

**KLASIFIKASI BERITA HOAKS PASCA PEMILIHAN UMUM  
PRESIDEN DAN WAKIL PRESIDEN REPUBLIK INDONESIA TAHUN  
2024 MENGGUNAKAN ALGORITMA *K-NEAREST NEIGHBOUR***

**TUGAS AKHIR**

Diajukan sebagai salah satu syarat memperoleh gelar Sarjana  
Program Studi S-1 Sistem Informasi



Disusun oleh:

**DENNIS FITRI SALSABILLA ARIANTI**

202103002

**PROGRAM STUDI S-1 SISTEM INFORMASI  
FAKULTAS TEKNIK & TEKNOLOGI INFORMASI  
UNIVERSITAS JENDERAL ACHMAD YANI YOGYAKARTA  
2024**

**HALAMAN PENGESAHAN**

**TUGAS AKHIR**

**KLASIFIKASI BERITA HOAKS PASCA PEMILIHAN UMUM PRESIDEN  
DAN WAKIL PRESIDEN REPUBLIK INDONESIA TAHUN 2024  
MENGUNAKAN ALGORITMA *K-NEAREST NEIGHBOUR***

Diajukan oleh:

**DENNIS FITRI SALSABILLA ARIANTI**  
202103002

Telah dipertahankan di depan dewan penguji dan dinyatakan sah  
sebagai salah satu syarat untuk memperoleh gelar Sarjana  
di Fakultas Teknik & Teknologi Informasi  
Universitas Jenderal Achmad Yani Yogyakarta

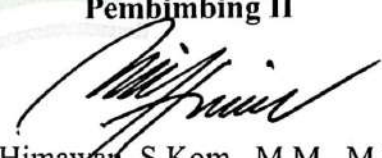
Tanggal: 12 Juli 2024

Mengesahkan:

**Pembimbing I**

  
Ulfi Saidata Aesyi, S.Kom., M.Cs.  
NIDN: 0515129002

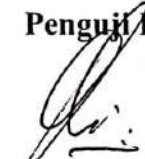
**Pembimbing II**

  
Arif Himawan, S.Kom., M.M., M.Eng.  
NIDN: 0517127402

**Penguji I**

  
Aris Wahyu Murdiyanto, S.Kom., M.Cs.  
NIDN: 0502098501

**Penguji II**

  
Kharisma, S.T., M.Cs.  
NIDN: 0502108201

Ketua Program Studi S-1 Sistem Informasi  
Fakultas Teknik & Teknologi Informasi  
Universitas Jenderal Achmad Yani Yogyakarta

  
Ahmad Hanafi, S.T., M.Eng.  
NPP: 2008.13.0020

## PERNYATAAN

Saya yang bertanda tangan di bawah ini, adalah mahasiswa Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta,

Nama : Dennis Fitri Salsabilla Arianti  
NPM : 202103002  
Program Studi : S-I Sistem Informasi  
Judul Tugas Akhir : Klasifikasi Berita Hoaks Paca Pemilihan Umum Presiden dan Wakil Presiden Republik Indonesia Tahun 2024 Menggunakan Algoritma *K-Nearest Neighbour*.

Menyatakan bahwa hasil penelitian dengan judul tersebut di atas adalah asli karya saya sendiri dan bukan hasil plagiarisme. Semua referensi dan sumber terkait yang dikutip dalam karya ilmiah ini telah ditulis sesuai kaidah penulisan ilmiah yang berlaku. Dengan ini, saya menyatakan untuk menyerahkan hak cipta penelitian kepada Universitas Jenderal Achmad Yani Yogyakarta guna kepentingan ilmiah.

Demikian surat pernyataan ini dibuat dengan sebenar-benarnya tanpa ada paksaan dari pihak mana pun. Apabila terdapat kekeliruan atau ditemukan adanya pelanggaran akademik di kemudian hari, maka saya bersedia menerima konsekuensi yang berlaku sesuai ketentuan akademik.

