

**ANALISIS SENTIMEN OPINI MASYARAKAT TERHADAP
PEMBELAJARAN DARING ERA PANDEMI DARI *TWITTER*
MENGUNAKAN ALGORITMA *K-NEAREST NEIGHBOR (K-NN)*
TUGAS AKHIR**

NO.970/WM.FT.H6/T.ILKOM/TA/2023

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh Gelar Sarjana
Komputer**



Oleh :

THERESIA EKAWATI INDRIANI

23119016

**PROGRAM STUDI ILMU KOMPUTER
FAKULTAS TEKNIK
UNIVERSITAS KATOLIK WIDYA MANDIRA
KUPANG
2023**

HALAMAN PERSETUJUAN

TUGAS AKHIR

NO. 970/WM.FT.H6/T.ILKOM/TA/2023

**ANALISIS SENTIMEN OPINI MASYARAKAT TERHADAP
PEMBELAJARAN DARING ERA PANDEMI DARI *TWITTER*
MENGUNAKAN ALGORITMA *K-NEAREST NEIGHBOR (K-NN)***

Oleh :

THERESIA EKAWATI INDRIANI


23119016

TELAH DIPERIKSA/DISETUJUI OLEH PENGUJI:

DI : KOTA KUPANG

PADA : JULI 2023

DOSEN PENGUJI I


Paskalis A. Nani, S.T., M.T
NIDN : 0831038602

DOSEN PENGUJI II

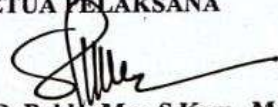

Yovinia C. Hoar Siki, S.T., M.T
NIDN : 0805058803

DOSEN PENGUJI III

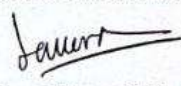

Sisilia D. Bakka Mau, S.Kom., M.T

NIDN : 0807098502

KETUA PELAKSANA


Sisilia D. Bakka Mau, S.Kom., M.T
NIDN : 0807098502

SEKRETARIS PELAKSANA


Emerensiana Ngaga, S.T., M.T
NIDN : 0802038601

HALAMAN PENGESAHAN

TUGAS AKHIR

NO. 970/WM.FT.H6/T.ILKOM/TA/2023

**ANALISIS SENTIMEN OPINI MASYARAKAT TERHADAP
PEMBELAJARAN DARING ERA PANDEMI DARI *TWITTER*
MENGUNAKAN ALGORITMA *K-NEAREST NEIGHBOR (K-NN)***

Oleh :

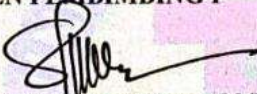
THERESIA EKAWATI INDRIANI

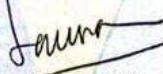
23119016

TELAH DIPERTAHANKAN DI DEPAN PEMBIMBING:

DOSEN PEMBIMBING I

DOSEN PEMBIMBING II


Sisilia D. Bakka Mau, S.Kom., M.T
NIDN : 0807098502


Emerensiana Ngaga, S.T., M.T
NIDN : 0802038601

**MENGETAHUI,
KETUA PROGRAM STUDI
ILMU KOMPUTER
UNIKA WIDYA MANDIRA**

**MEGESAHKAN,
DEKAN FAKULTAS TEKNIK
UNIKA WIDYA MANDIRA**


Sisilia D. Bakka Mau, S.Kom., M.T
NIDN : 0807098502


Dr. Don Gaspar Noesaku Da Costa, S.T., M.T
NIDN : 0820036801

HALAMAN PERSEMBAHAN

Karya ini secara khusus saya persembahkan untuk :

TUHAN YESUS KRISTUS,

Bapa Maximus Daur, Mama Titin Sulastri, Mama Susana Mues, Kakak Avelin Jedia,

Kakak-Kakak dan Adik-Adikku tersayang,

serta seluruh keluarga yang selalu mendukung saya.

Semua sahabat-sahabat yang selalu membantu dan mendukung saya,

dan juga teman-teman angkatan 2019.

Diri saya sendiri yang sudah berjuang dan telah menyelesaikan penulisan skripsi ini

Almamater UNWIRA tercinta

MOTTO

Tuhan tak'kan terlambat !

Juga tak akan lebih cepat

Semuanya

Dia jadikan indah tepat pada waktuNya.

(Pengkotbah 3 : 11a)

PERNYATAAN DAN KEASLIAN HASIL KARYA

Yang bertanda tangan dibawah ini:

Nama : Theresia Ekawati Indriani

No. Registrasi : 23119016

Fakultas : Teknik

Program Studi : Ilmu Komputer

Dengan ini menyatakan bahwa Tugas Akhir dengan judul “Analisis Sentimen Opini Masyarakat Terhadap Pembelajaran Daring Era Pandemi Dari *Twitter* Menggunakan Algoritma *K-Nearest Neighbor (K-NN)*” adalah benar-benar hasil penelitian yang dilakukan oleh penulis sendiri.

