

SKRIPSI

Penerapan Metode K-NN Untuk Analisis Sentimen

Ulasan Pengguna Aplikasi X di *Playstore*



ISNI ZULIANI

205610027

PROGRAM STUDI SISTEM INFORMASI

PROGRAM SARJANA

FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA

YOGYAKARTA

2024

SKRIPSI

PENERAPAN METODE KNN UNTUK ANALISIS SENTIMEN

ULASAN PENGGUNA APLIKASI X DI *PLAYSTORE*

Diajukan sebagai salah satu untuk menyelesaikan studi

Program Sarjana

Program Studi Sistem Informasi

Fakultas Teknologi Informasi

Universitas Teknologi Digital Indonesia

Yogyakarta

Disusun oleh:

ISNI ZULIANI

205610027

PROGRAM STUDI SISTEM INFORMASI

PROGRAM SARJANA

FAKULTAS TEKNOLOGI DIGITAL INFORMASI

UNIVERSITAS TEKNOLOGI DIGITAL INDONESIA

YOGYAKARTA

2024

HALAMAN PERSETUJUAN

UJIAN SKRIPSI

Judul : Penerapan Metode KNN Untuk Analisis Sentimen
Ulasan Pengguna aplikasi X di *Playstore*

Nama : Isnı Zuliani

NIM : 205610017

Program Studi : Sistem Informasi

Program : Sarjana

Semester : Gasal

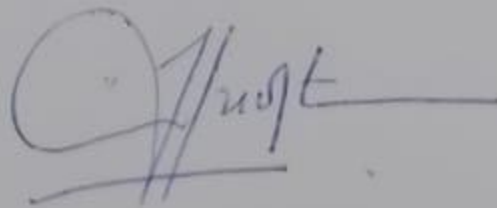
Tahun Akademik : 2023/2024

Telah diperiksa dan disetujui untuk dijadikan di hadapan

Dewan Penguji Skripsi

Yogyakarta, 29. Februari 2024

Dosen Pembimbing



Sumiyatun, S.Kom, M.Cs.

NIDN: 0515048402

HALAMAN PENGESAHAN

SKRIPSI

PENERAPAN METODE KNN UNTUK ANALISIS SENTIMEN

ULASAN PENGGUNA APLIKASI X DI *PLAYSTORE*

Telah dipertahankan di depan Dewan Penguji Skripsi dan dinyatakan diterima untuk memenuhi sebagai persyaratan guna memperoleh Gelar

Sarjana Komputer

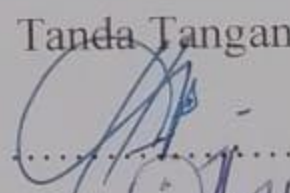
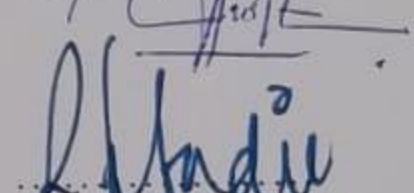

Program Studi Sistem Informasi

Fakultas Teknologi Informasi

Universitas Teknologi Digital Indonesia

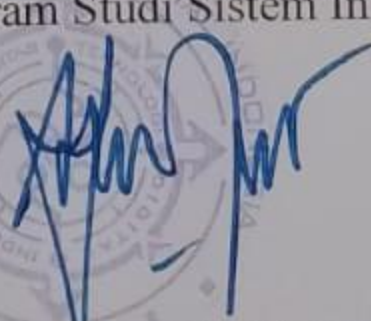
Yogyakarta

Yogyakarta, *25 Maret* 2024

| Dewan Penguji | NIDN | Tanda Tangan |
|--|------------|---|
| 1. Cosmas Haryawan, S.TP., S.Kom., M.Cs. | 0519067401 |  |
| 2. Sumiyatun, S.Kom., M.Cs. | 0515048402 |  |
| 3. Rikie Kartadie, S.T., M.Kom. | 0701037604 |  |

Mengetahui

Ketua Program Studi Sistem Informasi



Deborah Kurniawati, S.Kom., M.Cs.

