



SKRIPSI

***ASPECT-BASED SENTIMENT ANALYSIS PADA
ULASAN APLIKASI ACCESS BY KAI DENGAN
MENGUNAKAN METODE TF-IDF DAN
ALGORITMA SUPPORT VECTOR MACHINE***

MUHAMMAD NUR RACHMAN NIDHI SURYONO
NPM 21082010137

DOSEN PEMBIMBING

Amalia Anjani Arifiyanti, S.Kom., M.Kom.
Dhian Satria Yudha Kartika, S.Kom., M.Kom.

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

LEMBAR PENGESAHAN

ASPECT-BASED SENTIMENT ANALYSIS PADA ULASAN APLIKASI ACCESS BY KAI DENGAN MENGGUNAKAN METODE TF-IDF DAN ALGORITMA SUPPORT VECTOR MACHINE

Oleh:

MUHAMMAD NUR RACHMAN NIDHI SURYONO
NPM. 21082010137

Telah dipertahankan dihadapan dan diterima oleh Tim Penguji Skripsi Prodi Sistem Informasi
Fakultas Ilmu Komputer Universitas Pembangunan Nasional Veteran Jawa Timur Pada tanggal
4 Juni 2025

Menyetujui

Amalia Anjani Arifiyanti, S.Kom., M.Kom.
NIP. 19920812 2018032 001


..... (Pembimbing I)

Dhian Satria Yudha Kartika, S.Kom., M.Kom.
NPT. 20119860522249


..... (Pembimbing II)

Rizka Hadiwiyanti, S.Kom., M.Kom., MBA.
NIP. 19860727 2018032 001


..... (Ketua Penguji)

Anindo Saka Fitri, S.Kom., M. Kom.
NIP. 19930325 2024062 001


..... (Anggota Penguji II)

Iqbal Ramadhani Mukhlis, S.Kom., M. Kom.
NIP. 19930305 2024061 002


..... (Anggota Penguji III)

Mengetahui,

Dekan Fakultas Ilmu Komputer


Prof. Dr. Ir. Novirina Hendrasarie, MT

NIP. 19681126 199403 2 001

LEMBAR PERSETUJUAN

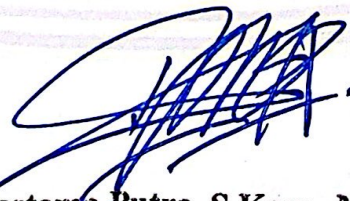
***ASPECT-BASED SENTIMENT ANALYSIS PADA ULASAN APLIKASI
ACCESS BY KAI DENGAN MENGGUNAKAN METODE TF-IDF DAN
ALGORITMA SUPPORT VECTOR MACHINE***

Oleh:

**MUHAMMAD NUR RACHMAN NIDHI SURYONO
NPM. 21082010137**

Menyetujui,

**Koordinator Program Studi Sistem Informasi
Fakultas Ilmu Komputer**



**Agung Brastama Putra, S.Kom., M.Kom.
NIP. 19851124 202121 1 003**

SURAT PERNYATAAN BEBAS PLAGIASI

Yang bertandatangan di bawah ini:

Nama Mahasiswa : MUHAMMAD NUR RACHMAN NIDHI SURYONO
NPM : 21082010137
Program : Sarjana(S1)
Program Studi : Sistem Informasi
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 daftar pustaka.

Dan saya menyatakan bahwa dokumen ilmiah ini bebas dari unsur-unsur plagiasi. Apabila dikemudian hari ditemulan indikasi plagiat pada Skripsi ini, saya bersedia menerima sanksi sesuai 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, 10 Juni 2025

Yang Membuat Pernyataan,



MUHAMMAD NUR RACHMAN NIDHI SURYONO

NPM. 21082010137

ABSTRAK

Nama Mahasiswa / NPM : Muhammad Nur Rachman Nidhi Suryono / 21082010137
Judul Skripsi : *Aspect-Based Sentiment Analysis* Pada Ulasan Aplikasi Access By Kai Dengan Menggunakan Metode TF-IDF Dan Algoritma *Support Vector Machine*
Dosen Pembimbing : 1. Amalia Anjani Arifiyanti, S.Kom., M.Kom.
2. Dhian Satria Yudha Kartika, S.Kom., M.Kom.

