

**IMPLEMENTASI ALGORITMA DECISION TREE DALAM
KLASIFIKASI OPINI PENGGUNA MEDIA SOSIAL X
(TWITTER) TERHADAP JNE DI INDONESIA**

Tugas Akhir ini disusun untuk memenuhi syarat mencapai gelar Kesarjanaan
Komputer pada Program Studi Teknik Informatika Jenjang Program Strata-1



Oleh :

Muhamad Firmansyah

18.01.53.2014

FAKULTAS TEKNOLOGI INFORMASI DAN INDUSTRI
UNIVERSITAS STIKUBANK (UNISBANK)

SEMARANG

2024

SURAT PERNYATAAN KESIAPAN UJIAN AKHIR

Saya, MUHAMAD FIRMANSYAH deng ini menyatakan bahwa laporan yang berjudul :

IMPLEMENTASI ALGORITMA DECISION TREE DALAM KLASIFIKASI OPINI PENGGUNA MEDIA SOSIAL X (TWITTER) TERHADAP JNE DI INDONESIA

adalah benar hasil karya saya dan belum pernah diajukan sebagai karya ilmiah sebagian atau seluruhnya atau pihak lain.



MUHAMAD FIRMANSYAH
18.01.53.2014

Disetujui Oleh Pembimbing
Kami setuju laporan tersebut diajukan untuk Ujian
Semarang, 09 Agustus 2024



Dr. EKA ARDHIANTO, S.Kom., M.Cs.
Pembimbing



SURAT PERNYATAAN KEASLIAN SKRIPSI

Yang bertanda tangan di bawah ini, saya menyatakan bahwa skripsi dengan Judul :
IMPLEMENTASI ALGORITMA DECISION TREE DALAM KLASIFIKASI OPINI PENGGUNA MEDIA SOSIAL X (TWITTER) TERHADAP JNE DI INDONESIA

yang telah diuji di depan tim penguji pada tanggal 14-08-2024, adalah benar hasil karya saya dan dalam skripsi ini tidak terdapat keseluruhan atau sebagian tulisan orang lain yang saya ambil dengan cara menyalin, atau meniru dalam bentuk rangkaian kalimat atau simbol yang saya aku seolah-olah sebagai tulisan saya sendiri dan atau tidak terdapat bagian atau keseluruhan tulisan yang saya salin, tiru atau yang saya ambil dari tulisan orang lain tanpa memberikan pengakuan pada penulis aslinya.

Apabila saya melakukan hal tersebut diatas, baik sengaja maupun tidak, dengan ini saya menyatakan menarik skripsi yang saya ajukan sebagai hasil tulisan saya sendiri.

Bila kemudian terbukti bahwa saya ternyata melakukan tindakan menyalin atau meniru tulisan orang lain seolah-olah hasil pemikiran saya sendiri, berarti gelar dan ijazah saya yang telah diberikan oleh Universitas Stikubank (UNISBANK) Semarang batal saya terima.



Semarang, 14-08-2024
Yang Menyatakan

(MUHAMAD FIRMANSYAH)
NIM. 18.01.53.2014

SAKSI 1
Tim Penguji



(Dr. EKA ARDHIANTO, S.Kom., M.Cs.)

SAKSI 2
Tim Penguji



(SAEFURROHMAN, S.KOM. M.Cs)

SAKSI 3
Tim Penguji



(SRI ENIYATI, S.KOM, M.Cs)



HALAMAN PENGESAHAN

Skripsi dengan Judul
**IMPLEMENTASI ALGORITMA DECISION TREE DALAM KLASIFIKASI OPINI
PENGGUNA MEDIA SOSIAL X (TWITTER) TERHADAP JNE DI INDONESIA**

Ditulis oleh
NIM : **18.01.53.2014**
Nama : **MUHAMAD FIRMANSYAH**

Telah dipertahankan di depan Tim Dosen Penguji skripsi dan diterima sebagai salah satu syarat guna menyelesaikan Jenjang Program S1 Program Studi Teknik Informatika (R2) pada Fakultas TEKNOLOGI INFORMASI DAN INDUSTRI Universitas Stikubank (UNISBANK) Semarang.