NPP: 051149

HALAMAN PERNYATAAN

Dengan ini saya menyatakan bahwa naskah skripsi ini belum pernah diajukan untuk memperoleh gelar Sarjana Komputer di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 29 Februari 2023



Isni Zuliani

NIM: 205610027

HALAMAN PERSEMBAHAN

Puji syukur Alhamdulillah saya panjatkan kepada Allah SWT, yang telah memberikan karunia kepada saya sehingga dapat menyelesaikan skripsi ini yang saya persembahkan kepada:

1. Kedua orang tua, Bapak Mugiyanto dan Mamak Mistinah yang telah membesarkan saya dengan penuh kasih sayang dan cinta yang tidak mungkin dapat saya balas, dan telah memberikan kepercayaan kepada saya atas setiap langkah yang saya pilih.
2. Kakak Isnu Hidayat, yang selalu memotivasi saya bahwa semua dapat dilalui tidak ada yang tidak mungkin.
3. Segenap keluarga besar yang telah memberikan dukungan berupa bimbingan, arahan, doa maupun materi.
4. Saudari-saudari yang telah menjadi teman sekaligus saudara disini dan selalu memberikan saya nasehat, semangat serta menemani dalam suka maupun duka.
5. Dan tidak lupa untuk pembaca skripsi ini, semoga dengan tulisan saya ini dapat membantu mendapatkan referensi bermanfaat.

HALAMAN MOTTO

“Sesungguhnya Allah tidak akan mengubah nasib suatu kaum sehingga mereka mengubah keadaan yang adapada diri mereka sendiri”

(QS: Ar-Ra’ad:11)

“Jangan terlalu berekpektasi, fokus saja berjalan, tidak perlu berharap tentang apapun. Dengan begitu ketika tidak sesuai harapan kita tidak akan begitu kecewa”

(Fiersa Besari)

KATA PENGANTAR

Puji syukur penulis panjatkan kepada Allah SWT, karena berkat rahmat -Nya penulis dapat menyelesaikan skripsi dengan judul “Penerapan Metode KNN Untuk Analisis Sentimen Ulasan Pengguna Aplikasi X di *Playstore*” dengan baik. Skripsi ini dilakukan dalam rangka memenuhi salah satu syarat dalam menyelesaikan Program Studi Sistem Informasi Universitas Teknologi Digital Indonesia.

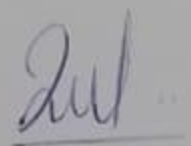
Penulis menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi sangatlah sulit bagi penulis untuk menyelesaikan skripsi ini. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Ibu Sri Redjeki, S.Si., M.Kom., PhD, selaku Rektor Universitas Teknologi Digital Indonesia
2. Ibu Deborah Kurniawati, S.Kom,M.Cs, selaku Kepala Program Studi Sistem Informasi.
3. Ibu Sumiyatun S.Kom,M.Cs, selaku Dosen Pembimbing Skripsi yang telah banyak memberikan arahan, masukan dan bimbingan pada saat penulisan skripsi ini. Semoga kebaikan beliau menjadi ladang pahala Aamiin.
4. Bapak Cosmas Haryawan S.TP., S,Kom., M.Cs. dan Bapak Rikie Kartadie S.T., M.Kom., selaku Dosen Penguji yang telah memberikan saran dan masukannya sehingga skripsi ini dapat lebih baik
5. Seluruh Staf Dosen dan Karyawan telah memberikan ilmu yang terbaik bagi penulis di dalam maupun luar kampus.

6. Kedua orang tua saya, Bapak Mugiyanto dan Mamak Mistinah yang telah memberikan segalanya tanpa kurang satu apapun serta doa yang selalu menyertai sehingga saya dapat berada dititik ini.
7. Saudari-saudari yang telah memberi semangat, arahan motivasi dan nasehat sehingga penulis dapat menyelesaikan kripsi ini, semoga apa yang kalian perbuat menjadi kebaikan untuk diri kalian sendiri Aamiin.
8. Semua pihak yang tidak dapat saya sebutkan satu persatu, telah banyak membantu dalam menyelesaikan kripsi ini baik secara langsung maupun tidak langsung. Terima kasih telah memberikan semuanya terutama waktu, kebersamaan dan canda tawanya selama ini.

Dalam penulisan skripsi ini, penulis sudah berusaha sebaik mungkin. Namun, masih banyak kesalahan dan kekurangan baik yang disengaja ataupun yang tidak disengaja. Penulis menerima kritik serta saran yang membangun sebagai perbaikan kedepannya dan semoga skripsi ini membawa manfaat bagi semua orang.