Access by KAI merupakan aplikasi layanan transportasi digital dari PT Kereta Api Indonesia yang mempermudah pengguna dalam mengakses layanan perjalanan kereta api. Untuk meningkatkan kualitas layanan dan pengalaman pengguna, penelitian ini melakukan analisis sentimen terhadap ulasan pengguna aplikasi *Google Play Store* menggunakan algoritma *Support Vector Machine* (SVM). Tiga aspek utama yang dianalisis yaitu *Financial Transactions*, *Technical Issues and Performance*, serta *User Experience and Interface*. Hasil terbaik untuk aspek *Financial Transactions* diperoleh dari model SMOTE dengan *kernel Linear* dan rasio 70:30, menghasilkan rata-rata *f1-score* sebesar 0.66 (*f1-score* -1: 0.58, *f1-score* 0: 0.95, *f1-score* 1: 0.46). Untuk *Technical Issues and Performance*, model terbaik adalah Non-SMOTE dengan *kernel Linear* dan rasio 80:20, dengan rata-rata *f1-score* sebesar 0.76 (*f1-score* -1: 0.94, *f1-score* 0: 0.76, *f1-score* 1: 0.59). Sedangkan untuk *User Experience and Interface*, model SMOTE dengan *kernel Linear* dan rasio 70:30 memberikan *f1-score* rata-rata tertinggi yaitu 0.76 (*f1-score* -1: 0.47, *f1-score* 0: 0.89, *f1-score* 1: 0.92). Model terbaik ini diimplementasikan dalam aplikasi web berbasis *Flask* yang dapat memprediksi sentimen, mengeksport hasil dalam bentuk *.csv*, serta menampilkan visualisasi data. Penelitian ini diharapkan membantu PT KAI dalam memahami persepsi pengguna dan meningkatkan layanan Access by KAI.

Kata Kunci: Access by KAI, Analisis Sentimen Berbasis Aspek, BERTopic, *Support Vector Machine*, *Flask*

ABSTRACT

Student Name / NPM : Muhammad Nur Rachman Nidhi Suryono / 21082010137
Thesis Title : Aspect-Based Sentiment Analysis on Access By Kai Application Reviews Using TF-IDF Method and Support Vector Machine Algorithm
Advisors : 1. Amalia Anjani Arifiyanti, S.Kom., M.Kom.
2. Dhian Satria Yudha Kartika, S.Kom., M.Kom.

Access by KAI is a digital transportation service application developed by PT Kereta Api Indonesia to simplify users' access to train travel services. To improve service quality and user experience, this study conducted sentiment analysis on user reviews from the Google Play Store using the Support Vector Machine (SVM) algorithm. The analysis focused on three key aspects: Financial Transactions, Technical Issues and Performance, and User Experience and Interface. The best result for Financial Transactions was achieved using the SMOTE model with a Linear kernel and a 70:30 data split, yielding an average f1-score of 0.66 (f1-score -1: 0.58, f1-score 0: 0.95, f1-score 1: 0.46). For Technical Issues and Performance, the highest performance came from the Non-SMOTE model with a Linear kernel and an 80:20 data split, resulting in an average f1-score of 0.76 (f1-score -1: 0.94, f1-score 0: 0.76, f1-score 1: 0.59). Meanwhile, the User Experience and Interface aspect achieved its best result with the SMOTE model using a Linear kernel and a 70:30 split, with an average f1-score of 0.76 (f1-score -1: 0.47, f1-score 0: 0.89, f1-score 1: 0.92). The best-performing model was then implemented into a Flask-based web application capable of predicting sentiment, exporting results in .csv format, and displaying data visualizations. This research aims to help PT KAI better understand user perceptions and enhance the Access by KAI service.

