

**ANALISIS SENTIMEN TERHADAP APLIKASI VIDIO
DI *GOOGLE PLAY STORE* MENGGUNAKAN
METODE *K-NEAREST NEIGHBOR* (KNN)**

SKRIPSI

**Oleh:
DAMAR WIJATI
201910225389**



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2024**

LEMBAR PERSETUJUAN PEMBIMBING
(DIGUNAKAN UNTUK TUGAS AKHIR)

Judul Tugas Akhir : Analisis Sentimen Terhadap Aplikasi Vidio
Di *Google Play Store* Menggunakan Metode
K-Nearest Neighbor (KNN)

Nama Mahasiswa : Damar Wijati

Nomor Pokok Mahasiswa : 201910225389

Program Studi / Fakultas : Informatika / Ilmu Komputer

Bekasi, 24 Januari 2024

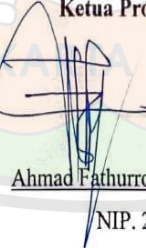
MENYETUJUI,
Pembimbing



Prima Dina Atika, S.Kom., M.Kom.

NIDN: 0311037107

Ketua Program Studi



Ahmad Fathurrozi, S.E., M.M.S.I.

NIP. 2012486

Program Studi Informatika
Fakultas Ilmu Komputer
Universitas Bhayangkara Jakarta Raya
2024

LEMBAR PENGESAHAN

Judul Tugas Akhir : ANALISIS SENTIMEN TERHADAP
APLIKASI VIDIO DI *GOOGLE PLAY*
STORE MENGGUNAKAN METODE *K-NEAREST NEIGHBOR* (KNN)

Nama Mahasiswa : Damar Wijati

Nomor Pokok Mahasiswa : 201910225389

Program Studi/Fakultas : Informatika/Illmu Komputer

Tanggal Lulus Ujian : 12 Februari 2024

Tugas Akhir

Jakarta, 12 Februari 2024

MENGESAHKAN,

Ketua Tim Penguji : Dani Yusuf, S.Kom., M.Kom.

NIDN : 0330067003

Penguji I : Mukhlis, S.Kom., M.T.

NIDN : 0312116802

Penguji II : Prima Dina Atika, S.Kom., M.Kom.

NIDN : 0311037107

MENGETAHUI,

Ketua

Program Studi Informatika

Ahmad Fathurrozi, S.E., M.M.S.I

NIP. 2012486

Dekan

Fakultas Ilmu Komputer

Dr. Dra. Tyastuti Sri Lestari, M.M

NIP. 1408206



UNIVERSITAS BHAYANGKARA JAKARTA RAYA
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI INFORMATIKA

LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Damar Wijati
NPM : 201910225389
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Terhadap Aplikasi Vidio di Google Play Store Menggunakan Metode *K-Nearest Neighbor* (KNN)

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Bekasi, 22 Januari 2024

Penulis



Damar Wijati

ABSTRAK

Damar Wijati . 201910225389 . analisis sentimen terhadap aplikasi vidio di google play store menggunakan metode *k-nearest neighbor* (knn). Bekasi: Fakultas Ilmu Komputer. Universitas Bhayangkara Jakarta Raya, 2024

Vidio merupakan salah satu media *over-the-top* (OTT) yang menyediakan berbagai konten hiburan seperti film, acara TV, acara olahraga, acara musik, program gaya hidup, dan lain-lain dalam satu *platform*. Aplikasi Vidio memberikan kesempatan kepada pengguna untuk memberikan ulasan mengenai kepuasan mereka dalam menggunakan aplikasi tersebut. Maka dari itu, penelitian ini dilakukan untuk menganalisis sentimen pengguna terhadap aplikasi Vidio dalam *platform Google Play Store* menggunakan metode *K-Nearest Neighbor* (KNN). Data untuk analisis sentimen diambil secara acak dari aplikasi Vidio berdasarkan ulasan yang paling relevan. Sebanyak 3.000 data dianalisis, dengan jumlah kelas negatif sebanyak 2238 data, kelas netral sebanyak 508 data serta kelas positif sebanyak 254 data. Penelitian ini menggunakan metode *K-Nearest Neighbor* (KNN) untuk pengklasifikasian ulasan berdasarkan kelas negatif, netral dan positif serta *Multiclass Confusion Matrix* untuk evaluasi model. Dengan pembagian data sebanyak 70% *data training* dan 30% *data testing* serta jumlah *n_neighbors* sebanyak 10 data, menghasilkan akurasi sebesar 81,6%, presisi sebesar 79%, *recall* sebesar 81.6% dan *F1-Score* sebesar 77%.