Yogyakarta, 29 Februari 2023



Penulis

DAFTAR ISI

| | |
|---|------|
| SKRIPSI..... | ii |
| HALAMAN PERSETUJUAN..... | iii |
| HALAMAN PENGESAHAN..... | iv |
| HALAMAN PERNYATAAN | v |
| HALAMAN PERSEMBAHAN | vi |
| HALAMAN MOTTO | vii |
| KATA PENGANTAR | viii |
| DAFTAR ISI..... | x |
| DAFTAR GAMBAR | xiii |
| DAFTAR TABEL..... | xiv |
| INTISARI..... | xv |
| <i>ABSTRACT</i> | xvi |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang Permasalahan | 1 |
| 1.2 Rumusan Masalah | 2 |
| 1.3 Ruang Lingkup | 2 |
| 1.4 Tujuan Penelitian..... | 3 |
| 1.5 Manfaat Penelitian..... | 3 |
| 1.6 Sistematika Penelitian | 3 |
| BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI..... | 5 |
| 2.1 Tinjauan Pustaka | 5 |
| 2.2 Landasan Teori | 8 |
| 2.2.1 Data mining..... | 8 |
| 2.2.2 K-Nearest Neighbor | 9 |
| 2.2.3 <i>Confusion Matrix</i> | 10 |
| 2.2.4 Analisis sentimen | 11 |
| 2.2.5 TF-IDF | 12 |
| 2.2.6 <i>Python</i> | 13 |
| 2.2.7 Sastrawi..... | 13 |
| 2.2.8 <i>Scikit learn</i> | 13 |
| 2.2.9 <i>Web Scrapping</i> | 14 |

| | |
|---|----|
| BAB III METODE PENELITIAN..... | 15 |
| 3.1 Bahan/Data | 15 |
| 3.1.1 Kebutuhan <i>input</i> | 15 |
| 3.1.2 Kebutuhan <i>output</i> | 15 |
| 3.2 Peralatan | 15 |
| 3.2.1 Kebutuhan Perangkat Lunak | 15 |
| 3.2.2 Kebutuhan Perangkat Keras | 16 |
| 3.3 Prosedur kerja dan pengumpulan data..... | 16 |
| 3.3.1 Analisis..... | 16 |
| 3.3.2 Identifikasi masalah | 17 |
| 3.3.3 Pengumpulan data | 17 |
| 3.3.4 <i>Pre-processing</i> | 18 |
| 3.3.5 Pelabelan Data..... | 21 |
| 3.3.6 Pembobotan Kata | 21 |
| 3.3.7 Pembagian Data | 22 |
| 3.3.8 Klasifikasi <i>K-Nearest Neighbor</i> | 22 |
| 3.3.9 Evaluasi | 23 |
| BAB IV IMPLEMENTASI DAN PEMBAHASAN | 24 |
| 4.1 Pengambilan Data..... | 24 |
| 4.2 <i>Preprocessing</i> | 25 |
| 4.2.1 <i>Cleaning</i> | 25 |
| 4.2.2 <i>Case Folding</i> | 26 |
| 4.2.3 <i>Tokenizing</i> | 27 |
| 4.2.4 <i>Normalizing</i> | 28 |
| 4.2.5 <i>Stopword Removal</i> | 29 |
| 4.2.6 <i>Stemming</i> | 30 |
| 4.3 <i>Labeling</i> | 31 |
| 4.4 Pembobotan | 32 |
| 4.5 Pembagian Data..... | 33 |
| 4.6 Klasifikasi KNN | 34 |
| BAB V PENUTUP..... | 43 |
| 32.1 Kesimpulan | 43 |

| | | |
|------|----------------------|----|
| 32.2 | Saran | 44 |
| | DAFTAR PUSTAKA | 45 |
| | LAMPIRAN | 47 |

DAFTAR GAMBAR

| | |
|--|----|
| Gambar 3. 1 Alur penelitian..... | 16 |
| Gambar 3. 2 <i>Preprocessing</i> | 18 |
| Gambar 4. 1 Kode Program <i>Scrapping</i> | 24 |
| Gambar 4. 2 Hasil <i>Scraper</i> | 25 |
| Gambar 4. 3 Kode Program <i>Cleaning</i> | 25 |
| Gambar 4. 4 Hasil <i>Cleaning</i> | 26 |
| Gambar 4. 5 Kode Program <i>Case Folding</i> | 26 |
| Gambar 4. 6 Hasil <i>Case Folding</i> | 27 |
| Gambar 4. 7 Kode Program <i>Tokenizing</i> | 27 |
| Gambar 4. 8 Hasil <i>Tokenizing</i> | 28 |
| Gambar 4. 9 Kode Program <i>Normalizing</i> | 28 |
| Gambar 4. 10 Hasil <i>Normalizing</i> | 29 |
| Gambar 4. 11 Kode Program <i>Stoword Removal</i> | 29 |
| Gambar 4. 12 Hasil <i>Stopword Removal</i> | 30 |
| Gambar 4. 13 Kode Program <i>Stemming</i> | 30 |
| Gambar 4. 14 Hasil <i>Stemming</i> | 31 |
| Gambar 4. 15 Hasil Pelabelan..... | 31 |
| Gambar 4. 16 Kode Program Pembobotan Kata..... | 32 |
| Gambar 4. 17 Hasil Pembobotan Kata..... | 33 |
| Gambar 4. 18 Kode Program Pembagian Data | 33 |
| Gambar 4. 19 Kode Program Klasifikasi KNN | 34 |
| Gambar 4. 20 Implementasi KNN Pembagian Data 90:10 | 35 |
| Gambar 4. 21 Implementasi KNN Pembagian Data 80:20 | 35 |
| Gambar 4. 22 Implementasi <i>Confusion Matrix</i> | 36 |
| Gambar 4. 23 Hasil <i>Confusion Matrix</i> Perbandingan Data 90:10 | 37 |
| Gambar 4. 24 Hasil <i>Confusion Matrix</i> Perbandingan Data 80:20 | 37 |
| Gambar 4. 25 Kode Program <i>classification report</i> | 39 |
| Gambar 4. 26 <i>Classification Report</i> Pembagian data 90:10 | 40 |
| Gambar 4. 27 <i>Classification Report</i> Pembagian data 80:20 | 40 |
| Gambar 4. 28 Kode Program ROC | 41 |
| Gambar 4. 29 Grafik <i>ROC</i> | 42 |

