

**ANALISIS SENTIMEN PADA ULASAN APLIKASI GOOGLE MAPS
TERHADAP PELAYANAN BADAN PENYELENGGARA JAMINAN
SOSIAL (BPJS) KESEHATAN SAMARINDA MENGGUNAKAN
METODE K-NEAREST NEIGHBOR DENGAN FITUR EKSTRAKSI
TF-IDF**

SKRIPSI

Diajukan Oleh :

Ikhsan Nuttakwa Takbirata Ihram Nabawi

2011102441072



PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR

JULI 2024

**ANALISIS SENTIMEN PADA ULASAN APLIKASI GOOGLE MAPS
TERHADAP PELAYANAN BADAN PENYELENGGARA JAMINAN
SOSIAL (BPJS) KESEHATAN SAMARINDA MENGGUNAKAN
METODE K-NEAREST NEIGHBOR DENGAN FITUR EKSTRAKSI
TF-IDF**

SKRIPSI

Diajukan Sebagai Salah Satu Persyaratan Untuk Memperoleh Gelar Sarjana Teknik
Informatika Fakultas Sains Dan Teknologi Universitas Muhammadiyah Kalimantan Timur

Diajukan Oleh :

Ikhsan Nuttakwa Takbirata Ihram Nabawi

2011102441072



PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS MUHAMMADIYAH KALIMANTAN TIMUR

JULI 2024

HALAMAN PERSETUJUAN

**ANALISIS SENTIMEN PADA ULASAN APLIKASI GOOGLE MAPS TERHADAP
PELAYANAN BADAN PENYELENGGARA JAMINAN SOSIAL (BPJS) KESEHATAN
SAMARINDA MENGGUNAKAN METODE K-NEAREST NEIGHBOR DENGAN
FITUR EKSTRAKSI TF-IDF**

SKRIPSI

Diajukan Oleh :

**Ikhsan Nuttakwa Takbirata Ihram Nabawi
2011102441072**

Disetujui untuk diujikan

Pada tanggal 27 Juni 2024

Pembimbing



Rudiman, S.Kom., M.Sc

NIDN 1105068202

Mengetahui,

Koordinator Tugas Akhir/Skripsi/Tesis/Disertasi



Abdul Rahim, S.Kom., M.Cs

NIDN 0009047901

HALAMAN PENGESAHAN

**ANALISIS SENTIMEN PADA ULASAN APLIKASI GOOGLE MAPS TERHADAP
PELAYANAN BADAN PENYELENGGARA JAMINAN SOSIAL (BPJS)
KESEHATAN SAMARINDA MENGGUNAKAN METODE K-NEAREST
NEIGHBOR DENGAN FITUR EKSTRAKSI TF-IDF**

Tugas akhir/skripsi



Diajukan oleh:

Ikhsan Nuttakwa Takbirata Ihram Nabawi

2011102441072

Diseminarkan dan Diujikan

Pada Tanggal 17 Juli 2024

Penguji I	Penguji II
 <u>Fendy Yulianto, S.Kom., M.Kom.</u> NIDN 1102079402	 <u>Rudiman, S.Kom, M.Sc</u> NIDN 1105068202

Mengetahui,

Ketua

Program Studi Teknik Informatika



Arbansyah, S.Kom, M.TI

NIDN. 1118019203

PERNYATAAN KEASLIAN PENELITIAN

Saya yang bertanda tangan di bawah ini:

Nama : Ikhsan Nuttakwa Takbirata Ihram Nabawi
NIM : 2011102441072
Program Studi : S1 Teknik Informatika
Judul Penelitian : Analisis Sentimen Pada Ulasan Aplikasi Google Maps Terhadap Pelayanan Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan Samarinda Menggunakan Metode K – Nearest Neighbor Dengan Fitur Ekstraksi TF-IDF

Menyatakan bahwa Skripsi yang saya tulis ini benar-benar hasil karya saya sendiri, dan bukan merupakan hasil plagiasi/falsifikasi/fabrikasi baik sebagian atau seluruhnya.

Atas pernyataan ini, saya siap menanggung resiko atau sanksi yang dijatuhkan kepada saya apa bila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam skripsi saya ini, atau klaim dari pihak lain terhadap keaslian karya saya ini.

