

**KLASIFIKASI PELAYANAN KESEHATAN
BERDASARKAN DATA SENTIMEN PELAYANAN
KESEHATAN MENGGUNAKAN *MULTICLASS SUPPORT
VECTOR MACHINE (SVM)***

TESIS



**UNIVERSITAS PENDIDIKAN GANESHA
PASCASARJANA
PROGRAM STUDI ILMU KOMPUTER
DESEMBER 2022**

**KLASIFIKASI PELAYANAN KESEHATAN
BERDASARKAN DATA SENTIMEN PELAYANAN
KESEHATAN MENGGUNAKAN *MULTICLASS SUPPORT
VECTOR MACHINE (SVM)***

TESIS

Diajukan kepada
Universitas Pendidikan Ganesha
Untuk Memenuhi Sebagai Persyaratan
Memperoleh Gelar Magister
Program Studi Ilmu Komputer

oleh

MOH. HERI SETIAWAN

NIM 2029101012



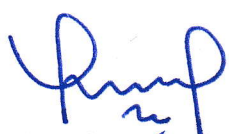
**UNIVERSITAS PENDIDIKAN GANESHA
PASCASARJANA
PROGRAM STUDI ILMU KOMPUTER
DESEMBER 2022**

LEMBAR PERSETUJUAN PEMBIMBING

Tesis oleh Moh. Heri Setiawan ini telah diperiksa dan disetujui untuk Mengikuti Ujian Tesis.

Singaraja, 16 Desember 2022

Pembimbing I



Dr. I Gede Aris Gunadi, S.Si, M.Kom.

NIP 19770318 200812 1 000

Pembimbing II



Dr. Gede Indrawan, S.T, M.T.

NIP 19760102 200312 1 001



LEMBAR PERSETUJUAN TIM PENGUJI

Tesis oleh Moh. Heri Setiawan ini telah dipertahankan di depan tim penguji dan dinyatakan diterima sebagai salah satu persyaratan untuk memperoleh gelar Magister di Program Studi Ilmu Komputer, Pascasarjana, Universitas Pendidikan Ganesha.

Disetujui pada tanggal: 26 Desember 2022

Oleh,
Tim Penguji

Ketua (Dr. I Gede Aris Gunadi, S.Si., M.Kom)

NIP 19770318 200812 1 000

Anggota (Prof. Dr. Sariyasa, M.Sc., Ph. D)

NIP 19640615 198902 1 000

Anggota (Dr. Komang Setemen, S.Si., M.T)

NIP 19760315 200112 1 000

Anggota (Dr. Gede Indrawan, S.T., M.T)

NIP 19760102 200312 1 001

Mengetahui Direktur

Pascasarjana Undiksha



Prof. Dr. I Gusti Putu Suharta, M.Si.

NIP 196212151988031002

LEMBAR PERNYATAAN

Saya menyatakan dengan sesungguhnya bahwa tesis yang saya susun sebagai syarat untuk memperoleh gelar Magister dari Pascasarjana Universitas Pendidikan Ganesha seluruhnya merupakan hasil karya saya sendiri. Bagian-bagian tertentu dalam penulisan tesis yang saya kutip dari hasil karya orang lain telah dituliskan sumbernya secara jelas dan sesuai dengan norma, kaidah, serta etika akademis.

Apabila di kemudian hari ditemukan seluruh atau sebagian tesis ini bukan hasil karya saya sendiri atau adanya plagiat dalam bagian-bagian tertentu, saya bersedia menerima sanksi pencabutan gelar akademik yang saya sandang dan sanksi-sanksi lainnya sesuai dengan peraturan perundang-undangan di wilayah Negara Kesatuan Republik Indonesia.

Singaraja, 26 Desember 2022

Yang Memberi Pernyataan



10000
SEPUTUH RIBU RUPIAH
TBL 20
METERAI
TEMPEL
0AFAKX284387448

Moh. Heri Setiawan

PRAKATA

Puji dan syukur penulis panjatkan kepada Allah SWT karena atas anugrah-Nya tesis yang berjudul “Klasifikasi Pelayanan Kesehatan Berdasarkan Data Sentimen Pelayanan Kesehatan Menggunakan *Multiclass Support Vector Machine* (SVM)” dapat diselesaikan sesuai dengan yang direncanakan.