Keywords: Access by KAI, Aspect-Based Sentiment Analysis, BERTopic, Flask, Support Vector Machine

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kehadirat Allah SWT yang telah melimpahkan segala rahmat-Nya, serta sholawat dan salam tercurahkan kepada junjungan Nabi Besar Muhammad SAW, sehingga penulis dapat menyelesaikan skripsi ini dengan berjudul “*ASPECT-BASED SENTIMENT ANALYSIS PADA ULASAN APLIKASI ACCESS BY KAI DENGAN MENGGUNAKAN METODE TF-IDF DAN ALGORITMA SUPPORT VECTOR MACHINE*” sebagai salah satu syarat dalam menyelesaikan Program Studi S1 Sistem Informasi di Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Penyusunan skripsi ini tentunya tidak lepas dari bantuan, dukungan, serta bimbingan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu kelancaran penulisan tugas akhir ini, khususnya kepada:

1. Kedua orang tua dan Abang yang selalu memberikan doa, dukungan moral maupun materi, serta menjadi sumber semangat yang tak ternilai selama proses perkuliahan hingga penyusunan skripsi ini.
2. Ibu Amalia Anjani Arifiyanti, S.Kom., M.Kom., selaku Dosen Pembimbing I yang telah meluangkan waktu, memberikan bimbingan, serta arahan selama proses penyusunan skripsi ini.
3. Bapak Dhian Satria Yudha Kartika, S.Kom., M.Kom., selaku Dosen Pembimbing II yang juga dengan sabar memberikan waktu, masukan, dan arahan dalam penyelesaian skripsi ini.
4. Bapak Agung Brastama Putra, S.Kom., M.Kom., selaku Koordinator Program Studi Sistem Informasi Universitas Pembangunan Nasional “Veteran” Jawa Timur atas dukungan dan arahnya selama masa studi.
5. Ibu Prof. Dr. Ir. Novirina Hendrasarie, M.T., selaku Dekan Fakultas Ilmu Komputer, yang telah memberikan arahan, dukungan, serta fasilitas yang mendukung kelancaran proses studi dan penyusunan skripsi ini. Terima kasih atas dedikasi dan kepemimpinan Ibu yang menginspirasi.

6. Seluruh dosen Program Studi Sistem Informasi yang telah membagikan ilmu, pengalaman, serta wawasan yang sangat berarti selama masa perkuliahan.
7. Putrila Maya Jannata, sebagai kekasih tercinta yang telah menjadi sosok penting selama proses penyusunan skripsi ini. Terima kasih atas dukungan, kesabaran dalam menghadapi segala keluh kesah, serta kehadiranmu yang selalu menjadi penguat di tengah tekanan. Kehangatan, perhatian, dan motivasimu telah membantu penulis menyelesaikan skripsi ini dengan lebih ringan dan penuh makna.
8. Rekan-rekan mahasiswa Sistem Informasi angkatan 2021 yang selalu memberikan semangat, dukungan, serta berbagi informasi dan pengalaman selama menyusun tugas hingga skripsi ini.
9. Seluruh teman-teman Program Studi Sistem Informasi yang tidak dapat disebutkan satu per satu, namun telah memberikan bantuan, motivasi, maupun dukungan dalam proses penyusunan skripsi ini.

Penulis menyadari bahwa skripsi ini masih memiliki kekurangan, baik dari segi isi maupun penyajian. Oleh karena itu, segala kritik dan saran yang membangun sangat penulis harapkan demi perbaikan ke depan. Semoga skripsi ini dapat bermanfaat dan menjadi referensi yang berguna, khususnya dalam bidang analisis sentimen.

Surabaya, 28 Mei 2025

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN.....	iii
LEMBAR PERSETUJUAN	v
SURAT PERNYATAAN BEBAS PLAGIASI	vii
ABSTRAK.....	ix
ABSTRACT	xi
KATA PENGANTAR.....	xiii
DAFTAR ISI	xv
DAFTAR TABEL	xix
DAFTAR GAMBAR.....	xxi
DAFTAR SINGKATAN, ISTILAH DAN ARTI SIMBOL	xxv
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 Penelitian.....	4
1.6. Sistematika Penelitian.....	5
BAB II TINJAUAN PUSTAKA	7
2.1. Landasan Teori.....	7
2.1.1. <i>Google Play Store</i>	7
2.1.2. <i>Access By KAI</i>	7
2.1.3. <i>Analisis Sentimen</i>	7
2.1.4. <i>Analisis Sentimen Text Mining</i>	8
2.1.5. <i>Aspect-Based Sentiment Analysis (ABSA)</i>	8
2.1.6. <i>Python</i>	8
2.1.7. <i>Text Preprocessing</i>	9
2.1.8. <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i>	10
2.1.9. <i>SMOTE</i>	11
2.1.10. <i>Support Vector Machine</i>	11
2.1.11. <i>BERTopic</i>	13
2.1.12. <i>Confusion Matrix</i>	13