Semarang, 20-08-2024
Ketua



(Dr. EKA ARDHIANTO, S.Kom., M.Cs.)
NIDN. 0610038201

Sekretaris



(SAEFURROHMAN, S.KOM, M.Cs)
NIDN. 0609027602

Anggota



(SRI ENIYATI, S.KOM, M.Cs)
NIDN. 0627096601

Mengetahui,
Universitas Stikubank (UNISBANK) Semarang
Fakultas TEKNOLOGI INFORMASI DAN INDUSTRI
Dekan



(Dr. AJI SUPRIYANTO, S.T., M.Kom.)
NIDN. 0628077101



MOTTO DAN PERSEMBAHAN

Moto :

1. "Setiap orang punya gilirannya sendiri-sendiri, bersabarlah dan tunggulah!." (Gol D. Roger)
2. "Selama Masih Hidup, Kesempatan Itu Tak Terbatas" (Monkey D. Luffy)
3. "Jika kamu lapar, makanlah." (Monkey D. Luffy)

Persembahan :

1. Kepada Kedua Orang tua, yang telah memberikan doa, dukungan, dan kasih sayang tanpa henti sehingga penulis dapat menyelesaikan skripsi ini.
2. Bapak Dr. Edy Winarno S.T., M.Eng. selaku Rektor Universitas Stikubank Semarang.
3. Bapak Dr. Aji Supriyanto S.T., M.Kom.selaku Dekan Fakultas Teknologi Informasi Dan Industri.
4. Bapak Dr. Drs. Eri Zuliarso M.Kom. Kaprodi Teknik Informatika,
5. Bapak Dr. Eka Ardhianto S.Kom., M.Cs.yang telah memberikan arahan, dukungan, dan bimbingan dalam setiap langkah penelitian ini. Terima kasih atas perhatian dan dedikasi dalam memastikan kemajuan akademis penulis.
6. Saudara Dimas Indra Purwadi selaku Rekan kakak tingkat saya di Universitas Stikubank Semarang.
7. Rekan-rekan di Universitas Stikubank yang Namanya tidak bisa saya sebutkan satu per satu.

8. Semua pihak-pihak yang telah membatu dan memberikan dukungan terhadap penulis dalam skripsi ini.

**FAKULTAS TEKNOLOGI INFORMASI DAN INDUSTRI
UNIVERSITAS STIKUBANK (UNISBANK) SEMARANG**

Program Studi Teknik Informatika

Tugas Akhir Sarjana Komputer

Semester 12 Tahun 2024

**IMPLEMENTASI ALGORITMA DECISION TREE DALAM
KLASIFIKASI OPINI PENGGUNA MEDIA SOSIAL X (TWITTER)
TERHADAP JNE DI INDONESIA**

Muhamad Firmansyah

18.01.53.2014

ABSTRAK

JNE, salah satu perusahaan ekspedisi terdepan di Indonesia, hadir dengan jaringan yang tersebar di seluruh penjuru tanah air, sehingga memudahkan masyarakat dalam mengirim barang ke seluruh daerah di Indonesia. Dengan banyaknya pengguna JNE, muncul beragam opini pengguna yang tersebar di media sosial, baik itu opini positif maupun negatif. Untuk mengelompokkan opini-opini ini secara efektif, diperlukan penggunaan program machine learning. Algoritma yang menjadi kunci untuk mengklasifikasikan opini adalah Decision Tree. Sebelum proses klasifikasi dilakukan, data tweet yang telah dikumpulkan harus melalui tahap pelabelan data dan preprocessing agar dapat dikenali oleh sistem secara maksimal. Saat penelitian ini dilakukan, dengan penggunaan rasio perbandingan antara data training dan data testing 90:10, maka algoritma Decision Tree dapat mencapai akurasi sebesar 88,10%. Ini menunjukkan potensi yang signifikan dalam klasifikasi opini pengguna.

