



SKRIPSI

ANALISIS FAKTOR YANG MEMPENGARUHI KEPUASAN PESERTA BPJS KESEHATAN PADA RUMAH SAKIT WILAYAH SURABAYA DENGAN PENDEKATAN ANALISIS SENTIMEN

RAHMATUL AMANILLAH
NPM 20083010002

DOSEN PEMBIMBING
Trimono, S.Si., M.Si.
Aviolla Terza Damaliana, S.Si., M.Stat.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**



SKRIPSI

ANALISIS FAKTOR YANG MEMPENGARUHI KEPUASAN PESERTA BPJS KESEHATAN PADA RUMAH SAKIT WILAYAH SURABAYA DENGAN PENDEKATAN ANALISIS SENTIMEN

RAHMATUL AMANILLAH
NPM 20083010002

DOSEN PEMBIMBING
Trimono, S.Si., M.Si.
Aviolla Terza Damaliana, S.Si., M.Stat.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**



SKRIPSI

ANALISIS FAKTOR YANG MEMPENGARUHI KEPUASAN PESERTA BPJS KESEHATAN PADA RUMAH SAKIT WILAYAH SURABAYA DENGAN PENDEKATAN ANALISIS SENTIMEN

RAHMATUL AMANILLAH

NPM 20083010002

DOSEN PEMBIMBING

Trimono, S.Si., M.Si.

Aviolla Terza Damaliana, S.Si., M.Stat.

**KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS PEMBANGUNAN NASIONAL VETERAN JAWA TIMUR
FAKULTAS ILMU KOMPUTER
PROGRAM STUDI SAINS DATA
SURABAYA
2025**

LEMBAR PENGESAHAN

ANALISIS FAKTOR YANG MEMPENGARUHI KEPUASAN PESERTA BPJS KESEHATAN PADA RUMAH SAKIT WILAYAH SURABAYA DENGAN PENDEKATAN ANALISIS SENTIMEN

Oleh:
Rahmatul Amanillah
NPM. 20083010002

Telah dipertahankan, dihadapan, dan diterima oleh Tim Penguji Sidang Skripsi Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal 16 Juni 2025:

Menyetujui,

Trimono, S.Si., M.Si.
NIP 19950908 202203 1 003


..... (Pembimbing I)

Aviolla Terza Damaliana, S.Si., M.Stat.
NIP 19940802 202203 2 005


..... (Pembimbing II)

Amri Muhaimin, S.Stat., M.Stat., M.S.
NIP. 19950723 202406 1 002


..... (Ketua Penguji)

Shindi Shella May Wara, M.Stat
NIP 19960518 202406 2 0003


..... (Penguji I)

Mengetahui,
Dekan Fakultas Ilmu Komputer


Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

**ANALISIS FAKTOR YANG MEMPENGARUHI KEPUASAN PESERTA
BPJS KESEHATAN PADA RUMAH SAKIT WILAYAH SURABAYA
DENGAN PENDEKATAN ANALISIS SENTIMEN**

Oleh:

RAHMATUL AMANILLAH

NPM. 20083010002

Telah disetujui untuk mengikuti Ujian Skripsi

Menyetujui,

**Koordinator Program Studi Sains Data
Fakultas Ilmu Komputer**

Dr. Eng. Ir. Dwi Arman Prasetya, S.T., M.T., IPU., Asean, Eng.

NIP. 19801205 200501 1002

Halaman ini sengaja dikosongkan

SURAT PERNYATAAN BEBAS PLAGIASI

Saya yang bertandatangan di bawah ini:

Nama Mahasiswa : Rahmatul Amanillah
NPM : 20083010002
Program : Sarjana (S1)
Program Studi : Sains Data
Fakultas : Fakultas Ilmu Komputer

Menyatakan bahwa dalam dokumen ilmiah Skripsi ini tidak terdapat bagian dari karya ilmiah lain yang telah diajukan untuk memperoleh gelar akademik di suatu lembaga Pendidikan Tinggi, dan juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang/lembaga lain, kecuali yang secara tertulis disitasi dalam dokumen ini dan disebutkan secara lengkap dalam pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila di kemudian hari ditemukan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi dengan peraturan perundang-undangan yang berlaku.

Demikian surat pernyataan ini saya buat dengan sesungguhnya tanpa ada paksaan dari siapapun juga dan untuk dipergunakan sebagaimana mestinya.