Disahkan/Diketahui,

Pembimbing I



Sisilia D. Bakka Mau, S.Kom., M.T

Kupang, Juli 2023

Mahasiswa



Theresia Ekawati Indriani

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan yang Maha Esa karena atas berkat dan kasih-Nya yang melimpah penulis dapat menyelesaikan Skripsi yang berjudul “ **Analisis Sentimen Opini Masyarakat Terhadap Pembelajaran Daring Era Pandemi Dari Twitter Menggunakan Algoritma *K-Nearest Neighbor (K-NN)***” dengan baik. Dalam menyelesaikan Skripsi ini, penulis telah banyak mendapat bantuan dari berbagai pihak yang telah membantu dengan caranya masing-masing baik dukungan moral maupun material. Oleh karena itu pada kesempatan berharga ini dengan penuh rasa hormat dan kerendahan hati penulis mengucapkan terima kasih kepada :

1. Allah Tri Tunggal Maha Kudus bersama Bunda Maria yang adalah sumber segala rahmat, kehidupan dan pengetahuan.
2. Pater Dr.Philipus Tule, SVD, selaku Rektor Universitas Katolik Widya Mandira.
3. Bapak Dr. Don Gaspar Noesaku da costa, S.T., M.T. selaku Dekan Fakultas Teknik Universitas Katolik Widya Mandira.
4. Ibu Sisilia D. Bakka Mau, S.Kom., M.T., selaku Ketua Program Studi Ilmu Komputer Universitas Katolik Widya Mandira.
5. Ibu Sisilia D. Bakka Mau, S.Kom., M.T., selaku Dosen Pembimbing I dan Ibu Emerensiana Ngaga, S.T., M.T., selaku Dosen Pembimbing II, terimakasih untuk kesabaran, arahan, nasehat, petunjuk, serta waktu yang telah dicurahkan bagi saya selama proses bimbingan skripsi ini.

6. Bapak Paskalis A. Nani, S.T., M.T., selaku Dosen Penguji I dan Ibu Yovinia C. Hoar Siki, S.T., M.T., selaku penguji II yang telah meluangkan waktu, tenaga dan pikiran untuk menguji dan membimbing dalam perbaikan skripsi ini.
7. Ibu Emerensiana Ngaga, S.T., M.T., selaku Dosen Pembimbing Akademik yang selalu memberikan motivasi dan dorongan.
8. Seluruh dosen dan staf karyawan pada Program Studi Ilmu Komputer Universitas Katolik Widya Mandira.
9. Keluarga tercinta Bapak Maximus, Mama Titin, Mama Susan Mues, juga kakak (Velin Jedia, Febi, Yungles, Kory) dan Adik (Kristi, Teo) atas segala bentuk dukungan, doa kasih sayang dan motivasi yang diberikan.
10. Teman-teman dan sahabat terbaik Desri, Mega, Nanci, Novi, Nunik, Ronal, Atay. Serta teman-teman kos (RDD) dan Anatha Doko yang telah mendoakan dan memberikan dukungan kepada penulis selama perkuliahan.

Penulis menyadari bahwa skripsi ini belum sempurna. Oleh karena itu, segala kritik dan saran serta masukan guna penyempurnaan skripsi ini sangat penulis harapkan. Akhir kata penulis menyampaikan semoga skripsi ini dapat bermanfaat bagi pembaca khususnya mahasiswa Ilmu Komputer Fakultas Teknik UNWIRA.

Kupang, Juli 2023

DAFTAR ISI

HALAMAN PERSETUJUAN	i
HALAMAN PENGESAHAN.....	ii
HALAMAN PERSEMBAHAN	iii
MOTTO	iv
PERNYATAAN DAN KEASLIAN HASIL KARYA	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	viii
DAFTAR TABEL	xi
DAFTAR GAMBAR.....	xii
ABSTRAK	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4

1.6	Sistematika Penulisan	4
BAB II LANDASAN TEORI		6
2.1	Penelitian Terdahulu	6
2.2	Teori Penunjang	12
2.3	Klasifikasi	16
2.4	Algoritma <i>K-Nearest Neighbor (K-NN)</i>	16
BAB III METODOLOGI PENELITIAN		19
3.1	Rancangan Penelitian	19
3.2	Perancangan Sistem	22
BAB IV HASIL DAN PEMBAHASAN		24
4.1	Pembahasan	24
4.1.1	Analisa Data	24
4.1.2	<i>Representasi Data</i>	29
4.1.2.1	<i>Preprocessing Data</i>	29
4.1.3	Hasil Analisis Data	33
4.1.3.1	<i>Klasifikasi K-Nearest Neighbor</i>	33
4.1.3.2	<i>Confusion Matrix</i>	39
4.2	Hasil	42

BAB V PENUTUP	46
5.1 Kesimpulan	46
5.2 Saran	47
DAFTAR PUSTAKA	48

DAFTAR TABEL

Tabel 2. 1 Perbandingan Terhadap Penelitian Sebelumnya.....	8
Tabel 4. 1 <i>Stopwords</i>	37
Tabel 4. 2 <i>Confusion Matrix</i>	40
Tabel 4. 3 Contoh Perhitungan <i>Confusion Matrix</i>	41
Tabel 4. 4 Daftar K dan <i>Number of Folds</i>	42