2.1.13. <i>Flask</i>	15
2.2. Penelitian Terdahulu.....	15
BAB III METODOLOGI PENELITIAN	19
3.1. Alur Penelitian	19
3.2. Studi Literatur	19
3.3. Pengumpulan Data	20
3.4. <i>Data Preprocessing</i>	20
3.4.1. <i>Case-folding</i>	20
3.4.2. <i>Cleaning</i>	20
3.4.3. <i>Normalization</i>	21
3.4.4. <i>Tokenizing</i>	21
3.4.5. <i>Stopword Removal</i>	21
3.4.6. <i>Stemming</i>	21
3.5. <i>Topic Modelling</i>	22
3.6. Pelebelan Aspek dan Sentimen.....	22
3.7. Pembobotan TF-IDF.....	22
3.8. Pelatihan Model	26
3.9. Evaluasi Model	27
3.10. <i>Model Deployment</i>	27
BAB IV HASIL DAN PEMBAHASAN.....	31
4.1. Pengumpulan Data	31
4.2. <i>Data Preprocessing</i>	32
4.2.1. <i>Case-folding</i>	32
4.2.2. <i>Cleaning</i>	33
4.2.3. <i>Normalization</i>	34
4.2.4. <i>Tokenizing</i>	35
4.2.5. <i>Stopword Removal</i>	36
4.2.6. <i>Stemming</i>	37
4.3. <i>Topic Modelling</i>	38
4.4. <i>Labeling Aspect and Sentiment</i>	46
4.5. <i>Exploratory Data Analysis (EDA)</i>	53
4.6. Pembobotan TF-IDF.....	57

4.7. Pembagian Data	58
4.8. Pelatihan Model	60
4.9. Evaluasi Model	62
4.10. <i>Model Deployment</i>	104
BAB V PENUTUP	115
5.1. Kesimpulan	115
5.2. Saran Pengembangan	116
DAFTAR PUSTAKA.....	117

DAFTAR TABEL

Tabel 2. 1 <i>Confusion Matrix</i>	14
Tabel 2. 2 Penelitian Terdahulu.....	15
Table 3. 1 Contoh Hasil TF-IDF	25
Table 3. 2 Skenario Pengujian.....	26
Tabel 4. 1 Contoh <i>Case-folding</i>	32
Tabel 4. 2 Contoh <i>Cleaning</i>	33
Tabel 4. 3 Contoh <i>Normalization</i>	34
Tabel 4. 4 Contoh <i>Tokenizing</i>	35
Tabel 4. 5 Contoh <i>Stopword Removal</i>	36
Tabel 4. 6 Contoh <i>Stemming</i>	37
Tabel 4. 7 Pengelompokkan Topik.....	44
Tabel 4. 8 Kata Kunci Topik.....	45
Tabel 4. 9 Hasil Perbandingan <i>F1-Score Financial Transaction</i>	98
Tabel 4. 10 Hasil Perbandingan <i>F1-Score Technical Issues and Performance</i>	99
Tabel 4. 11 Hasil Perbandingan <i>F1-Score User Experience and Interface</i>	100