Surabaya, 12 Oktober 2025
Yang Membuat Pernyataan,


RAHMATUL AMANILLAH
NPM. 20083010002

Halaman ini sengaja dikosongkan

ABSTRAK

Nama Mahasiswa / NPM : Rahmatul Amanillah / 20083010002
Judul Skripsi : Analisis Faktor Yang Mempengaruhi Kepuasan Peserta BPJS Kesehatan pada Rumah Sakit Wilayah Surabaya dengan Pendekatan Analisis Sentimen
Dosen Pembimbing : 1. Trimono, S.Si., M.Si.
2. Aviolla Terza Damaliana, S.Si., M.Stat.

Program Jaminan Kesehatan Nasional (JKN) oleh BPJS telah hadir sejak 2014 untuk menjamin kebutuhan dasar kesehatan masyarakat. Namun, program ini masih menghadapi berbagai masalah, seperti pembatasan kuota, pelayanan kurang maksimal, serta kendala fasilitas dan kebijakan. Sehingga, peninjauan berkala diperlukan untuk menilai tingkat kepuasan peserta terhadap pelayanan BPJS Kesehatan. Oleh karena itu, peneliti melakukan analisis sentimen dengan menggunakan model *ServQual* pada pembuatan kuisisioner dan menghasilkan data teks yang digunakan dalam analisis sentimen. Analisis sentimen dilakukan untuk menyerap sentimen negatif dan sentimen positif yang ada pada data. Proses analisis sentimen dimulai dari pembersihan data, pelabelan data, pembobotan kata dengan TF-IDF, model klasifikasi dengan algoritma *support vector machine* (SVM), lalu evaluasi model dengan *confusion matrix*, dan terakhir visualisasi data sentimen positif dan negatif dengan *wordcloud*. Juga menggunakan metode SMOTE untuk menangani ketidakseimbangan data. Hasil analisis menunjukkan bahwa model SVM dengan kernel RBF lebih konsisten dalam pengklasifikasian data dari 6 aspek yang diteliti, dengan performa model terbaik yaitu dengan parameter $C = 1$ setelah dilakukan SMOTE dengan nilai akurasi 90%, presisi 93%, recall 88%, dan F1_Score 89%. Model mampu mengklasifikasikan sentimen positif dan sentimen negatif dengan baik. Beberapa penemuan yang perlu ditingkatkan meliputi antrean yang panjang, keterbatasan fasilitas medis, ketidakefisienan administrasi, serta inkonsistensi layanan tenaga medis.

Kata kunci : *Analisis sentiment, BPJS Kesehatan, support vector machine, pelayanan kesehatan, kepuasan pasien*

Halaman ini sengaja dikosongkan

ABSTRACT

Student Name / NPM : Rahmatul Amanillah / 20083010002
Thesis Title : *Analysis of Factors Influencing BPJS Health Participants' Satisfaction in Hospitals in the Surabaya Region Using Sentimen Analysis Approach*
Advisor : 1. Trimono, S.Si., M.Si.
2. Aviolla Terza Damaliana, S.Si., M.Stat.

The National Health Insurance (JKN) program by BPJS has been present since 2014 to guarantee the basic health needs of the community. However, this program still faces various problems, such as limited quotas, less than optimal services, and constraints on facilities and policies. Thus, periodic studies are needed to assess the level of participant satisfaction with BPJS Kesehatan services. Therefore, researchers conducted sentiment analysis using the ServQual model in creating questionnaires and producing text data used in sentiment analysis. Sentiment analysis is carried out to absorb negative and positive sentiments in the data. The sentiment analysis process starts from data cleaning, data labeling, word weighting with TF-IDF, classification models with the support vector machine (SVM) algorithm, then model evaluation with configuration matrix, and finally visualization of positive and negative sentiment data with wordcloud. In addition, using the SMOTE method to handle data imbalance. The results of the analysis show that the SVM model with the RBF kernel is more consistent in classifying data from the 6 aspects studied, with the best model performance being with parameter $C = 1$ after SMOTE with an accuracy value of 90%, precision of 93%, recall of 88%, and F1_Score of 89%. Some of the findings that needed improvement included long queues, limited medical facilities, administrative inefficiencies, and inconsistencies in medical personnel services.

Keywords: *Sentiment analysis, BPJS Kesehatan, support vector machine, health services, patient satisfaction*

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala rahmat, hidayah dan karunia-Nya kepada penulis sehingga skripsi dengan judul **“Analisis Faktor yang Mempengaruhi Kepuasan Peserta BPJS Kesehatan Pada Rumah Sakit Wilayah Surabaya dengan Pendekatan Analisis Sentimen”** dapat terselesaikan dengan baik.