Kata Kunci : JNE, Analisis Sentimen, Text Mining,, Decision Tree.

ABSTRACT

JNE is one of the leading goods delivery service companies in Indonesia which has a wide network, making it easier for people to send goods to all regions in Indonesia. With so many JNE users, various user opinions have emerged on social media, both positive and negative views. Grouping these opinions effectively requires the use of machine learning programs. One of the algorithms used to classify opinions is a Decision Tree. Before the classification process is carried out, the tweet data that has been collected must go through the data labeling and preprocessing stages so that it can be recognized by the system properly. In this research, by using a ratio of training data to testing data of 90:10, the Decision Tree algorithm can achieve an accuracy of 88.10%. This shows significant potential in user opinion classification.

Keywords: JNE, Sentiment Analysis, Text Mining,, Decision Tree.

Pembimbing



Dr. Eka Ardhiyanto, S.Kom, M.Cs
NIDN 06100338201

KATA PENGANTAR

Segala puji dan syukur kita panjatkan ke hadirat Allah SWT yang telah memberikan rahmat dan karunia-Nya, Sehingga penulis dapat menyelesaikan penelitian ini. Bersama-sama kita panjatkan sholawat dan salam kepada Nabi Muhammad SAW, sebagai ungkapan cinta dan penghormatan..

Dalam upaya menyusun skripsi ini, penulis ingin menyampaikan rasa terima kasih kepada semua pihak yang turut serta dalam menyelesaikan karya tulis ini. Skripsi ini dihasilkan dengan penuh dedikasi dan semangat, sebagai upaya kontribusi kecil penulis dalam pengembangan ilmu pengetahuan. Semoga hasil penelitian ini dapat memberikan manfaat dan kontribusi positif bagi perkembangan teknologi informasi di Indonesia.

Akhir kata, penulis mengakui bahwa penelitian ini masih memiliki banyak kekurangan dan keterbatasan. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun dari pembaca sangat berguna untuk perbaikan di masa yang akan datang.

Semarang, 25 Januari 2024

Penulis



Muhamad Firmansyah

DAFTAR ISI

| | |
|---|------|
| SURAT PERNYATAAN KESIAPAN TUGAS AKHIR | ii |
| SURAT PERNYATAAN KEASLIAN TUGAS AKHIR..... | iii |
| HALAMAN PENGESAHAN..... | iv |
| MOTTO DAN PERSEMBAHAN | v |
| ABSTRAK | vii |
| ABSTRACT | viii |
| KATA PENGANTAR..... | ix |
| DAFTAR ISI | x |
| DAFTAR GAMBAR | xv |
| DAFTAR TABEL..... | xvii |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang Masalah..... | 1 |
| 1.2 Rumusan Masalah | 2 |
| 1.3 Batasan Masalah..... | 2 |
| 1.4 Tujuan Penelitian..... | 3 |
| 1.5 Manfaat Penelitian | 3 |
| 1.6 Usulan Pemecahan Masalah..... | 4 |
| BAB II TINJAUAN PUSTAKA | 5 |

| | | |
|--------------------------------|---------------------------|----|
| 2.1 | Penelitian Terkait..... | 5 |
| 2.2 | Landasan Teori | 10 |
| 2.2.1 | Data Mining..... | 10 |
| 2.2.2 | Text Mining | 11 |
| 2.2.3 | Sentiment Analyst..... | 11 |
| 2.2.4 | Decision Tree..... | 12 |
| 2.2.5 | Media Sosial | 14 |
| 2.2.6 | X (Twitter)..... | 14 |
| 2.2.7 | JNE Indonesia..... | 15 |
| 2.2.8 | Python..... | 18 |
| 2.2.9 | Twitter API | 18 |
| 2.2.10 | Crawling Data..... | 20 |
| 2.2.11 | Preprocessing..... | 21 |
| 2.2.12 | Sastrawi | 21 |
| 2.2.13 | TF-IDF..... | 21 |
| 2.2.14 | Pemodelan Splitting | 23 |
| 2.2.15 | Confusion Matrix | 24 |
| 2.3 | Kerangka Pemikiran..... | 25 |
| BAB III METODE PENELITIAN..... | | 26 |
| 3.1 | Instrumen Penelitian..... | 26 |