Yogyakarta, 20 Juni 2024



Dennis Fitri Salsabilla Arianti

## KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat-Nya sehingga penulis dapat menyelesaikan laporan tugas akhir yang berjudul: “Klasifikasi Berita Hoaks Pasca Pemilihan Umum Presiden dan Wakil Presiden Republik Indonesia Tahun 2024 Menggunakan Algoritma *K-Nearest Neighbour*”. Penyusunan laporan ini merupakan salah satu persyaratan untuk menyelesaikan studi di Program Studi S-1 Sistem Informasi Fakultas Teknik & Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta. Laporan ini dapat diselesaikan atas bimbingan, arahan, dan bantuan dari berbagai pihak. Pada kesempatan ini penulis dengan rendah hati mengucapkan terima kasih dengan setulus-tulusnya kepada:

1. Bapak Aris Wahyu Murdiyanto, S.Kom., M.Cs. selaku Dekan Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta;
2. Bapak Ahmad Hanafi, S.T., M.Eng. selaku Ketua Program Studi S-1 Sistem Informasi Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta;
3. Ibu Ulfi Saidata Aesyti, S.Kom., M.Cs. dan Bapak Arif Himawan, S.Kom., M.M., M.Eng. selaku Dosen Pembimbing Tugas Akhir;
4. Para dosen yang telah memberikan banyak bekal ilmu pengetahuan kepada penulis selama menjadi mahasiswa di Fakultas Teknik dan Teknologi Informasi Universitas Jenderal Achmad Yani Yogyakarta;
5. Ayah, ibu, kakak dan adik saya, yang telah memberikan dukungan semangat serta doa restu kepada saya, sehingga dapat menyelesaikan studi saya;
6. Sahabatku Krisna, Arum, Rachel, Latifah, Laela, dan Moudy yang telah memberikan semangat dan doa kepada saya;
7. Rekan-rekan mahasiswa Prodi S-1 Sistem Informasi di Universitas Jenderal Achmad Yani Yogyakarta yang sudah memberi dukungan dan kerja sama selama pembuatan tugas akhir.

Penulis menyadari bahwa laporan tugas akhir ini masih jauh dari kata sempurna. Maka dari itu dengan segala kerendahan hati penulis sangat menghargai adanya kritik dan saran yang membangun dari semua pihak yang bersedia meluangkan waktu untuk membaca laporan tugas akhir ini.

Yogyakarta, 20 Juni 2024

Dennis Fitri Salsabilla Arianti

PERPUSTAKAAN  
UNIVERSITAS JENDERAL ACHMAD YANI  
YOGYAKARTA

## DAFTAR ISI

|                                                                       |             |
|-----------------------------------------------------------------------|-------------|
| <b>Halaman Pengesahan .....</b>                                       | <b>ii</b>   |
| <b>Pernyataan .....</b>                                               | <b>iii</b>  |
| <b>Kata Pengantar .....</b>                                           | <b>iv</b>   |
| <b>Daftar Isi.....</b>                                                | <b>vi</b>   |
| <b>Daftar Tabel.....</b>                                              | <b>ix</b>   |
| <b>Daftar Gambar .....</b>                                            | <b>xi</b>   |
| <b>Daftar Lampiran .....</b>                                          | <b>xii</b>  |
| <b>Daftar Singkatan .....</b>                                         | <b>xiii</b> |
| <b>Intisari.....</b>                                                  | <b>xiv</b>  |
| <b>Abstract .....</b>                                                 | <b>xv</b>   |
| <b>Bab 1   Pendahuluan .....</b>                                      | <b>1</b>    |
| 1.1 Latar Belakang .....                                              | 1           |
| 1.2 Perumusan Masalah .....                                           | 2           |
| 1.3 Batasan Masalah .....                                             | 2           |
| 1.4 Pertanyaan Penelitian .....                                       | 3           |
| 1.5 Tujuan Penelitian .....                                           | 3           |
| 1.6 Manfaat Hasil Penelitian .....                                    | 4           |
| <b>Bab 2   Tinjauan Pustaka dan Landasan Teori.....</b>               | <b>5</b>    |
| 2.1 Tinjauan Pustaka .....                                            | 5           |
| 2.2 Landasan Teori .....                                              | 8           |
| 2.2.1 <i>Text Mining</i> .....                                        | 8           |
| 2.2.2 Klasifikasi .....                                               | 8           |
| 2.2.3 Hoaks .....                                                     | 9           |
| 2.2.4 <i>Text Preprocessing</i> .....                                 | 10          |
| 2.2.5 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i> ..... | 12          |
| 2.2.6 <i>K-Nearest Neighbour (KNN)</i> .....                          | 13          |
| 2.2.7 <i>Confusion Matrix</i> .....                                   | 14          |
| 2.2.8 Probabilitas .....                                              | 17          |