Penulis mengucapkan terima kasih kepada Bapak Bapak Trimono, S.Si., M.Si. selaku Dosen Pembimbing 1 dan Ibu Aviolla Terza Damaliana, S.Si., M.Stat. selaku Dosen Pembimbing 2 yang telah meluangkan waktunya untuk memberikan bimbingan, nasehat serta motivasi kepada penulis. Dan penulis juga banyak menerima bantuan dari berbagai pihak, baik itu berupa moril, spiritual maupun materiil. Untuk itu penulis mengucapkan terima kasih kepada:

1. Ibu Prof. Dr. Ir. Novirina Hendrasarie, MT. selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
2. Bapak Dr. Eng. Ir. Dwi Arman Prasetya, ST., MT., IPU., Asean. Eng selaku Ketua Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran “ Jawa Timur.
3. Bapak Dr. Ir. I Gede Susrama Mas Diyasa, ST. MT. IPU selaku Dosen Wali Mahasiswa selama menempuh Pendidikan di Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran “ Jawa Timur.
4. Bapak/Ibu Dosen Program Studi Sains Data Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran “ Jawa Timur yang telah membimbing selama masa perkuliahan.
5. Kedua Orang tua dan keluarga yang telah memberikan dorongan semangat, do’a, dan kasih sayangnya.
6. Teman-teman yang banyak membantu dari awal perkuliahan sampai dengan sekarang melewati masa perkuliahan.

Penulis menyadari bahwa di dalam penyusunan skripsi ini banyak terdapat kekurangan. Untuk itu kritik dan saran yang membangun dari semua pihak sangat

diharapkan demi kesempurnaan penulisan skripsi ini. Akhirnya, dengan segala keterbatasan yang penulis miliki semoga laporan ini dapat bermanfaat bagi semua pihak umumnya dan penulis pada khususnya.

Surabaya, oktober 2025

Penulis

DAFTAR ISI

HALAMAN SAMPUL	i
LEMBAR PENGESAHAN	ii
LEMBAR PERSETUJUAN.....	iii
SURAT PERNYATAAN BEBAS PLAGIASI.....	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xiii
DAFTAR GAMBAR.....	xv
DAFTAR TABEL	xvii
DAFTAR LAMPIRAN	xix
DAFTAR NOTASI.....	xxi
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	5
1.3. Batasan Masalah.....	5
1.4. Tujuan Penelitian	6
1.5. Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA.....	7
2.1. Penelitian Terdahulu	7
2.2. Dasar Teori.....	10
2.2.1. Badan Penyelenggara Jaminan Sosial (BPJS).....	10
2.2.2. Kepuasan Pasien dalam Layanan Kesehatan.....	10
2.2.3. <i>Natural Language Processing</i> (NLP).....	13
2.2.4. Analisis Sentimen	14
2.2.5. Leksikon dalam Analisis Sentimen	18
2.2.6. <i>Term Frequency – Inverse Document Frequency</i> (TF-IDF)	18
2.2.7. <i>Support Vector Machine</i> (SVM).....	19
2.2.8. <i>Data Imbalance</i>	24
2.2.9. <i>Synthetic Minority Oversampling Technique</i> (SMOTE)	24

2.2.10. Evaluasi Model	26
BAB III DESAIN DAN IMPLEMENTASI SISTEM.....	29
3.1. Variabel Penelitian dan Sumber Data.....	29
3.1.1. Variabel Penelitian.....	29
3.1.2. Sumber Data.....	32
3.2. Langkah Analisis	32
3.2.1. Pengumpulan Data	32
3.2.2. Praproses dan Pembersihan Data	33
3.2.3. Analisis Data	34
3.2.4. Pembuatan Graphical User Interface (GUI).....	36
3.2.5. Pembahasan dan Penyusunan Laporan	36
3.3. Desain Sistem	37
BAB IV HASIL DAN PEMBAHASAN	39
4.1. Pengambilan Data Penelitian.....	39
4.2. Praproses dan Pembersihan Data.....	40
4.3. Eksplorasi Data.....	47
4.4. Eksplorasi Data Sentimen.....	51
4.5. <i>Synthetic Minority Oversampling Technique (SMOTE)</i>	52
4.6. Ekstraksi Fitur TF-IDF	53
4.7. Pemodelan Support Vector Machine (SVM).....	55
4.8. Evaluasi Model.....	59
4.9. Confusion Matrix.....	64
4.10. Visualisasi Word Cloud.....	65
4.11. <i>Graphical User Interface (GUI)</i>	73
4.12. Gap Analisis	77
BAB V PENUTUP	79
5.1. Kesimpulan.....	79
5.2. Saran Pengembangan.....	80
DAFTAR PUSTAKA	81
LAMPIRAN	87