| | | |
|---|---|----|
| 3.1.1 | Kebutuhan Perangkat Lunak | 26 |
| 3.1.2 | Kebutuhan Perangkat Keras | 26 |
| 3.2 | Alur Metode | 27 |
| 3.3 | Teknik Pengumpulan Data | 28 |
| 3.4 | Teknik Analisis Data | 30 |
| 3.5 | Metode yang diusulkan | 31 |
| 3.6 | Metode Pengujian..... | 32 |
| BAB IV HASIL PENELITIAN & PEMBAHASAN..... | | 34 |
| 4.1 | Pengambilan Data | 34 |
| 4.2 | Labeling Data | 37 |
| 4.3 | Preprocessing Data..... | 38 |
| 4.4 | Pembobotan TF-IDF | 41 |
| 4.5 | Klasifikasi | 44 |
| 4.5.1 | Menentukan Root(Akar)..... | 45 |
| 4.5.2 | Menentukan Node(simpul)..... | 48 |
| 4.5.3 | Menentukan Leaf Node (Simpul Cabang)..... | 51 |
| 4.6 | Evaluasi | 53 |
| 4.7 | Prediksi..... | 54 |
| 4.8 | Hasil Penelitian | 55 |
| 4.9 | Tampilan Program | 56 |

| | | |
|----------------------|--|----|
| 4.9.1 | Menginstall Paket Phyton yang Diperlukan | 56 |
| 4.9.2 | Mengimport Library Python..... | 57 |
| 4.9.3 | Crawling Data..... | 57 |
| 4.9.4 | Labeling Data | 59 |
| 4.9.5 | Preprocessing..... | 60 |
| 4.9.6 | Menghitung Frekuensi kata | 62 |
| 4.9.7 | Hasil Preprocessing | 63 |
| 4.9.8 | Menghitung Kata Unik..... | 64 |
| 4.9.9 | Ubah Label ke Binary Number..... | 64 |
| 4.9.10 | Cek Total Data Positif dan Negatif..... | 65 |
| 4.9.11 | TF-IDF | 66 |
| 4.9.12 | Data Spliting..... | 66 |
| 4.9.13 | Membuat Model Decision Tree | 68 |
| 4.9.14 | Evaluasi | 69 |
| 4.9.15 | Prediksi Tweet | 69 |
| 4.9.16 | Hasil Penelitian..... | 71 |
| BAB V PENUTUP..... | | 73 |
| 5.1 | Kesimpulan | 73 |
| 5.2 | Saran..... | 74 |
| DAFTAR PUSTAKA | | 75 |

| | |
|---------------|----|
| CODING | 79 |
| LAMPIRAN..... | 95 |

DAFTAR GAMBAR

| | |
|---|----|
| Gambar 2. 1 Kerangka Pemikiran..... | 25 |
| Gambar 3. 1 Alur Metode..... | 27 |
| Gambar 3. 2 Teknik Pengumpulan Data | 29 |
| Gambar 3. 3 Teknik Analisis Data | 30 |
| Gambar 3. 4 Confusion Matrix | 32 |
| Gambar 4. 1 Crawling Data Twitter | 35 |
| Gambar 4. 2 Menentukan jalur ke file CSV..... | 35 |
| Gambar 4. 3 Data Hasil Crawling yang telah di Download..... | 36 |
| Gambar 4. 4 Informasi Data Hasil Crawling | 36 |
| Gambar 4. 5 Data Labeling | 37 |
| Gambar 4. 6 Root(Akar) | 48 |
| Gambar 4. 7 Node(Simpul)..... | 51 |
| Gambar 4. 8 Leaf Node (Simpul Cabang) | 52 |
| Gambar 4. 9 Install Paket Python..... | 56 |
| Gambar 4. 10 Import Library Python..... | 57 |
| Gambar 4. 11 Menginstal Node.js..... | 58 |
| Gambar 4. 12 Proses Crawling Data | 58 |
| Gambar 4. 13 Jalur Dataframe file CSV | 59 |
| Gambar 4. 14 Membaca Data Dari File Excel | 59 |
| Gambar 4. 15 Cleansing | 60 |
| Gambar 4. 16 Case Folding..... | 61 |