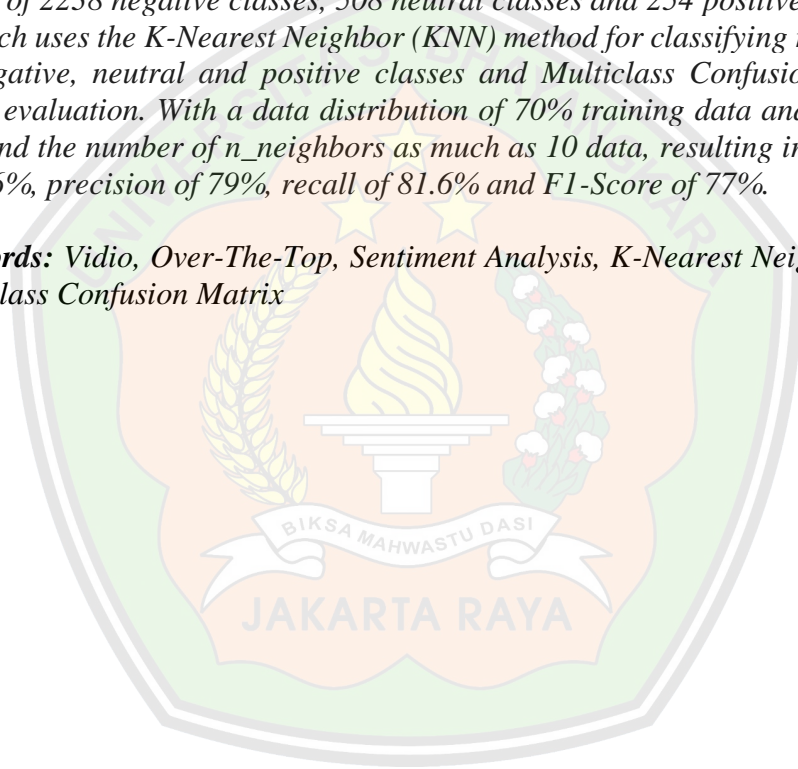
Kata kunci: Vidio, *Over-The-Top*, Analisis Sentimen, *K-Nearest Neighbor* (KNN), *Multiclass Confusion Matrix*

ABSTRACT

Damar Wijati . 201910225389 . sentiment analysis of vidio applications in the google play store using the k-nearest neighbor (knn) method. Bekasi: Faculty of Computer Science. Jakarta Bhayangkara University, 2024

Vidio is one of the over-the-top (OTT) media that provides various entertainment content such as movies, TV shows, sporting events, music events, lifestyle programs, and others in one platform. The Vidio application provides an opportunity for users to provide reviews regarding their satisfaction in using the application. Therefore, this research was conducted to analyze user sentiment towards the Vidio application on the Google Play Store platform using the K-Nearest Neighbor (KNN) method. Data for sentiment analysis was randomly drawn from the Vidio app based on the most relevant reviews. A total of 3,000 data were analyzed, with a total of 2238 negative classes, 508 neutral classes and 254 positive classes. This research uses the K-Nearest Neighbor (KNN) method for classifying reviews based on negative, neutral and positive classes and Multiclass Confusion Matrix for model evaluation. With a data distribution of 70% training data and 30% testing data and the number of n_neighbors as much as 10 data, resulting in an accuracy of 81.6%, precision of 79%, recall of 81.6% and F1-Score of 77%.

Keywords: *Vidio, Over-The-Top, Sentiment Analysis, K-Nearest Neighbor (KNN), Multiclass Confusion Matrix*



**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK**

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Damar Wijati
NPM : 201910225389
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (Non-Exclusive Royalty-Free Right)**, atas karya ilmiah saya yang berjudul :

Analisis Sentimen Terhadap Aplikasi Vidio di Google Play Store Menggunakan Metode K-Nearest Neighbor (KNN)

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-ekklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Bekasi
Pada tanggal : 22 Januari 2024
Yang Menyatakan



Damar Wijati

KATA PENGANTAR

Dengan rasa syukur yang mendalam kepada Allah SWT., yang telah melimpahkan rahmat, taufik, dan petunjuk-Nya, penulis berhasil menyelesaikan skripsi ini dengan judul “**Analisis Sentimen Terhadap Aplikasi Vidio Di Google Play Store Menggunakan Metode K-Nearest Neighbor (KNN)**”, yang disusun sebagai salah satu persyaratan untuk menyelesaikan Program Sarjana (S1) Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.