Tesis ini ditulis untuk memenuhi sebagai persyaratan dalam menyelesaikan studi di Program Pascasarjana Universitas Pendidikan Ganesha, Program Studi Ilmu Komputer. Pada halaman ini, ijin penulis menyampaikan terimakasih dan penghargaan yang sebesar-besarnya kepada:

1. Dr. I Gede Aris Gunadi, S.Si, M.Kom sebagai pembimbing I yang dengan sabar memberikan bimbingan, arahan, dan motivasi yang demikian bermakna, sehingga penulis mampu melewati berbagai permasalahan dalam perjalanan studi dan penyelesaian proposal tesis ini.
2. Dr. Gede Indrawan, M.T sebagai pembimbing II yang telah memberikan bimbingan, arahan, dan memberikan motivasi selama penyusunan proposal tesis ini sehingga penulis dapat menyelesaikannya sesuai harapan.
3. Bapak Rektor Universitas Pendidikan Ganesha yang telah memberikan bantuan secara moral dan memfasilitasi berbagai kepentingan studi, selama penulis menempuh perkuliahan di Program Pascasarjana Undiksha.
4. Bapak Direktur Program Pascasarjana Undiksha dan Staff yang telah banyak membantu selama penulis mengikuti dan menyelesaikan penulisan proposal tesis ini.
5. Ketua Program Studi Ilmu Komputer dan Staf Dosen pengajar di Program Studi Ilmu Komputer yang telah banyak membantu dan

memotivasi penulis selama perjalanan studi dan penyusunan proposal tesis ini.

6. Rekan-rekan seangkatan di Program Studi Ilmu Komputer yang dengan karakternya masing-masing telah banyak berkontribusi membentuk kemandirian penulis selama menjalani studi dan menyelesaikan tesis ini.
7. Kedua orang tua penulis, serta saudara-saudara dan keluarga yang telah banyak membantu secara material dan moral selama perjalanan studi yang penulis tempuh di Program Studi Ilmu Komputer.
8. Istriku Fatimah dan anak-anakku: Muhammad Fatih Al Khawarizmi, Zafina Khanza Al Farabi, dan Azmi Abdullah Al Amin yang penulis banggakan dan penulis cintai, dimana selalu mendukung dan membantu selama penulis mengikuti perkuliahan di Program Studi Ilmu Komputer.
9. Ketua Yayasan Al Ma'ruf Denpasar yang selalu mendukung dan membantu penulis untuk bisa melanjutkan pendidikan ke Program Studi Ilmu Komputer ini.
10. Rekan-rekan dari seluruh Guru dan Staff SMK Bina Madina Denpasar yang terus memberikan dukungan selama menyelesaikan program studi ini.

Semoga semua bantuan yang telah mereka berikan dalam perjalanan studi penulis dapat menjadi amal baik yang diterima oleh Tuhan Yang Maha Esa, sehingga mereka mendapatkan kemudahan dalam segala hal yang ingin mereka capai.

Penulis menyadari bahwa sepenuhnya penulisan ini masih jauh dari kata sempurna, oleh karena itu seluruh jenis saran, kritikan dan masukan yang dapat membangun tulisan ini sangat penulis harapkan. Akhir kata, semoga tulisan ini dapat memberikan manfaat dan memberikan wawasan bagi para pembacanya.

Denpasar, 2 Juni 2022

Penulis

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	ii
LEMBAR PERSETUJUAN TIM PENGUJI.....	iii
LEMBAR PERNYATAAN	iv
PRAKATA	v
ABSTRAK	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiii
DAFTAR RUMUS	xv
DAFTAR LAMPIRAN	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Penelitian	1
1.2 Identifikasi Masalah.....	7
1.3 Pembatasan Masalah.....	8
1.4 Rumusan Masalah.....	8
1.5 Tujuan	9
1.6 Manfaat Penelitian	10
BAB II KAJIAN PUSTAKA.....	11
2.1 Kajian Teori	11
2.1.1 Pelayanan Kesehatan.....	11
2.1.2 Analisis Sentimen.....	12
2.1.3 <i>Machine Learning</i>	13
2.1.4 <i>Support Vector Machine (SVM)</i>	16
2.1.5 SVM optimasi <i>One Against One</i> dan SVM optimasi <i>One Against All/Rest</i>	18
2.1.6 <i>Term Frequency – Inverse Document Frequency (TF-IDF)</i>	19
2.2 Penelitian Relevan.....	20
2.3 Proses Penelitian	23