Samarinda 30 Juni 2024

Yang membuat pernyataan



Ikhsan Nuttakwa Takbirata Ihram Nabawi

NIM: 2011102441072

ABSTRAK

Penelitian ini bertujuan untuk menganalisis sentimen masyarakat terhadap pelayanan BPJS Kesehatan Samarinda berdasarkan ulasan di aplikasi Google Maps. Metode yang digunakan dalam penelitian ini adalah *K-Nearest Neighbor* (KNN) dengan fitur ekstraksi TF-IDF (*Term Frequency-Inverse Document Frequency*). Data yang digunakan terdiri dari 500 ulasan berbahasa Indonesia yang dikumpulkan melalui teknik *web scraping*. Setelah proses pengumpulan data, data dilabelling oleh seorang ahli (*expert*) dilakukan tahap pre-processing yang terdiri dari *case folding*, *cleaning*, *tokenizing*, *stop word removal*, dan *stemming*. Data kemudian diberi bobot menggunakan metode TF-IDF untuk mengidentifikasi kata-kata penting. Pengujian dilakukan dengan menggunakan rasio data latih dan data uji sebesar 70:30 serta nilai $k = 5$. Hasilnya menunjukkan bahwa metode KNN mampu mengklasifikasikan sentimen positif dan negatif dengan tingkat akurasi sebesar 93,3%. Analisis ini memberikan gambaran tentang kualitas pelayanan BPJS Kesehatan di Samarinda dan dapat digunakan sebagai dasar untuk perbaikan layanan. Selain itu, penelitian ini juga memberikan kontribusi dalam penggunaan KNN dan TF-IDF untuk analisis sentimen, membuka peluang untuk penelitian lebih lanjut di bidang ini.

Kata kunci: Analisis Sentimen, BPJS Kesehatan, K-Nearest Neighbor

ABSTRACT

This study aims to analyze public sentiment towards the services of BPJS Kesehatan Samarinda based on reviews on the Google Maps application. The method used in this research is K-Nearest Neighbor (KNN) with TF-IDF (Term Frequency-Inverse Document Frequency) feature extraction. The data used consists of 500 Indonesian-language reviews collected through web scraping techniques. After the data collection process, the data was labeled by an expert, and then a pre-processing stage was carried out, including case folding, cleaning, tokenizing, stop word removal, and stemming. The data was then weighted using the TF-IDF method to identify important words. The testing was conducted using a training and testing data ratio of 70:30 and a k value of 5. The results showed that the KNN method was able to classify positive and negative sentiments with an accuracy rate of 93.3%. This analysis provides an overview of the service quality of BPJS Kesehatan in Samarinda and can be used as a basis for service improvements. Additionally, this research contributes to the use of KNN and TF-IDF for sentiment analysis, opening opportunities for further research in this field.

Keywords: *Sentiment Analysis, BPJS Kesehatan, K-Nearest Neighbor*

PRAKATA

Puji dan syukur saya panjatkan kepada Allah SWT. Atas ridho-Nya penulis dapat menyelesaikan penyusunan skripsi ini. Adapun judul skripsi yang penulis ajukan adalah “Analisis Sentimen Pada Ulasan Aplikasi Google Maps Terhadap Pelayanan Badan Penyelenggara Jaminan Sosial (BPJS) Kesehatan Samarinda Menggunakan Metode *K – Nearest Neighbor* Dengan Fitur Ekstraksi TF-IDF”

Skripsi ini diajukan untuk memenuhi syarat kelulusan mata kuliah Skripsi di Fakultas Sains dan Teknologi, Universitas Muhammadiyah Kalimantan Timur. Tidak dapat disangka bahwa butuh usaha yang keras dalam penyelesaian pengerjaan skripsi ini. Namun, karya ini tidak akan selesai tanpa orang-orang tercinta di sekeliling penulis yang mendukung dan membantu. Terima kasih penulis sampaikan kepada:

1. Bapak Dr. Muhammad Musiyam, M.T. Selaku Rektor Universitas Muhammadiyah Kalimantan Timur yang telah memberi kesempatan pada peneliti untuk menempuh pendidikan di Universitas Muhammadiyah Kalimantan Timur
2. Bapak Rudiman, M.Sc. Selaku Dosen Pembimbing penulis yang telah membagi ilmu, bimbingan dan arahan selama proses perkuliahan.
3. Orang tua dan saudari penulis yang selalu mendukung, memotivasi dan mendoakan penulis sehingga proses pengerjaan skripsi ini selesai sampai akhir.
4. Pacar penulis yang bernama Nur Adelia Raihanda yang telah membantu penulis, menjadi *support system* dan mendoakan penulis sampai akhirnya penulis bisa mengerjakan skripsi ini sampai selesai.

5. Sahabat penulis, Khoirul Huda, Rayhan Elfansyah, Ipan Hasmadi, Farhat Jundullah, Zulkarnaen, dan anggota kos lainnya terimakasih atas dukungan dan semangat yang diberikan kepada penulis selama ini.
6. Teman satu bimbingan dan teman kuliah peneliti yang bernama Takhta Perlawanan, Ricky, Andi Nur Halim, Aditya dll yang sudah membantu peneliti dan mensupport peneliti, bersama-sama menyelesaikan skripsi ini sampai akhir.