DAFTAR GAMBAR

Gambar 2. 1 Alur <i>Text Preprocessing</i>	9
Gambar 3. 1 Alur Metode Skripsi.....	19
Gambar 3. 2 <i>Wireframe Dashboard Aplikasi 1</i>	28
Gambar 3. 3 <i>Wireframe Dashboard Aplikasi 2</i>	29
Gambar 4. 1 <i>Source Code Scraping Data</i>	31
Gambar 4. 2 <i>Source Code Case-folding</i>	32
Gambar 4. 3 <i>Source Code Cleaning</i>	33
Gambar 4. 4 <i>Source Code Normalization</i>	34
Gambar 4. 5 <i>Source Code Tokenizing</i>	35
Gambar 4. 6 <i>Source Code Stopword Removal</i>	36
Gambar 4. 7 <i>Source Code Stemming</i>	37
Gambar 4. 8 <i>Install Package</i>	38
Gambar 4. 9 <i>Data Preparation</i>	38
Gambar 4. 10 <i>Pre-calculate Embeddings</i>	39
Gambar 4. 11 <i>Additional Representations</i>	40
Gambar 4. 12 <i>Component Settings</i>	41
Gambar 4. 13 <i>Pipeline & Clustering Code</i>	42
Gambar 4. 14 <i>Topic Result</i>	43
Gambar 4. 15 <i>Hierarchial Structure</i>	43
Gambar 4. 16 <i>Source Code Data Preparation</i>	46
Gambar 4. 17 <i>Source Code Labeling Aspect 1</i>	47
Gambar 4. 18 <i>Source Code Labeling Aspect 2</i>	48
Gambar 4. 19 Hasil Pelebelan Otomatis Aspek & Sentimen	49
Gambar 4. 20 <i>Source Code Default Sentiment</i>	51
Gambar 4. 21 <i>Labeling Aspect and Sentiment Result</i>	52
Gambar 4. 22 Hasil Persebaran Sentimen.....	52
Gambar 4. 23 Hapus Komentar Tidak Bersentimen.....	52
Gambar 4. 24 <i>Source Code Pie Chart</i>	53
Gambar 4. 25 <i>Sentiment Distribution Financial Transaction</i>	54
Gambar 4. 26 <i>Sentiment Distribution Technical Issues and Performance</i>	54
Gambar 4. 27 <i>Sentiment Distribution User Experience and Interface</i>	55

Gambar 4. 28 <i>Wordcloud Positive Sentiments</i>	56
Gambar 4. 29 <i>Wordcloud Negative Sentiments</i>	56
Gambar 4. 30 <i>Source Code TF-IDF</i>	57
Gambar 4. 31 <i>Contoh Hasil TF-IDF</i>	58
Gambar 4. 32 <i>Potongan Source Code Split Data 80:20</i>	59
Gambar 4. 33 <i>Potongan Source Code Split Data 70:30</i>	59
Gambar 4. 34 <i>Source Code SVM Non-SMOTE</i>	60
Gambar 4. 35 <i>Source Code SVM SMOTE</i>	61
Gambar 4. 36 <i>Hasil Skenario 1 Financial Transactions</i>	62
Gambar 4. 37 <i>Hasil Skenario 1 Technical Issues and Performance</i>	63
Gambar 4. 38 <i>Hasil Skenario 1 User Experience and Interface</i>	64
Gambar 4. 39 <i>Hasil Skenario 2 Financial Transactions</i>	65
Gambar 4. 40 <i>Hasil Skenario 2 Technical Issues and Performance</i>	66
Gambar 4. 41 <i>Hasil Skenario 2 User Experience and Interface</i>	67
Gambar 4. 42 <i>Hasil Skenario 3 Financial Transactions</i>	68
Gambar 4. 43 <i>Hasil Skenario 3 Technical Issues and Performance</i>	69
Gambar 4. 44 <i>Hasil Skenario 3 User Experience and Interface</i>	70
Gambar 4. 45 <i>Hasil Skenario 4 Financial Transactions</i>	71
Gambar 4. 46 <i>Hasil Skenario 4 Technical Issues and Performance</i>	72
Gambar 4. 47 <i>Hasil Skenario 4 User Experience and Interface</i>	73
Gambar 4. 48 <i>Hasil Skenario 5 Financial Transactions</i>	74
Gambar 4. 49 <i>Hasil Skenario 5 Technical Issues and Performance</i>	75
Gambar 4. 50 <i>Hasil Skenario 5 User Experience and Interface</i>	76
Gambar 4. 51 <i>Hasil Skenario 6 Financial Transactions</i>	77
Gambar 4. 52 <i>Hasil Skenario 6 Technical Issues and Performance</i>	78
Gambar 4. 53 <i>Hasil Skenario 6 User Experience and Interface</i>	79
Gambar 4. 54 <i>Hasil Skenario 7 Financial Transactions</i>	80
Gambar 4. 55 <i>Hasil Skenario 7 Technical Issues and Performance</i>	81
Gambar 4. 56 <i>Hasil Skenario 7 User Experience and Interface</i>	82
Gambar 4. 57 <i>Hasil Skenario 8 Financial Transactions</i>	83
Gambar 4. 58 <i>Hasil Skenario 8 Technical Issues and Performance</i>	84
Gambar 4. 59 <i>Hasil Skenario 8 User Experience and Interface</i>	85