DAFTAR GAMBAR

Gambar 2.1 Model <i>Support Vector Machine</i> (SVM)	20
Gambar 2.2 <i>Hyperplane</i> klasifikasi	21
Gambar 2.3 Kernel <i>Polynomial</i>	23
Gambar 2.4 Kernel <i>Radian Basic Function</i> (RBF)	23
Gambar 2.5 Kernel <i>Signoid</i>	24
Gambar 2.6 Ilustrasi algoritma SMOTE	25
Gambar 3.1 Desain alir kegiatan penelitian	36
Gambar 3.2 Desain <i>Graphical User Interface</i> (GUI)	37
Gambar 4.1 Dokumentasi Kuisisioner <i>google form</i>	39
Gambar 4.2 Dokumentasi angket	39
Gambar 4.3 Diagram batang distribusi jenis kelamin	48
Gambar 4.4 Diagram batang distribusi pendidikan terakhir	49
Gambar 4.5 Diagram batang distribusi pekerjaan	50
Gambar 4.6 Diagram batang distribusi jumlah kunjungan	51
Gambar 4.7 Kata yang sering muncul pada sentimen negatif aspek <i>tangibles</i>	65
Gambar 4.8 <i>Word cloud</i> sentimen negatif aspek <i>tangibles</i>	65
Gambar 4.9 Kata yang sering muncul pada sentimen positif aspek <i>tangibles</i>	66
Gambar 4.10 <i>Word cloud</i> sentimen positif aspek <i>tangibles</i>	66
Gambar 4.11 Kata yang sering muncul pada sentimen negatif aspek <i>reliability</i> ..	67
Gambar 4.12 <i>Word cloud</i> sentimen negatif aspek <i>reliability</i>	67
Gambar 4.13 Kata yang sering muncul pada sentimen positif aspek <i>reliability</i> ...	67
Gambar 4.14 <i>Word cloud</i> sentimen positif aspek <i>reliability</i>	67
Gambar 4.15 Kata yang sering muncul pada sentimen negatif aspek <i>responsiveness</i>	68
Gambar 4.16 <i>Word cloud</i> sentimen negatif aspek <i>responsiveness</i>	68
Gambar 4.17 Kata yang sering muncul pada sentimen positif aspek <i>responsiveness</i>	68
Gambar 4.18 <i>Word cloud</i> sentimen positif aspek <i>responsiveness</i>	68
Gambar 4.19 Kata yang sering muncul pada sentimen negatif aspek <i>assurance</i> .	69
Gambar 4.20 <i>Word cloud</i> sentimen negatif aspek <i>assurance</i>	69

Gambar 4.21 Kata yang sering muncul pada sentimen positif aspek <i>assurance</i> ...	70
Gambar 4.22 <i>Word cloud</i> sentimen positif aspek <i>assurance</i>	70
Gambar 4.23 Kata yang sering muncul pada sentimen negatif aspek <i>empathy</i>	71
Gambar 4.24 <i>Word cloud</i> sentimen negatif aspek <i>empathy</i>	71
Gambar 4.25 Kata yang sering muncul pada sentimen positif aspek <i>empathy</i>	71
Gambar 4.26 <i>Word cloud</i> sentimen positif aspek <i>empathy</i>	71
Gambar 4.27 Kata yang sering muncul pada sentimen negatif aspek sentimen umum	72
Gambar 4.28 <i>Word cloud</i> sentimen negatif aspek sentimen umum	72
Gambar 4.29 Kata yang sering muncul pada sentimen positif aspek sentimen umum	72
Gambar 4.30 <i>Word cloud</i> sentimen positif aspek sentimen umum	72
Gambar 4.31 Tampilan Utama GUI	73
Gambar 4.32 Contoh inputan teks Prediksi Manual.....	74
Gambar 4.33 Contoh mengimpor file Excel untuk prediksi.....	74
Gambar 4.34 Contoh hasil WordCloud GUI.....	75
Gambar 4.35 Contoh tampilan tabel hasil prediksi GUI	76
Gambar 4.36 Contoh tampilan Performa Model pada GUI	76

