

## PERNYATAAN KEASLIAN TULISAN

Saya menyatakan bahwa skripsi ini telah dibuat dengan tingkat plagiasi yang minimal dan sesuai standar akademik yang berlaku, apabila di kemudian hari terbukti terdapat plagiat dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan peraturan perundang-undangan.

Bojonegoro, 29 Juni 2024



Vittri Witdyya Ningsih

NIM : 2120200531

## HALAMAN PERSETUJUAN

Nama : Vittri Witdyya Ningsih

NIM : 2120200531

Judul : Analisis Sentimen Ulasan Pengguna *Google Maps* Terhadap Rumah Sakit di Bojonegoro Menggunakan Metode *K-Nearest Neighbor*

Telah disetujui dan dinyatakan memenuhi syarat untuk diajukan dalam ujian skripsi.

Bojonegoro, 28 Juni 2024

Pembimbing I



**Muhammad Jauhar Vikri, M.Kom.**

NIDN : 0712078803

Pembimbing II



**Sunu Wahyudhi, M.Pd.**

NIDN : 0709058902

## LEMBAR PENGESAHAN

Nama : Vitri Witdyia Ningsih

NIM : 2120200531

Judul : Analisis Sentimen Ulasan Pengguna *Google Maps* Terhadap Rumah Sakit di Bojonegoro Menggunakan Metode *K-Nearest Neighbor*

Telah dipertahankan di hadapan penguji pada tanggal 06 Juli 2024.

Dewan Penguji

Penguji I



Guruh Putro Dingantoro, M. Kom.

NIDN : 0722049201

Tim Pembimbing

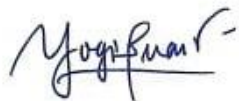
Pembimbing I



Muhammad Jauhar Vikri, M.Kom.

NIDN : 0712078803

Penguji II



Dr. H. Yogi Prana Izza, Lc., M.A.

NIDN : 0731127601

Pembimbing II



Sunu Wahvudhi, M.Pd.

NIDN : 0709058902

Mengetahui,

Dekan Fakultas Sains dan Teknologi



Muhammad Jauhar Vikri, M.Kom.

NIDN : 0712078803

Mengetahui,

Ketua Program Studi



Ucta Pradema Sanjaya, M.Kom.

NIDN : 0729128903

## MOTTO DAN PERSEMBAHAN

### MOTTO

Jangan pernah menyerah, karena sesungguhnya kita tidak tahu seberapa dekat kita dengan kesuksesan.

~ Vitri Witdyya Ningsih ~

*Being Thankful and Giving Thanks is One of The Keys to be Happy*

~ Mark Lee ~

### PERSEMBAHAN

Persembahan tugas akhir ini dan rasa terima kasih saya ucapkan untuk :

1. Allah SWT yang telah melancarkan semua proses pembuatan tugas akhir ini.
2. Keluarga yang cintanya tak terbatas, Bapak Sadar dan Ibu Tasmi yang telah memberikan doa dan dukungan sehingga penulis dapat melewati segala proses perkuliahan serta pengerjaan tugas akhir.
3. Sahabat-sahabat saya yang luar biasa, terimakasih telah bersedia berbagi ilmu, mengulurkan bantuan, mendengarkan keluh-kesah, dan kebersamai saya selama pengerjaan tugas akhir.
4. Semua pihak yang terlibat dalam pembuatan maupun proses pengerjaan tugas akhir ini.
5. Terima kasih juga kepada diri sendiri yang telah mampu berusaha dan berjuang sejauh ini.

## KATA PENGANTAR

Puji syukur atas kehadiran Allah SWT yang telah memberikan rahmat dan hidayah-Nya sehingga penulis menyelesaikan skripsi yang berjudul “Analisis Sentimen Ulasan Pengguna *Google Maps* Terhadap Rumah Sakit di Bojonegoro Menggunakan Metode *K-Nearest Neighbor*”. Skripsi ini merupakan salah satu syarat untuk memperoleh gelar Sarjana (S.Kom) pada Program Studi Teknik Informatika Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.