2.3.1	Tujuan Penelitian	24
2.3.2	Studi Literatur	25
2.3.3	Pengumpulan <i>Dataset</i>	26
2.3.4	<i>Preprocessing Data</i>	26
2.3.5	Data <i>Training</i> dan Data <i>Testing</i>	28
2.3.6	<i>Support Vector Machine</i>	29
2.3.7	Evaluasi dan Analisis	29
BAB III METODE PENELITIAN		30
3.1	Sumber Data.....	31
3.2	Pengelolaan <i>Raw Dataset</i>	31
3.3	<i>Preprocessing Dataset</i>	32
3.4	Model SVM OvO dan OvR.....	36
3.5	Pembuatan Model.....	37
3.6	Evaluasi dan Pengujian Model.....	39
3.7	Analisis.....	39
BAB IV HASIL PENELITIAN.....		41
4.1	Pengolahan <i>Raw Dataset</i>	41
4.2	<i>Preprocessing Data</i>	43
4.2.1	<i>Casefolding</i>	43
4.2.2	<i>Cleaning</i>	43
4.2.3	<i>Tokenizing</i>	45
4.2.4	<i>Normalization</i>	46
4.2.5	<i>Stopwords</i>	47
4.2.6	<i>Stemming</i>	48
4.2.7	<i>Weighting</i>	51
4.3	Pembuatan Model SVM.....	52
4.3.1	SVM <i>One versus One</i> (OvO)	52
4.3.2	SVM <i>One versus Rest</i> (OvR)	59
4.4	Pengujian Model	62
4.4.1	SVM OvO	64
4.4.2	SVM OvR.....	65
4.5	Analisis Model	66

BAB V PENUTUP.....	73
5.1 Simpulan	73
5.2 Saran.....	74
DAFTAR PUSTAKA	75
LAMPIRAN.....	80



DAFTAR TABEL

Tabel 2.1 Daftar penelitian relevan.....	20
Tabel 3.1 Sampel <i>dataset</i> sentimen pelayanan kesehatan.....	31
Tabel 3.2 Sampel <i>dataset</i> sentimen dengan <i>label</i>	32
Tabel 3.3 Contoh proses <i>casefolding</i>	33
Tabel 3.4 Contoh proses <i>stemming</i>	34
Tabel 3.5 Contoh proses <i>tokenizing</i>	34
Tabel 3.6 Contoh proses <i>stopwords</i>	35
Tabel 3.7 Dataset setelah melalui tahapan <i>casefolding-stopwords</i>	38
Tabel 3.8 Proses pembobotan dengan metode TF-IDF <i>Unigram</i>	38
Tabel 3.9 Matriks vektor dari data sentimen.....	38
Tabel 3.10 Contoh matriks perbandingan.....	39
Tabel 4.1 <i>Encoding</i> label dari data teks ke angka.....	41
Tabel 4.2 Contoh dataset 3 kelas.....	53
Tabel 4.3 Fungsi <i>hyperplane</i> pendekatan OvO.....	54
Tabel 4.4 Hasil substitusi nilai x_1 dan x_2 ke persamaan <i>hyperplane</i> OvO	55
Tabel 4.5 Hasil <i>vote strategy</i> pada pendekatan OvO	56
Tabel 4.6 Fungsi <i>hyperplane</i> pendekatan OvR.....	59
Tabel 4.7 Hasil substitusi x_1 dan x_2 pada <i>hyperplane</i> OvR.....	60
Tabel 4.8 Prediksi kelas pada pendekatan OvR.....	61
Tabel 4.9 Hasil <i>Training</i> data dengan <i>KFold</i>	64
Tabel 4.10 Hasil validasi data dengan <i>KFold</i>	64
Tabel 4.11 Hasil <i>Training</i> data pendekatan OvR.....	65
Tabel 4.12 Hasil <i>test</i> data pendekatan OvR	66
Tabel 4.13 Perbandingan model SVM OvO <i>Unigram</i> dengan SVM OvR <i>Unigram</i>	67
Tabel 4.14 Perbandingan model SVM OvO <i>Bigram</i> dengan SVM OvR <i>Bigram</i>	68
Tabel 4.15 Perbandingan model SVM OvO <i>Unigram</i> dengan <i>Bigram</i>	69
Tabel 4.16 Perbandingan model SVM OvR <i>Unigram</i> dengan <i>Bigram</i>	70
Tabel 4.17 Hasil skor seluruh model SVM.....	71