DAFTAR TABEL

| | |
|--|----|
| Tabel 2. 1 Penelitian Terdahulu | 7 |
| Tabel 2. 2 Tabel <i>Confusion Matrix Multi Class</i> | 10 |
| Tabel 3. 1 Hasil <i>Scrapping</i> | 17 |
| Tabel 3. 2 Contoh Hasil <i>Preprocsseing</i> | 20 |
| Tabel 3. 3 Contoh Hasil Pelabelan | 21 |
| Tabel 3. 4 Contoh Pembobotan | 22 |
| Tabel 4. 1 Pembagian Data | 34 |
| Tabel 4. 2 Perhitungan Manual <i>Precision</i> | 38 |
| Tabel 4. 3 Perhitungan Manual <i>Recall</i> | 38 |
| Tabel 4. 4 Perhitungan <i>F1-Score</i> | 39 |

INTISARI

Aplikasi media sosial *X* atau yang dahulunya disebut dengan *Twitter* merupakan sosial media yang sangat banyak digunakan oleh masyarakat Indonesia. *Google Playstore* merupakan salah satu tempat utama bagi pengguna untuk mengunduh dan mengulas sebuah aplikasi. Melakukan analisis sentimen dengan menggunakan metode KNN dapat memungkinkan pengembang aplikasi untuk dapat melihat perasaan pengguna dan mengidentifikasi permasalahan yang ada pada aplikasi tersebut.

Penelitian ini berfokus pada ulasan pengguna aplikasi *X* di *Playstore* untuk mengetahui presentase tingkat akurasi, presentase negatif, positif dan netral terhadap media sosial *X*. Ulasan yang dikumpulkan mulai tanggal 04 November 2023 sampai 15 Januari 2024. Ulasan yang telah diambil kemudian akan melalui beberapa tahap yaitu preprocessing, labeling, pembobotan dan klasifikasi KNN.

Penelitian ini menggunakan sebanyak 3000 ulasan, setelah dilakukan tahap preprocessing data menjadi 2.241. Pembagian data untuk klasifikasi menggunakan 2 perbandingan pembagian data 90:10 dan 80:10. Pembagian data 90:10 mendapatkan nilai akurasi sebesar 80,00%, pembagian data 80:20 mendapatkan nilai akurasi sebesar 80.40%. Dari hasil akurasi perbandingan 2 pembagian data, hasil akurasi terbaik didapatkan saat menggunakan data latih 80% sebanyak 1792 data dan data uji 20% sebanyak 449 data dengan nilai $K=7$ yaitu 80,40%.

Kata Kunci: *Analisis Sentimen, KNN, Python.*

ABSTRACT

The social media application X or what was previously called Twitter is a social media that is very widely used by Indonesian people. Google Playstore is one of the main places for users to download and review applications. Carrying out sentiment analysis using the KNN method can enable application developers to see user feelings and identify problems with the application.

This research focuses on user reviews of the X application on Playstore to determine the percentage of accuracy, negative, positive and neutral percentages regarding social media preprocessing, labeling, weighting and KNN classification.

This research used 3000 reviews, after carrying out the preprocessing stage the data became 2,241. Data division for classification uses 2 data division ratios 90:10 and 80:10. Data division 90:10 gets an accuracy value of 80.00%, data division 80:20 gets an accuracy value of 80.40%. From the accuracy results of comparing 2 data divisions, the best accuracy results were obtained when using 80% training data of 1792 data and 20% test data of 449 data with a value of $K=7$, namely 80.40%.

Keywords: Sentiment Analysis, KNN, Python.