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	7
Tabel 2.2 Contoh hasil <i>case folding</i>	15
Tabel 2.3 Contoh hasil <i>tokenization</i>	16
Tabel 2.4 Contoh hasil <i>stemming/lemmatization</i>	16
Tabel 2.5 Contoh hasil <i>normalization</i>	17
Tabel 2.6 Contoh hasil <i>stopword</i>	17
Tabel 2.7 Tabel <i>confusional matrix</i>	26
Tabel 3.1 Jenis data demografis peserta.....	29
Tabel 3.2 Jenis data sentimen peserta	29
Tabel 3.3 Struktur data informasi demografis.....	29
Tabel 3.4 Struktur data sentimen aspek tangibles.....	30
Tabel 3.5 Struktur data sentimen aspek reliability.....	30
Tabel 3.6 Struktur data sentimen aspek responsiveness	30
Tabel 3.7 Struktur data sentimen aspek assurance.....	31
Tabel 3.8 Struktur data sentimen aspek empathy.....	31
Tabel 3.9 Struktur data sentimen aspek sentimen umum.....	32
Tabel 4.1 Dokumentasi data pada <i>Microsof Excel</i>	40
Tabel 4.2 Data sentimen peserta asspek <i>tangibles</i>	40
Tabel 4.3 Data sentimen peserta asspek <i>reliability</i>	41
Tabel 4.4 Data sentimen peserta asspek <i>responsiveness</i>	41
Tabel 4.5 Data sentimen peserta asspek <i>assurance</i>	41
Tabel 4.6 Data sentimen peserta asspek <i>empathy</i>	42
Tabel 4.7 Data sentimen peserta asspek sentimen umum	42
Tabel 4.8 Data sebelum pra-proses dan pembersihan data	42
Tabel 4.9 Dokumentasi pembersihan teks	43
Tabel 4.10 Dokumentasi hasil tokenisasi.....	44
Tabel 4.11 Dokumentasi hasil penghapusan stop words	45
Tabel 4.12 Dokumentasi hasil Stemming	45
Tabel 4.13 Dokumentasi hasil konfersi kata informal	46

Tabel 4.14 Dokumentasi hasil pelabelan data	47
Tabel 4.15 Distribusi sentimen negatif dan positif.....	51
Tabel 4.16 Data contoh ekstraksi fitur <i>TF_IDF</i>	53
Tabel 4.17 Hasil ekstraksi fitur TF-IDF	54
Tabel 4.18 Hasil <i>kernel Polynomial</i> tanpa SMOTE	59
Tabel 4.19 Hasil <i>kernel polynomial</i> dengan SMOTE	60
Tabel 4.20 Hasil <i>kernel RBF</i> tanpa SMOTE	60
Tabel 4.21 Hasil <i>kernel RBF</i> dengan SMOTE	60
Tabel 4.22 Hasil <i>kernel Sigmoid</i> tanpa SMOTE	61
Tabel 4.23 Hasil <i>kernel Sigmoid</i> dengan SMOTE	61
Tabel 4.24 Perbandingan nilai akurasi parameter C.....	62
Tabel 4.25 Perbandingan nilai presisi parameter C.....	62
Tabel 4.26 Perbandingan nilai recall parameter C	63
Tabel 4.27 Perbandingan nilai F1_score parameter C.....	63
Tabel 4.28 Hasil evaluasi model SVM oleh confusion matrix.....	64
Tabel 4.29 Contoh hasil evaluasi model data <i>empathy</i>	64

DAFTAR LAMPIRAN

Lampiran 1 LoA artikel penelitian.....	87
Lampiran 2 Daftar kuisisioner penelitian	88
Lampiran 3 Data penelitian.....	89
Lampiran 4 <i>Sintaks</i> pemrograman	90

Halaman ini sengaja dikosongkan

DAFTAR NOTASI

TP	:	<i>True Positif</i>
TN	:	<i>True Negative</i>
FP	:	<i>False Positif</i>
FN	:	<i>False Negative</i>
$TF(t, d)$	:	Fungsi untuk menghitung TF dari sebuah kata t dalam sebuah dokumen d
$n_{t,d}$	:	Jumlah kemunculan t dalam dokumen d
n_d	:	Jumlah total kata dalam dokumen d
$IDF(t, D)$	:	Fungsi untuk menghitung IDF dari sebuah kata t di dalam sebuah koleksi dokumen D
log	:	Operasi logaritma
N	:	Jumlah total dokumen dalam korelasi
n_t	:	Jumlah dokumen dalam koleksi yang mengandung kata t
$TF - IDF(t, d, D)$	:	Fungsi untuk pembobotan TF-IDF
$f(x)$	:	Fungsi keputusan
W	:	<i>Vector</i> bobot
X	:	<i>Vector fitur input</i>
b	:	Bias
γ	:	Parameter gamma
α	:	<i>alpha</i>
exp	:	Fungsi eksponensial
$\ x_1 - x_2\ $	:	Aturan <i>Euclidean</i> yang berupa jarak dari x_1 ke x_2
X_{new}	:	Data sintetis baru
X_i	:	Data minoritas asli
X_{knn}	:	Tetangga minoritas terdekat
δ	:	Parameter acak antara 0 dan 1