DAFTAR GAMBAR

Gambar 2.1 Alternatif bidang pemisah (Nugroho et al., 2003)	17
Gambar 2.2 Bidang pemisah terbaik dengan <i>Margin</i> (M) terbesar (Nugroho et al., 2003)	17
Gambar 2.3 Alur kerangka berpikir	25
Gambar 2.4 Proses <i>casefolding</i> teks.....	26
Gambar 2.5 Proses <i>cleaning</i> teks	26
Gambar 2.6 Proses <i>tokenizing</i> teks.....	27
Gambar 2.7 Proses <i>normalization</i> teks.....	27
Gambar 2.8 Proses <i>stopwords</i> teks.....	28
Gambar 2.9 Proses <i>stemming</i> teks.....	28
Gambar 3.1 Tahapan penelitian	30
Gambar 3.2 Tahapan <i>preprocessing dataset</i>	32
Gambar 3.3 Alur kerja <i>machine learning</i>	37
Gambar 4.1 Isi data teks dalam <i>raw</i> dataset pelayanan kesehatan.....	41
Gambar 4.2 <i>Encoding</i> data teks label ke data angka.....	42
Gambar 4.3 Proses <i>casefolding</i> dataset.....	43
Gambar 4.4 Proses <i>cleaning</i> dataset.....	44
Gambar 4.5 Hasil proses <i>cleaning</i> dataset	45
Gambar 4.6 Proses <i>tokenizing</i> menggunakan <i>library</i> NLTK.....	46
Gambar 4.7 Proses <i>normalization</i>	47
Gambar 4.8 Proses <i>stopwrods</i> menggunakan <i>library</i> NLTK dan <i>Corpus</i> dari Tala	48
Gambar 4.9 Hasil proses <i>stopwords</i>	48
Gambar 4.10 Kode program untuk proses <i>stemming</i>	49
Gambar 4.11 Hasil proses <i>stemming</i>	50
Gambar 4.12 Hasil <i>preprocessing</i> data	51
Gambar 4.13 Fungsi pembobotan dengan TF-IDF	52
Gambar 4.14 Visualisasi dataset 3 kelas	53
Gambar 4.15 Visualisasi <i>hyperplane</i> pendekatan OvO	55

Gambar 4.16 Model SVM OvO	58
Gambar 4.17 Visualisasi <i>hyperplane</i> pendekatan OvR.....	60
Gambar 4.18 Model SVM OvR	62
Gambar 4.19 Ilustrasi pendekatan <i>KFold Cross Validation</i>	63
Gambar 4.20 <i>Source code KFold Cross Validation</i>	63
Gambar 4.21 Perbandingan SVM OvO <i>bigram</i> dan <i>unigram</i>	65
Gambar 4.22 Perbandingan OvR <i>Unigram</i> dan <i>Bigram</i>	66
Gambar 4.23 Visualisasi model SVM OvO <i>Unigram</i> dengan SVM OvR <i>Unigram</i>	67
Gambar 4.24 Visualisasi perbandingan SVM OvO <i>Bigram</i> dengan OvR <i>Bigram</i>	68
Gambar 4.25 Visualisasi perbandingan SVM OvO <i>Unigram</i> dengan <i>Bigram</i>	69
Gambar 4.26 Visualisasi perbandingan SVM OvR <i>Unigram</i> dengan <i>Bigram</i>	70
Gambar 4.27 Visualisasi seluruh skor model SVM.....	71



DAFTAR RUMUS

Persamaan 2.1 Persamaan jumlah <i>binary classification</i>	19
Persamaan 2.2 Persamaan penghitungan <i>term frequency</i>	19
Persamaan 2.3 Persamaan penghitungan <i>inverse document frequency</i>	19
Persamaan 4.4 Persamaan untuk mendapatkan <i>hyperplane</i>	54
Persamaan 4.5 Persamaan untuk menghitung nilai <i>w</i> dan <i>b</i>	54
Persamaan 4.6 Persamaan optimasi untuk mendapatkan SVM	54



DAFTAR LAMPIRAN

Lampiran 1 <i>Import Document</i> ke <i>library pandas</i>	80
Lampiran 2 <i>Encoding label</i>	80
Lampiran 3 Proses <i>casefolding</i>	81
Lampiran 4 Proses <i>cleaning</i>	82
Lampiran 5 Hasil proses <i>cleaning</i>	83
Lampiran 6 Proses <i>tokenizing</i>	83
Lampiran 7 Proses <i>normalization</i>	84
Lampiran 8 Proses <i>stopwords</i>	85
Lampiran 9 Hasil proses <i>stopwords</i>	85
Lampiran 10 Proses <i>stemming</i>	86
Lampiran 11 Hasil proses <i>stemming</i>	87
Lampiran 12 Hasil proses <i>casefolding</i> hingga <i>stemming</i>	88
Lampiran 13 Fungsi pembobotan dengan TF-IDF.....	88
Lampiran 14 Fungsi <i>KFold Crossvalidation</i>	88
Lampiran 15 Fungsi model SVM OvO.....	89
Lampiran 16 Fungsi model SVM OvR.....	89