|              |                                |           |
|--------------|--------------------------------|-----------|
| <b>Bab 3</b> | <b>Metode Penelitian</b>       | <b>18</b> |
| 3.1          | Bahan dan Alat Penelitian      | 20        |
| 3.2          | Jalan Penelitian               | 21        |
| <b>Bab 4</b> | <b>Hasil Penelitian</b>        | <b>24</b> |
| 4.1          | Ringkasan Hasil Penelitian     | 24        |
| 4.2          | <i>Bussiness Understanding</i> | 24        |
| 4.3          | Data Understanding             | 25        |
| 4.3.1        | Turnbackhoax                   | 25        |
| 4.3.2        | Infopublik                     | 27        |
| 4.3.3        | RRI                            | 28        |
| 4.3.4        | Kominfo                        | 29        |
| 4.3.5        | Kompas                         | 30        |
| 4.3.6        | Liputan6                       | 31        |
| 4.3.7        | Detik                          | 32        |
| 4.3.8        | AntaraNews                     | 33        |
| 4.3.9        | CNN                            | 34        |
| 4.3.10       | Okezone                        | 35        |
| 4.3.11       | Sindonews                      | 36        |
| 4.3.12       | Kumparan                       | 37        |
| 4.3.13       | Pikiran Rakyat                 | 38        |
| 4.3.14       | Wartatransparansi              | 39        |
| 4.4          | Data Preparation               | 40        |
| 4.4.1        | Normalisasi                    | 40        |
| 4.4.2        | Pre-processing Text            | 41        |
| 4.4.3        | Pembobotan Kata (TF-IDF)       | 46        |
| 4.5          | Modeling                       | 47        |
| 4.6          | Evaluasi                       | 48        |
| 4.7          | Deployment                     | 51        |
| 4.7.1        | Kompas                         | 56        |
| 4.7.2        | Liputan6                       | 58        |
| 4.7.3        | Detik                          | 59        |

|                                                   |           |
|---------------------------------------------------|-----------|
| 4.7.4 Prediksi Berita Antaranews .....            | 60        |
| 4.7.5 Prediksi Berita CNN Indonesia .....         | 61        |
| 4.7.6 Klasifikasi Berita Okezone .....            | 62        |
| 4.7.7 Klasifikasi Berita Sindonews .....          | 64        |
| 4.7.8 Kumparan .....                              | 65        |
| 4.7.9 Klasifikasi Berita Pikiran Rakyat .....     | 66        |
| 4.7.10 Klasifikasi Berita Wartatransparansi ..... | 67        |
| <b>Bab 5 Kesimpulan dan Saran .....</b>           | <b>70</b> |
| 4.8 Kesimpulan .....                              | 70        |
| 4.9 Saran .....                                   | 70        |
| <b>Daftar Pustaka .....</b>                       | <b>72</b> |
| <b>Lampiran .....</b>                             | <b>76</b> |