Dalam penulisan proposal skripsi ini, penulis mendapatkan banyak motivasi, inspirasi, dukungan, dan bimbingan dari berbagai pihak sehingga penulis dapat menyelesaikan skripsi ini tepat waktu. Penulis ingin menggunakan kesempatan ini untuk mengucapkan terima kasih kepada :

1. Bapak M. Jauharul Ma'arif, M. Pd. I. selaku Rektor Universitas Nahdlatul Ulama Sunan Giri Bojonegoro.
2. Bapak Muhammad Jauhar Vikri, M. Kom. selaku Dekan Fakultas Sains dan Teknologi Universitas Nahdlatul Ulama Sunan Giri Bojonegoro dan selaku Dosen Pembimbing I yang senantiasa membimbing penulis.
3. Bapak Ucta Pradema Sanjaya, M. Kom. selaku Kaprodi Teknik Informatika Universitas Nahdlatul Ulama Sunan Giri Bojonegoro dan selaku Dosen Pembimbing Akademik yang senantiasa membimbing penulis.
4. Bapak Sunu Wahyudhi, M.Pd. selaku Dosen Pembimbing 2 yang selalu membimbing penulis dan memberi masukan dalam penulisan skripsi ini.

Penulis menyadari, bahwa skripsi ini masih banyak kekurangan dan belum sempurna. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan agar penulis dapat menyempurnakan skripsi ini. Penulis berharap skripsi ini dapat bermanfaat bagi pembaca dan berbagai pihak.

Bojonegoro, 28 Juni 2024

Penulis

## **ABSTRACT**

*Ningsih, Vitri Witdyya. Sentiment Analysis of Google Maps User Reviews of Hospitals in Bojonegoro Using the K-Nearest Neighbor Method. Thesis. Department of Informatics Engineering, Faculty of Science and Technology, Sunan Giri Nahdlatul Ulama University. First Supervisor Muhammad Jauhar Vikri, M.Kom. and Second Supervisor Sunu Wahyudhi, M.Pd.*

*This research was conducted with the aim of analyzing Google Maps user reviews of hospitals in Bojonegoro using the K-Nearest Neighbor (KNN) method. The sentiment classes used were positive and negative. The dataset used in the study consisted of 754 review data points. The research process consisted of several stages: data scraping, performed with the web scraper extension on Chrome called Instant Data Scraper; data preprocessing, which included case folding, cleaning, tokenizing, stopword removal, and stemming; word weighting using Term Frequency-Inverse Document Frequency (TF-IDF); and the creation of a K-Nearest Neighbor classification model with (  $k$  ) values of 5, 7, and 9. This study used K-Fold Cross Validation with a scenario of (  $k = 10$  ) for validation. The model performance evaluation process was carried out using a Confusion Matrix. The results of this study showed that the highest accuracy was achieved with (  $k = 5$  ), with an accuracy of 83%, precision of 97%, and recall of 83%.*

**UNUGIRI**

## ABSTRAK

Ningsih, Vitri Witdyya. Analisis Sentimen Ulasan Pengguna *Google Maps* Terhadap Rumah Sakit di Bojonegoro Menggunakan Metode *K-Nearest Neighbor*. Skripsi. Jurusan Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Nahdlatul Ulama Sunan Giri. Pemimbing I Muhammad Jauhar Vikri, M.Kom. dan Pembimbing II Sunu Wahyudhi, M.Pd.

Penelitian ini dilakukan dengan tujuan untuk melakukan analisis terhadap ulasan pengguna google maps terhadap rumah sakit di Bojonegoro menggunakan metode K-Nearest Neighbor (KNN). Kelas sentimen yang digunakan yaitu kelas positif dan negatif. Dataset yang digunakan dalam penelitian berjumlah 754 data ulasan. Proses penelitian terdiri dari beberapa tahap, yaitu scraping data yang dilakukan dengan ekstensi *web scraper* yang terdapat pada Chrome yaitu *Instan Data Scraper*. Kemudian *Preprocessing data* yang terdiri dari *Case Folding*, *Cleaning*, *Tokenizing*, *Stopword Removal*, dan *Stemming*. Selanjutnya dilakukan pembobotan kata menggunakan *Term Frequency-Inverse Document Frequency* (TF-IDF). Lalu pembuatan model klasifikasi *K-Nearest Neighbor* dengan nilai  $k = 5, 7, \text{ dan } 9$ . Penelitian ini menggunakan *K-Fold Cross Validation* dengan skenario  $k = 10$  untuk melakukan validasi. Proses evaluasi kinerja model yang dilakukan menggunakan *Confusion Matrix*. Hasil penelitian ini menunjukkan bahwa akurasi tertinggi dicapai oleh nilai  $k = 5$  dengan *accuracy* 83%, *precision* 97 %, dan *recall* 83%.

