik yang sama [25].....	15
Gambar 2.5 Logo Python [29]	16
Gambar 2.6 Logo VS Code [32]	17
Gambar 2.7 Logo Flask [35].....	18
Gambar 3.1 <i>Flowchart</i> Kerangka Penelitian	25
Gambar 3.2 Blok Diagram Sistem	26
Gambar 3.3 <i>Flowchart</i> Model <i>Machine Learning</i>	28
Gambar 3.4 <i>Flowchart</i> Use Case Website	33
Gambar 4.1 Hasil Pengumpulan Dataset	36
Gambar 4.2 Histogram Probabilitas <i>Naïve Bayes</i>	43
Gambar 4.3 Grafik Hasil Ekstraksi Fitur <i>Naïve Bayes</i>	44
Gambar 4.4 Visualisasi Hasil <i>Cosine Similarity</i>	45
Gambar 4.5 Tampilan Kode Pemrograman Implementasi ke Website	47
Gambar 4.6 Tampilan Halaman <i>Home</i>	49
Gambar 4.7 Tampilan Halaman <i>About</i>	51
Gambar 4.8 Tampilan Halaman Cek Plagiarisme Melalui Input Teks	52
Gambar 4.9 Tampilan Halaman <i>Upload File</i>	53
Gambar 4.10 Halaman Hasil Plagiarisme.....	54

DAFTAR TABEL

Tabel 2.1 Penelitian-Penelitian Sebelumnya	19
Tabel 4.1 Hasil Preprocessing Data.....	37
Tabel 4.2 Hasil TF Abstrak	40
Tabel 4.3 Hasil IDF Abstrak	41
Tabel 4.4 Hasil TF-IDF Abstrak.....	42
Tabel 4.5 Hasil Pengujian Blackbox pada Website	48
Tabel 4.6 Hasil Uji Skenario Pertama	55
Tabel 4.7 Hasil Uji Skenario Kedua.....	57
Tabel 4.8 Hasil Uji Skenario Ketiga.....	60
Tabel 4.9 Hasil Uji Skenario Keempat	61

DAFTAR PERSAMAAN MATEMATIS

(2. 1 Persamaan Teorema Bayes).....	12
(2. 2 Persamaan TF).....	14
(2. 3 Persamaan IDF).....	14
(2. 4 Persamaan TF-IDF).....	14

DAFTAR LAMPIRAN

- Lampiran 1** Daftar Riwayat Hidup
- Lampiran 2** Lembar Kesepakatan Bimbingan Tugas Akhir Pembimbing I
- Lampiran 3** Lembar Kesepakatan Bimbingan Tugas Akhir Pembimbing II
- Lampiran 4** Lembar Konsultasi Pembimbing I
- Lampiran 5** Lembar Konsultasi Pembimbing II
- Lampiran 6** Lembar Rekomendasi Ujian Tugas Akhir
- Lampiran 7** Lembar Revisi Ujian Tugas Akhir
- Lampiran 8** *Letter of Acceptance (LoA)*
- Lampiran 9** *Submitted Journal Article*
- Lampiran 10** *Source Coding*