## DAFTAR TABEL

|                                                                              |    |
|------------------------------------------------------------------------------|----|
| Tabel 2. 1 Tinjauan Pustaka .....                                            | 6  |
| Tabel 2. 2 <i>Confusion Matrix</i> .....                                     | 14 |
| Tabel 4. 1 Sampel Hasil Scraping Website Turnbackhoax.id .....               | 26 |
| Tabel 4. 2 Sampel Hasil Scraping Website Infopublik.id .....                 | 27 |
| Tabel 4. 3 Sampel Hasil Scraping Website Rri.co.id .....                     | 28 |
| Tabel 4. 4 Sampel Hasil Scraping Website Kominfo.go.id .....                 | 29 |
| Tabel 4. 5 Sampel Hasil Scraping Website Kompas.com .....                    | 31 |
| Tabel 4. 6 Sampel Hasil Scraping Website Liputan6.com .....                  | 31 |
| Tabel 4. 7 Sampel Hasil Scraping Website Detik.com .....                     | 32 |
| Tabel 4. 8 Sampel Hasil Scraping Website Antaraneews.com .....               | 33 |
| Tabel 4. 9 Sampel Hasil Scraping Website Cnnindonesia.com .....              | 34 |
| Tabel 4. 10 Sampel Hasil Scraping Website Okezone.com .....                  | 35 |
| Tabel 4. 11 Sampel Hasil Scraping Website Sindonews.com .....                | 36 |
| Tabel 4. 12 Sampel Hasil Scraping Website Kumparan.com .....                 | 37 |
| Tabel 4. 13 Sampel Hasil Scraping Website Pikiranrakyat.com .....            | 38 |
| Tabel 4. 14 Sampel Hasil Scraping Website wartatransparansi.com .....        | 39 |
| Tabel 4. 15 Penerapan Case Folding .....                                     | 41 |
| Tabel 4. 16 Penerapan Replace Texts .....                                    | 42 |
| Tabel 4. 17 Penerapan Cleaning .....                                         | 43 |
| Tabel 4. 18 Penerapan Cleaning .....                                         | 44 |
| Tabel 4. 19 Penerapan Stemming .....                                         | 45 |
| Tabel 4. 20 Penerapan Tokenizing .....                                       | 46 |
| Tabel 4. 22 Hasil <i>Testing Accuracy</i> dan <i>Training Accuracy</i> ..... | 47 |
| Tabel 4. 23 Hasil TF-IDF Matrix .....                                        | 54 |
| Tabel 4. 24 Sampel Hasil Implementasi Model pada Website Kompas.com .....    | 56 |
| Tabel 4. 25 Sampel Hasil Implementasi Model pada Website Liputan6.com .....  | 58 |
| Tabel 4. 26 Sampel Hasil Implementasi Model pada Website Detik.com .....     | 59 |
| Tabel 4. 27 Sampel Hasil Implementasi Model pada Website Antaraneews.com ..  | 60 |
| Tabel 4. 28 Sampel Hasil Implementasi Model pada Website Cnnindonesia.com    | 61 |

|                                                                                      |    |
|--------------------------------------------------------------------------------------|----|
| Tabel 4. 29 Sampel Hasil Implementasi Model pada Website Okezone.com .....           | 62 |
| Tabel 4. 30 Sampel Hasil Implementasi Model pada Website Sindonews.com ...           | 64 |
| Tabel 4. 31 Sampel Hasil Implementasi Model pada Website Kumparan.com ....           | 65 |
| Tabel 4. 32 Sampel Hasil Implementasi Model pada Website Pikiranrakyat.com           | 66 |
| Tabel 4. 33 Sampel Hasil Implementasi Model pada Website Wartatransparansi.com ..... | 67 |
| Tabel 4. 34 Hasil Implementasi Model .....                                           | 69 |