UNUGIRI

## DAFTAR ISI

	Halaman
PERNYATAAN KEASLIAN TULISAN.....	ii
HALAMAN PERSETUJUAN.....	iii
LEMBAR PENGESAHAN .....	iv
MOTTO DAN PERSEMBAHAN .....	v
KATA PENGANTAR .....	vi
<i>ABSTRACT</i> .....	vii
ABSTRAK .....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR BAGAN.....	xiv
DAFTAR GAMBAR .....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat.....	4
BAB II TINJAUAN PUSTAKA DAN DASAR TEORI.....	5
2.1 Penelitian Terkait.....	5
2.2 Perbedaan dengan penelitian sebelumnya.....	12
2.3 Dasar Teori .....	12
2.3.1 Analisis Sentimen.....	12
2.3.2 Standar Pelayanan Rumah Sakit Menurut Permenkes .....	12
2.3.3 <i>K-Nearest Neighbor</i> .....	12
2.3.4 <i>Text Preprocessing</i> .....	13
2.3.5 TF-IDF .....	14
2.3.6 <i>K-Fold Cross Validation</i> .....	15
2.3.7 <i>Confusion Matrix</i> .....	15
2.3.8 <i>Data Labeling</i> .....	16
2.3.9 <i>Web Scraping</i> .....	16



2.3.10	<i>Google Maps</i> .....	17
BAB III METODE PENELITIAN.....		18
3.1	Objek Penelitian .....	18
3.2	<i>Pipeline Research</i> .....	18
3.3	Dataset .....	18
3.3.1	Populasi dan Sampel Penelitian .....	18
3.3.2	Metode Pengumpulan Data .....	18
3.3.4	Pelabelan Data.....	20
3.4	<i>Preprocessing</i> .....	20
3.4.1	<i>Case Folding</i> .....	21
3.4.2	<i>Cleaning</i> .....	21
3.4.3	<i>Tokenizing</i> .....	22
3.4.4	<i>Stopword Removal</i> .....	22
3.4.5	<i>Stemming</i> .....	23
3.5	TF-IDF.....	23
3.6	Algoritma <i>K-Nearest Neighbor</i> .....	25
3.7	<i>Validation</i> .....	26
3.8	<i>Evaluation</i> .....	27
3.9	Skenario Pengujian.....	28
3.10	Metode Pengembangan Sistem.....	29
3.10.1	Analisis.....	29
3.10.2	Desain.....	31
3.10.3	Implementasi .....	34
3.10.4	Pengujian.....	35
BAB IV HASIL DAN PEMBAHASAN .....		39
4.1	Pengujian Algoritma.....	39
4.1.1	Dataset.....	39
4.1.2	<i>Preprocessing Data</i> .....	40
4.1.3	TF-IDF .....	49
4.1.4	Algoritma <i>K-Nearest Neighbor</i> dan Validasi.....	52
4.1.5	<i>Evaluation</i> .....	56
4.2	Tampilan Website.....	58

4.2.1	Tampilan Login.....	58
4.2.2	Tampilan Beranda.....	58
4.2.3	Tampilan Dataset.....	59
4.2.4	Tampilan Klasifikasi.....	59
4.2.5	Tampilan Evaluasi.....	60
BAB V KESIMPULAN DAN SARAN.....		61
5.1	Kesimpulan.....	61
5.2	Saran.....	62
DAFTAR PUSTAKA.....		63
LAMPIRAN.....		67