| | |
|---|----|
| Gambar 4. 17 Tokenize | 61 |
| Gambar 4. 18 Stopword Removal..... | 62 |
| Gambar 4. 19 Stemming | 62 |
| Gambar 4. 20 Menghitung Frekuensi Kata | 63 |
| Gambar 4. 21 Hasil Preprocessing..... | 63 |
| Gambar 4. 22 Menghitung Kata Unik..... | 64 |
| Gambar 4. 23 Ubah Label ke Binary Number | 65 |
| Gambar 4. 24 Total Data Positif dan Negatif dalam bentuk Pie Bar | 65 |
| Gambar 4. 25 Perhitungan TF-IDF | 66 |
| Gambar 4. 26 50% Data Training & 50% Data Testing..... | 67 |
| Gambar 4. 27 60% Data Training & 40% Data Testing..... | 67 |
| Gambar 4. 28 70% Data Training & 30% Data Testing..... | 67 |
| Gambar 4. 29 80% Data Training & 20% Data Testing..... | 68 |
| Gambar 4. 30 90% Data Training & 10% Data Testing..... | 68 |
| Gambar 4. 31 Membuat Model Decision Tree..... | 68 |
| Gambar 4. 32 Mengevaluasi Model Decision Tree..... | 69 |
| Gambar 4. 33 Kata-Kata untuk tes Model | 70 |
| Gambar 4. 34 Prediksi..... | 70 |
| Gambar 4. 35 Program untuk Membuat Diagam Batang Tweet Label Positif | 71 |
| Gambar 4. 36 Diagram Batang Kata yang Sering Muncul di Label Positif..... | 72 |
| Gambar 4. 37 Program untuk Membuat Diagam Batang Tweet Label Negatif... | 72 |
| Gambar 4. 38 Diagram Batang Kata yang Sering Muncul di Label Negatif..... | 72 |

DAFTAR TABEL

| | |
|---|----|
| Tabel 2. 1 Penelitian Terkait..... | 8 |
| Tabel 4. 1 Data Tweet Asli | 38 |
| Tabel 4. 2 Cleansing | 39 |
| Tabel 4. 3 Case Folding..... | 39 |
| Tabel 4. 4 Tokenize | 40 |
| Tabel 4. 5 Stopword Removal..... | 40 |
| Tabel 4. 6 Stemming | 41 |
| Tabel 4. 7 Hasil Preprocessing | 41 |
| Tabel 4. 8 Contoh Tweet sesudah Preprocessing | 42 |
| Tabel 4. 9 Perhitungan TF | 42 |
| Tabel 4. 10 TF Normalisasi..... | 43 |
| Tabel 4. 11 Perhitungan TF-IDF | 43 |
| Tabel 4. 12 Nilai Data Training 1 | 45 |
| Tabel 4. 13 Perhitungan Entropy dan Gain | 47 |
| Tabel 4. 14 Nilai Data Training 2..... | 48 |
| Tabel 4. 15 Perhitungan Entropy dan Gain | 50 |
| Tabel 4. 16 Nilai Data Training..... | 51 |
| Tabel 4. 17 Perhitungan Entropy dan Gain | 52 |
| Tabel 4. 18 Contoh Tweet yang akan diprediksi | 54 |
| Tabel 4. 19 Contoh Tweet setelah diprediksi | 54 |
| Tabel 4. 20 Tabel Kemunculan Kata | 55 |