PERPUSTAKAAN  
UNIVERSITAS JENDERAL ACHMAD YANI  
YOGYAKARTA

## DAFTAR GAMBAR

|                                                                                            |    |
|--------------------------------------------------------------------------------------------|----|
| Gambar 2. 1 Case Folding.....                                                              | 10 |
| Gambar 2. 1 Replace Texts.....                                                             | 10 |
| Gambar 2. 1 Filtering.....                                                                 | 11 |
| Gambar 2. 1 Cleaning.....                                                                  | 11 |
| Gambar 2. 1 Stemming.....                                                                  | 12 |
| Gambar 2. 1 Tokenizing.....                                                                | 12 |
| Gambar 3. 1 Metode Penelitian.....                                                         | 18 |
| Gambar 3. 2 Jalan Penelitian.....                                                          | 22 |
| Gambar 4. 1 Proses Scraping Website Turnbackhoax.id.....                                   | 25 |
| Gambar 4. 2 Proses Scraping Website Kominfo.go.id.....                                     | 25 |
| Gambar 4. 3 Proses Scraping Website Kompas.com.....                                        | 25 |
| Gambar 4. 4 Proses Scraping Website Cnnindonesia.com.....                                  | 50 |
| Gambar 4. 5 Proses Scraping Website Kumparan.com.....                                      | 37 |
| Gambar 4. 6 Proses Scraping Website Wartatransparansi.com.....                             | 39 |
| Gambar 4. 7 Proses Normalisasi Data.....                                                   | 40 |
| Gambar 4. 8 Kode Program Case Folding.....                                                 | 41 |
| Gambar 4. 9 Kode Program Cleaning.....                                                     | 42 |
| Gambar 4. 10 Kode Program Filtering.....                                                   | 43 |
| Gambar 4. 11 Kode Program Stemming.....                                                    | 44 |
| Gambar 4. 12 Kode Program Tokenzing.....                                                   | 45 |
| Gambar 4. 13 Grafik Perbandingan <i>Training Accuracy</i> dan <i>Testing Accuracy</i> .... | 48 |
| Gambar 4. 14 <i>Classification Report</i> .....                                            | 50 |
| Gambar 4. 15 <i>Cofussion Matrix</i> .....                                                 | 51 |
| Gambar 4. 16 Kode Program Import Library untuk Modelling.....                              | 52 |
| Gambar 4. 17 Kode Program Memasukan dan Membaca File.....                                  | 53 |
| Gambar 4. 18 Kode Program TF-IDF.....                                                      | 54 |
| Gambar 4. 19 Kode Implementasi Model.....                                                  | 55 |
| Gambar 4. 20 Kode Program Menyimpan File Hasil Prediksi.....                               | 56 |

## DAFTAR LAMPIRAN

|                                                               |    |
|---------------------------------------------------------------|----|
| Lampiran 1 Kode Program Scraping Website Infopublik .....     | 77 |
| Lampiran 2 Kode Program Scraping Website Liputan6 .....       | 78 |
| Lampiran 3 Kode Program Scraping Website Detik .....          | 80 |
| Lampiran 4 Kode Program Scraping Website Antaranews .....     | 81 |
| Lampiran 5 Kode Program Scraping Website Okezone .....        | 82 |
| Lampiran 6 Kode Program Scraping Website Sindonews .....      | 83 |
| Lampiran 7 Kode Program Replace Texts .....                   | 84 |
| Lampiran 8 Kode Program TF-IDF .....                          | 86 |
| Lampiran 9 Kode Program Pemodelan KNN .....                   | 87 |
| Lampiran 10 Hasil Cross Validation .....                      | 89 |
| Lampiran 11 Kode Program Preprocessing Tahap Deployment ..... | 93 |
| Lampiran 12 Jadwal Penelitian .....                           | 96 |
| Lampiran 13 Lembar Bimbingan Dosen .....                      | 97 |
| Lampiran 14 Hasil Cek Plagiarisme .....                       | 98 |

## DAFTAR SINGKATAN

|          |                                                 |
|----------|-------------------------------------------------|
| KNN      | K-Nearest Neighbour                             |
| TF-IDF   | Term Frequency-Inverse Document Frequency       |
| CRISP-DM | Cross-Industry Standard Process for Data Mining |
| LSTM     | Long Short Term Memory                          |
| RNN      | Recurrent Neural Network                        |

PERPUSTAKAAN  
UNIVERSITAS JENDERAL ACHMAD YANI  
YOGYAKARTA