Semoga segala kebaikan dan pertolongan semuanya mendapat berkah dari Allah SWT. Dan akhirnya penulis menyadari bahwa skripsi ini tidak luput dari berbagai kekurangan, karena keterbatasan ilmu yang saya miliki. Untuk itu dengan kerendahan hati saya mengharapkan saran dan kritik yang sifatnya membangun dari semua pihak demi membangun laporan penelitian ini,

Samarinda, 30 Juni 2024



Ikhsan Nuttakwa Takbirata Ihram Nabawi

DAFTAR ISI

	Halaman
HALAMAN PERSETUJUAN	i
HALAMAN PENGESAHAN.....	ii
PERNYATAAN KEASLIAN PENELITIAN	iii
ABSTRAK	iv
<i>ABTRACT</i>	v
PRAKATA	vi
DAFTAR ISI.....	viii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
BAB II METODELOGI PENELITIAN	5
2.1. Objek Penelitian	5
2.2. Alat dan Bahan	5
2.3. Prosedur Penelitian	5
2.3.1. Pengumpulan Data	6
2.3.2. <i>Labeling Data</i>	7
2.3.3. <i>Pre-Processing</i>	8
2.3.4. Pembobotan Kata (TF-IDF)	9
2.3.5. <i>Split Data</i>	10
2.3.6. Klasifikasi.....	10
2.3.7. <i>K-Nearest Neighbors</i> (KNN)	10
2.3.8. Evaluasi	11
2.4 Jadwal Penelitian	12
BAB III HASIL DAN PEMBAHASAN.....	14
3.1 Pengumpulan Data.....	14
3.2 Labelling Data	15

3.3	<i>Pre-Processing</i>	16
3.3.1	<i>Case Folding</i>	16
3.3.2	<i>Cleaning</i>	17
3.3.3	<i>Tokenizing</i>	18
3.3.4	<i>Stopword Removal</i>	18
3.3.5	<i>Stemming</i>	19
3.4	Pembobotan Kata (TF-IDF).....	19
3.5	Visualisasi.....	21
3.6	Evaluasi	22
BAB IV PENUTUP		28
4.1	Simpulan.....	28
4.2	Implikasi	28
DAFTAR RUJUKAN		30
LAMPIRAN		32
RIWAYAT HIDUP		47

DAFTAR TABEL

Tabel	Halaman
Tabel 2.1 Jadwal Penelitian.....	13
Tabel 3.1 Pengumpulan Data	14
Tabel 3.2 Labeling Data	15
Tabel 3.3 Hasil Data Case Folding.....	17
Tabel 3.4 Hasil Data Cleaning	17
Tabel 3.5 Hasil Data Tokenizing.....	18
Tabel 3.6 Hasil Data Stopword Removal.....	18
Tabel 3.7 Hasil Data Stemming	19
Tabel 3.8 Pembagian Data KNN.....	22
Tabel 3.9 Hasil Data Uji.....	26

DAFTAR GAMBAR

Gambar	Halaman
Gambar 2.1 Alur Penelitian.....	6
Gambar 2.2 Hasil Crawling Data	7
Gambar 3.1 Visualisasi Sentimen	16
Gambar 3.2 Pembobotan Kata (TF-IDF)	20
Gambar 3.3 Scoring TF	21
Gambar 3.4 WordCloud Positif.....	21
Gambar 3.5 WordCloud Negatif.....	22
Gambar 3.6 Hasil Akurasi Nilai K.....	24
Gambar 3.7 Plot Akurasi Nilai K.....	25
Gambar 3.8 Hasil Akurasi Metode KNN.....	26
Gambar 3.9 Hasil Confusion Matrix	26

DAFTAR LAMPIRAN

Lampiran	Halaman
L1 CV Expert Labeling.....	32
L2 Surat Izin Penelitian.....	34
L3 Pengumpulan Data.....	35
L4 Install Library	37
L5 Membaca Dataset.....	37
L6 Cek Informasi Dataset	38
L7 Cek Jumlah Sentimen Positif dan Negatif.....	38
L8 Persentase Sentimen	38
L9 Membaca Dataset Kolom Review dan Sentimen	38
L10 Pre-Processing Data.....	39
L11 Pembobotan Kata TF-IDF	40
L12 Split Data dan Evaluasi Confusion Matrix Metode KNN	41
L13 Hasil Akurasi Pada Nilai K.....	41
L14 Plot Akurasi Nilai K	42
L15 Visualisasi Wordcloud Positif dan Negatif.....	42
L16 Kartu Kendali Bimbingan.....	43