Penulis menyadari bahwa skripsi ini tidak akan selesai tanpa adanya dukungan, bantuan, bimbingan, maupun wejangan dari berbagai pihak selama penyusunan skripsi ini. Pada kesempatan ini, penulis menyampaikan terima kasih yang sebanyak-banyaknya, terutama kepada kedua orang tua penulis, bapak Janji dan ibu Suridah, serta beberapa pihak yang terlibat didalamnya, yaitu:

1. Bapak Irjen Pol (Prof) Dr. Drs. Bambang Karsono, S.H., M.M., selaku Rektor Universitas Bhayangkara Jakarta Raya.
2. Ibu Dr. Dra. Tyastuti Sri Lestari, M.M., selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
3. Bapak Ahmad Fathurrozi, S.E., M.M.S.I., selaku Ketua Program Studi Informatika Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
4. Ibu Prima Dina Atika, S.Kom., M.Kom., selaku Dosen Pembimbing Skripsi atas segala bimbingan, arahan, serta saran yang diberikan kepada penulis sehingga skripsi ini dapat diselesaikan dengan baik.
5. Bapak Joni Warta, S.Si., M.Si., selaku Dosen Pembimbing Akademik yang telah membantu penulis dalam mengikuti dan menyelesaikan studi di Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
6. Seluruh staff pengajar Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya, yang sudah memberikan ilmu pengetahuan dan pengalaman yang tak ternilai selama penulis menempuh pendidikan di Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
7. Kakak yang penulis sayangi, Siwi Pramesti, yang sudah memberikan *support* kepada penulis dalam hal pembiayaan perkuliahan selama penulis menempuh

pendidikan di Ubhara Jaya. Terima kasih atas perhatiannya selama ini, mohon maaf jika adikmu ini selalu merepotkan dirimu.

8. Teman-teman penulis, Dimas, Zacky, dan Jumi. Terima kasih atas bantuan, motivasi, saran, diskusi, serta kerja samanya dan terima kasih juga karena sudah mau menjadi teman yang terbaik bagi penulis selama ini.
9. Seluruh teman-teman penulis di Fakultas Ilmu Komputer angkatan 2019 yang tidak bisa disebutkan namanya satu per satu, terima kasih atas pertemanannya selama ini.
10. Teruntuk diri sendiri, terima kasih sudah mau bertahan sampai sejauh ini dan selamat telah menyelesaikan perjalanan ini. Setiap halaman adalah bukti tekad dan cinta pada ilmu. Teruslah berkembang dan jadilah pribadi yang lebih baik.
11. Semua pihak yang tidak bisa disebutkan namanya satu per satu yang sudah memberikan doa dan motivasi kepada penulis, sehingga penulis dapat menyelesaikan skripsi ini dengan baik.

Dalam penulisan skripsi ini mungkin masih banyak terdapat kekurangan dalam penulisannya, maka dari itu segala kritik dan saran yang membangun akan menyempurnakan penulisan skripsi ini serta bermanfaat bagi penulis dan para pembaca.

Bekasi, 24 Januari 2024

Penulis,



Damar Wijati

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PENGESAHAN	iii
LEMBAR PERNYATAAN BUKAN PLAGIASI	iv
ABSTRAK	v
ABSTRACT	vi
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI	vii
KATA PENGANTAR	viii
DAFTAR ISI	x
DAFTAR TABEL	xiii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	5
1.3 Rumusan Masalah	5
1.4 Tujuan Penelitian	5
1.5 Manfaat Penelitian	6
1.6 Batasan Masalah	6
1.7 Sistematika Penulisan	7
BAB II TINJAUAN PUSTAKA	9
2.1 Penelitian Terdahulu	9
2.2 Landasan Teori	20
2.2.1 Aplikasi Vidio	20
2.2.2 <i>Data Mining</i>	21
2.2.3 <i>Text Mining</i>	23
2.2.4 Klasifikasi	24
2.2.5 <i>K-Nearest Neighbor (KNN)</i>	25
2.2.6 Analisis Sentimen	30
2.2.7 TF-IDF	31
2.2.8 <i>Split Validation</i>	32