DAFTAR GAMBAR

Gambar 2. 1 Contoh <i>KNN Classification</i>	17
Gambar 3. 1 Kerangka Penelitian	19
Gambar 3. 2 Flowchart Sistem.....	22
Gambar 4. 1 Alur Sistem <i>Crawling Data Twitter</i>	24
Gambar 4. 2 Tampilan akun <i>API Twitter</i>	25
Gambar 4. 3 Tampilan <i>Bearer Token</i>	26
Gambar 4. 4 Tampilan <i>Bearer Token</i> di <i>Orange</i>	27
Gambar 4. 5 Tampilan <i>Crawling</i> data pada aplikasi <i>Orange</i>	27
Gambar 4. 6 Kumpulan Data <i>Crawling</i>	28
Gambar 4. 7 Widget <i>Impute</i>	29
Gambar 4. 8 Alur Proses <i>Cleaning</i>	30
Gambar 4. 9 Hasil <i>Preprocessing Cleaning</i>	30
Gambar 4. 10 Hasil <i>Preprocessing Case Folding</i>	31
Gambar 4. 11 Hasil <i>Preprocessing Tokenizing</i>	32
Gambar 4. 12 Menentukan Nilai Parameter <i>k</i>	33
Gambar 4. 13 Data <i>Training</i>	34
Gambar 4. 14 Data <i>Testing</i>	34
Gambar 4. 15 Alur Pelabelan Data <i>Testing</i>	35
Gambar 4. 16 <i>Testing Data</i>	36
Gambar 4. 17 Visualisasi Pada <i>Orange</i>	37
Gambar 4. 18 <i>Wordcloud</i>	38

Gambar 4. 19 Hasil *Confusion Matrix* 44

Gambar 4. 20 Nilai Akurasi, Presisi, *Recall*, *F1-Score* 45

ABSTRAK

Penyebaran virus *Covid-19* di Indonesia pada tahun 2019 menyebabkan 463.000 orang terkonfirmasi positif dan korban meninggal mencapai 15.148 orang. Pembelajaran daring menjadi *trending topic* di *twitter* karena banyak masyarakat yang memberikan opini terkait pembelajaran daring ini. Kumpulan opini pada *twitter* dapat dianalisis untuk mengetahui sentimen negatif, positif dan netral yang terkandung dalam opini masyarakat. Penelitian ini menggunakan metode *K-Nearest Neighbor (K-NN)* untuk menganalisis sentimen. Data penelitian merupakan kumpulan data tweet berbahasa Indonesia sebanyak 50.000 data dan setelah di-*preprocessing* data tersebut menjadi 9.000 data yang diambil pada tanggal 17 November 2022 – 14 Desember 2022 dengan memanfaatkan *API Twitter* dan menggunakan kata kunci “Pembelajaran Online” dan “Daring”. Analisis sentimen hasil penerapan klasifikasinya pada data opini masyarakat terhadap pembelajaran daring pada media sosial *twitter* bahwa sentimen positif lebih unggul berjumlah 4055 dibandingkan kelas sentimen negatif berjumlah 3264 dan netral berjumlah 1681. Dengan melihat hasil dari *confusion matrix*, didapatkan nilai akurasi, *presisi*, *recall*, dan *f-1 score*. Berdasarkan nilai $K=95$ dan *number of folds* 20 didapatkan nilai *accuracy* sebesar 95.5%, *precision* sebesar 92%, *recall* sebesar 92% dan *f1-score* sebesar 91.9%.

Kata Kunci : Sentimen Analisis, *K-Nearest Neighbor (K-NN)*, Pembelajaran Daring.

ABSTRACT

The spread of the Covid-19 virus in Indonesia in 2019 resulted in 463,000 people being confirmed positive and the death toll reaching 15,148 people. Online learning is a trending topic on Twitter because many people give their opinions regarding online learning. A collection of opinions on Twitter can be analyzed to find out the negative, positive and neutral sentiments contained in public opinion. This study uses the K-Nearest Neighbor (K-NN) method to analyze sentiment. The research data is a collection of 50,000 Indonesian language tweet data and after preprocessing the data becomes 9,000 data taken on November 17 2022 - December 14 2022 by utilizing the Twitter API and using the keywords "Online Learning" and "Online". Sentiment analysis results from applying the classification to public opinion data on online learning on Twitter social media that positive sentiments are superior to 4055 compared to 3264 negative sentiment classes and 1681 neutral sentiments. By looking at the results of the confusion matrix, the values for accuracy, precision, recall, and f1 scores. Based on the value of $K = 95$ and the number of folds 20, the accuracy value is 95.5%, precision is 92%, recall is 92% and f1-score is 91.9%.

Keywords: Sentiment Analysis, K-Nearest Neighbor (K-NN), Online Learning.