**UNUGIRI**

## DAFTAR TABEL

Tabel 2. 1 Penelitian pertama.....	7
Tabel 2. 2 Penelitian kedua.....	8
Tabel 2. 3 Penelitian ketiga.....	9
Tabel 2. 4 Penelitian keempat.....	10
Tabel 2. 5 Penelitian kelima.....	11
Tabel 2. 6 Perbedaan dengan penelitian sebelumnya.....	12
Tabel 2. 7 <i>Confusion Matrix</i> .....	15
Tabel 3. 1 Hasil data scraper.....	20
Tabel 3. 2 Hasil pelabelan data.....	20
Tabel 3. 3 <i>Case Folding</i> .....	21
Tabel 3. 4 <i>Cleaning</i> .....	21
Tabel 3. 5 Tokenizing.....	22
Tabel 3. 6 <i>Stopword Removal</i> .....	22
Tabel 3.7 <i>Stemming</i> .....	23
Tabel 3.8 Perhitungan TF dan DF.....	23
Tabel 3.9 Perhitungan IDF.....	24
Tabel 3.10 Perhitungan TF-IDF.....	25
Tabel 3.11 Kebutuhan fungsional.....	29
Tabel 3. 12 Kebutuhan Perangkat Lunak.....	31
Tabel 3.13 Kebutuhan Perangkat Keras.....	31
Tabel 3. 14 Rencana Pengujian <i>Black Box</i> .....	35
Tabel 4.1 Dataset Ulasan Google Maps.....	39
Tabel 4.2 Dataset Ulasan yang Telah Diberi Label.....	40
Tabel 4.3 Hasil Proses <i>Case Folding</i> .....	42
Tabel 4.4 Hasil Proses <i>Cleaning</i> .....	43
Tabel 4.5 Hasil Proses <i>Tokenizing</i> .....	45
Tabel 4.6 Hasil <i>Stopword Removal</i> .....	47
Tabel 4.7 Hasil Proses <i>Stemming</i> .....	49
Tabel 4.8 Contoh Data Ulasan.....	53
Tabel 4.9 Hasil <i>Preprocessing Data</i> .....	53

Tabel 4.10 Hasil TF-IDF.....	53
Tabel 4.11 Perhitungan Jarak.....	55
Tabel 4.12 Hasil <i>Confusion Matrix</i> .....	57
Tabel 4.13 <i>Precision, Recall, dan Accuracy</i> .....	57



## DAFTAR BAGAN

Bagan 2. 1 Penelitian pertama.....	7
Bagan 2. 2 Penelitian kedua.....	8
Bagan 2. 3 Penelitian ketiga.....	9
Bagan 2. 4 Penelitian keempat.....	10
Bagan 2. 5 Penelitian kelima.....	11
Bagan 3. 1 Skenario Pengujian.....	28
Bagan 3.2 Model Waterfall.....	29



# UNUGIRI

## DAFTAR GAMBAR

Gambar 3. 1 Pipeline Research .....	18
Gambar 3. 2 Web Scraper .....	19
Gambar 3.3 Rumah sakit Dr Sosodoro pada <i>google maps</i> .....	19
Gambar 3.4 Proses <i>Scraper</i> .....	19
Gambar 3.5 Visualisasi K-Fold Cross Validation.....	26
Gambar 3.6 <i>Use Case Diagram</i> .....	31
Gambar 3. 5 Tampilan login .....	32
Gambar 3. 6 Tampilan halaman utama .....	33
Gambar 3. 7 Tampilan dataset .....	33
Gambar 3. 8 Tampilan klasifikasi .....	34
Gambar 3. 9 Tampilan evaluasi .....	34
Gambar 4.1 Tahapan <i>Preprocessing Data</i> .....	40
Gambar 4.2 Proses <i>Case Folding</i> .....	41
Gambar 4.3 Source Code Case Folding .....	41
Gambar 4.4 Proses <i>Cleaning Data</i> .....	42
Gambar 4.5 Source Code Proses <i>Cleaning Data</i> .....	43
Gambar 4.6 Proses <i>Tokenizing</i> .....	44
Gambar 4.7 Source Code Proses <i>Tokenizing</i> .....	45
Gambar 4.8 Proses <i>Stopword Removal</i> .....	46
Gambar 4.9 Source Code <i>Stopword Removal</i> .....	47
Gambar 4.10 Proses <i>Stemming</i> .....	48
Gambar 4.11 Source Code <i>Stemming</i> .....	49
Gambar 4.12 Proses TF-IDF .....	50
Gambar 4.13 Source Code TF-IDF .....	50
Gambar 4.14 Hasil Proses TF-IDF.....	51
Gambar 4.15 Proses Klasifikasi dengan KNN .....	52
Gambar 4.16 Source Code Klasifikasi dengan KNN.....	55
Gambar 4.17 Prediksi Label Klasifikasi .....	56
Gambar 4.18 Proses <i>Confusion Matrix</i> .....	56
Gambar 4.19 Source Code <i>Confusion Matrix</i> .....	57

Gambar 4.20 Tampilan Login.....	58
Gambar 4.21 Tampilan Beranda.....	58
Gambar 4.22 Tampilan Dataset.....	59
Gambar 4.23 Tampilan Preprocessing Data.....	59
Gambar 4.24 Tampilan Klasifikasi.....	60
Gambar 4.25 Tampilan Evaluasi.....	60