2.2.9 Evaluasi	33
BAB III METODOLOGI PENELITIAN	37
3.1 Tipe Penelitian.....	37
3.2 Lokasi dan Waktu Penelitian	37
3.3 Teknik Analisis Data	38
3.4 Metode Pengumpulan Data	38
3.5 Subjek dan Objek Penelitian	39
3.6 Populasi dan Sampel	39
3.7 Jenis Data	40
3.8 Analisis Sistem Berjalan	40
3.9 Analisis Sistem yang Diusulkan	41
3.10 Analisis Kebutuhan Sistem	42
3.10.1 Analisis Kebutuhan Perangkat Keras (<i>Hardware</i>)	43
3.10.2 Analisis Kebutuhan Perangkat Lunak (<i>Software</i>)	43
3.11 Kerangka Pemikiran.....	44
BAB IV HASIL DAN PEMBAHASAN	46
4.1 Pengambilan Data	46
4.2 Pelabelan Data	49
4.3 <i>Text Preprocessing</i>	52
4.3.1 <i>Case Folding</i>	53
4.3.2 <i>Cleaning</i>	55
4.3.3 <i>Tokenizing</i>	57
4.3.4 <i>Stopwords Removal</i>	59
4.3.5 <i>Stemming</i>	61
4.4 TF-IDF	66
4.5 Klasifikasi <i>K-Nearest Neighbor</i> (KNN)	76
4.6 Pengujian Model	79
4.7 Evaluasi Model	80
BAB V PENUTUP	91
5.1 Kesimpulan	91
5.2 Saran	91
DAFTAR PUSTAKA	93



DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	12
Tabel 2.2 Tinjauan Pustaka	15
Tabel 2.3 Data Makanan	27
Tabel 2.4 Pencarian 3 Data Dengan Jarak Terdekat	28
Tabel 2.5 Tiga Data Dengan Jarak Terdekat	29
Tabel 2.6 Klasifikasi Data Baru	30
Tabel 2.7 <i>Confusion Matrix</i>	33
Tabel 2.8 <i>Multiclass Confusion Matrix 3*3</i>	36
Tabel 4.1 Hasil <i>Web Scraping</i>	49
Tabel 4.2 Hasil Pelabelan Data	51
Tabel 4.3 Hasil <i>Case Folding</i>	54
Tabel 4.4 Hasil <i>Cleaning</i>	56
Tabel 4.5 Hasil <i>Tokenizing</i>	58
Tabel 4.6 Hasil <i>Stopwords Removal</i>	60
Tabel 4.7 Hasil <i>Stemming</i>	64
Tabel 4.8 Hasil Perhitungan TF	71
Tabel 4.9 Hasil Perhitungan DF	72
Tabel 4.10 Hasil Perhitungan IDF	74
Tabel 4.11 Hasil Perhitungan TF-IDF	75
Tabel 4.12 <i>Multiclass Confusion Matrix 3*3</i>	79
Tabel 4.13 Hasil <i>Multiclass Confusion Matrix</i>	81
Tabel 4.14 Hasil Kalkulasi Nilai Kelas Negatif	82
Tabel 4.15 Hasil Kalkulasi Nilai Kelas Netral	83
Tabel 4.16 Hasil Kalkulasi Nilai Kelas Positif	85
Tabel 4.17 Hasil Kalkulasi Performa setiap Kelas	86

DAFTAR GAMBAR

Gambar 1.1 Grafik Jumlah Pengguna Aplikasi Hiburan di Indonesia	4
Gambar 3.1 <i>Flowchart</i> Analisis Sistem Berjalan	41
Gambar 3.2 <i>Flowchart</i> Analisis Sistem yang Diusulkan	42
Gambar 3.3 Kerangka Pemikiran “Analisis Sentimen Terhadap Aplikasi Vidio Di <i>Google Play Store</i> Menggunakan Metode <i>K-Nearest Neighbor</i> (KNN)”	44
Gambar 4.1 <i>Flowchart</i> Pengambilan Data	48
Gambar 4.2 <i>Flowchart</i> Pelabelan Data	50
Gambar 4.3 <i>Flowchart Text Preprocessing</i>	65
Gambar 4.4 Hasil Keseluruhan Sentimen setelah <i>Text Preprocessing</i>	65
Gambar 4.5 Hasil Pembobotan Kata dengan TF-IDF	68
Gambar 4.6 <i>Flowchart</i> TF-IDF	69
Gambar 4.7 Hasil Pengujian Model	80
Gambar 4.8 Hasil Perhitungan Evaluasi Model KNN	88
Gambar 4.9 <i>Flowchart</i> Klasifikasi KNN dan Evaluasi Model	90

DAFTAR LAMPIRAN

Lampiran 1 Data Primer Kasar (Ulasan Aplikasi Vidio)	97
Lampiran 2 Lembar Validasi Instrumen	99
Lampiran 3 Surat Pernyataan Validitas Instrumen	108
Lampiran 4 Lembar Evaluasi Seminar Proposal	109
Lampiran 5 Surat Rekomendasi dari Pembimbing	110
Lampiran 6 Lembar Evaluasi Sidang Skripsi	111
Lampiran 7 Lembar Persetujuan Pembimbing (Proposal Tugas Akhir)	113
Lampiran 8 Kartu Bimbingan Proposal Skripsi	114