Gambar 4. 60 Hasil Skenario 9 <i>Financial Transactions</i>	86
Gambar 4. 61 Hasil Skenario 9 <i>Technical Issues and Performance</i>	87
Gambar 4. 62 Hasil Skenario 9 <i>User Experience and Interface</i>	88
Gambar 4. 63 Hasil Skenario 10 <i>Financial Transactions</i>	89
Gambar 4. 64 Hasil Skenario 10 <i>Technical Issues and Performance</i>	90
Gambar 4. 65 Hasil Skenario 10 <i>User Experience and Interface</i>	91
Gambar 4. 66 Hasil Skenario 11 <i>Financial Transactions</i>	92
Gambar 4. 67 Hasil Skenario 11 <i>Technical Issues and Performance</i>	93
Gambar 4. 68 Hasil Skenario 11 <i>User Experience and Interface</i>	94
Gambar 4. 69 Hasil Skenario 12 <i>Financial Transactions</i>	95
Gambar 4. 70 Hasil Skenario 12 <i>Technical Issues and Performance</i>	96
Gambar 4. 71 Hasil Skenario 12 <i>User Experience and Interface</i>	97
Gambar 4. 72 Persebaran <i>Data Validasi</i>	101
Gambar 4. 73 <i>Confusion Matrix Financial Transactions</i>	101
Gambar 4. 74 <i>Confusion Matrix Technical Issues and Performance</i>	102
Gambar 4. 75 <i>Confusion Matrix User Experience and Interface</i>	103
Gambar 4. 76 <i>Source Code Backend Environment</i>	104
Gambar 4. 77 <i>Source Code Input Word Predict</i>	105
Gambar 4. 78 <i>Source Code Input File Predict 1</i>	106
Gambar 4. 79 <i>Source Code Input File Predict 2</i>	107
Gambar 4. 80 Tampilan Halaman Index.....	108
Gambar 4. 81 <i>Source Code Text Input</i>	108
Gambar 4. 82 <i>Source Code Prediksi Text</i>	109
Gambar 4. 83 <i>Input file CSV</i>	110
Gambar 4. 84 Grafik Distribusi Aspek Per Sentimen.....	110
Gambar 4. 85 <i>Wordcloud Sentimen</i>	111
Gambar 4. 86 Tabel Komentar Hasil Analisis	112
Gambar 4. 87 <i>Source Code Input File CSV</i>	112
Gambar 4. 88 <i>Source Code Wordcloud</i>	112
Gambar 4. 89 <i>Source Code Menampilkan Tabel</i>	113

DAFTAR SINGKATAN, ISTILAH DAN ARTI SIMBOL

SINGKATAN/ISTILAH/ SIMBOL	ARTI DAN KETERANGAN
$f_{(t,d)}$	Jumlah kemunculan kata t dalam dokumen d .
T_d	Total kata dalam dokumen d .
$TF(t,d)$	Nilai frekuensi term: $f_{(t,d)} / T_d$.
IDF	<i>Inverse Document Frequency</i>
N	Jumlah total dokumen dalam korpus.
n_t	Jumlah dokumen yang mengandung kata t .
γ (<i>gamma</i>)	Parameter yang menentukan pengaruh satu data terhadap lainnya dalam RBF kernel.
C	Konstanta dalam kernel polinomial
d	Derajat polinomial dalam <i>kernel polynomial</i>
TP	Kasus positif yang berhasil diprediksi benar oleh model
TN	Kasus negatif yang berhasil diprediksi benar oleh model
FP	Kasus negatif yang salah diprediksi sebagai positif
FN	Kasus positif yang salah diprediksi sebagai negatif