

ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI WONDR MENGGUNAKAN *HYBRID* RFSVM MODEL

Skripsi

**Disusun untuk memenuhi salah satu syarat
memperoleh gelar Sarjana Matematika**



**Rifa Khairunnisa Pratiwi
1305621004**

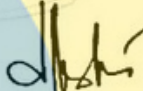
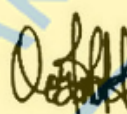
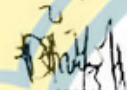
**PROGRAM STUDI MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI JAKARTA**

2026

LEMBAR PENGESAHAN UJIAN SKRIPSI
ANALISIS SENTIMEN ULASAN PENGGUNA APLIKASI WONDR
MENGGUNAKAN *HYBRID* RFSVM MODEL

Nama : Rifa Khairunnisa Pratiwi

NIM : 1305621004

	Nama	Tanda Tangan	Tanggal
Penanggung Jawab			
Dekan :	Dr. Hadi Nasbey, S.Pd., M.Si. NIP. 197909162005011004		26/01 2026
Wakil Penanggung Jawab			
Wakil Dekan I :	Dr. Meiliasari, S.Pd., M.Sc. NIP. 197905042009122002		26/01 2026
Ketua :	Dr. Sudarwanto, M.Si, DEA. NIP. 196503251993031003		14/01 2026
Penguji Ahli :	Ibnu Hadi, M.Si. NIP. 198107182008011017		14/01 2026
Sekretaris :	Dr. Yudi Mahatma, M.Si. NIP. 197610202008121001		14/01 2026
Pembimbing I :	Qorry Meidianingsih, S.Si., M.Si. NIP. 199105192019032019		15/01 2026
Pembimbing II :	Devi Eka Wardani M., S.Pd., M.Si. NIP. 199005162019032014		19/01 2026

Dinyatakan lulus ujian skripsi pada tanggal 22 Desember 2025

SURAT PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini, mahasiswa Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta:

Nama : Rifa Khairunnisa Pratiwi
No. Registrasi : 1305621004
Program Studi : Matematika

Dengan ini menyatakan bahwa skripsi yang saya buat dengan judul
"Analisis Sentimen Ulasan Pengguna Aplikasi Wondr Menggunakan Hybrid RFSVM Model"

1. Dibuat sendiri, mengadopsi hasil kuliah, buku-buku, dan referensi acuan yang tertera di dalam daftar pustaka pada skripsi saya.
2. Bukan merupakan hasil duplikasi skripsi yang telah dipublikasikan atau pernah dipakai untuk mendapatkan gelar sarjana di universitas lain, kecuali pada bagian-bagian sumber informasi yang dicantumkan berdasarkan tata cara referensi yang semestinya.

Pernyataan ini dibuat dengan sesungguhnya dan saya bersedia menanggung segala akibat yang timbul jika pernyataan saya tidak benar.

Jakarta, 9 Desember 2025



Rifa Khairunnisa Pratiwi



KEMENTERIAN PENDIDIKAN TINGGI, SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
PERPUSTAKAAN DAN KEARSIPAN

Jalan Rawamangun Muka Jakarta 13220

Telepon/Faksimili: 021-4894221

Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Rifa Khairunnisa Pratiwi
NIM : 1305621004
Fakultas/Prodi : FMIPA/Matematika
Alamat email : pratiwirifa@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan dan Kearsipan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

☒ Skripsi ☐ Tesis ☐ Disertasi ☐ Lain-lain (.....)

yang berjudul :

Analisis Sentimen Ulasan Pengguna Aplikasi Wondr Menggunakan

Hybrid RFSVM Model

Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan dan Kearsipan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya, dan menampilkan/mempublikasikannya di internet atau media lain secara **fulltext** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 5 Februari 2026

Penulis

(Rifa Khairunnisa Pratiwi)
nama dan tanda tangan

ABSTRAK

RIFA KHAIRUNNISA PRATIWI. Analisis Sentimen Ulasan Pengguna Aplikasi Wondr Menggunakan *Hybrid* RFSVM Model. Skripsi, Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta. Desember 2025.

Wondr merupakan *platform mobile banking* BNI yang membantu nasabah mengelola keuangan dengan efisien. Untuk menilai kualitas layanan dan kepuasan pengguna, dilakukan analisis sentimen terhadap ulasan pengguna. Penelitian ini menerapkan metode *hybrid* yang menggabungkan algoritma *Random Forest* (RF) dan *Support Vector Machine* (SVM), disebut *hybrid* RFSVM. Tujuan penelitian ini adalah menjelaskan proses penggabungan model serta mengukur kinerjanya dalam klasifikasi ulasan aplikasi Wondr. Data diambil dari Google Play Store pada 8 Maret hingga 30 April 2025. Hasil menunjukkan akurasi 95% dengan nilai presisi untuk kelas negatif, netral, dan positif masing-masing 82%, 94%, dan 97%. Nilai *recall* untuk kelas negatif, positif, dan netral berturut-turut sebesar 85%, 99%, dan 58%, sedangkan *F1-score* masing-masing 84%, 98%, dan 72%. Berdasarkan hasil klasifikasi diperoleh 5.419 ulasan positif, 749 ulasan negatif, dan 137 ulasan netral. Secara keseluruhan, mayoritas pengguna memiliki persepsi positif terhadap aplikasi Wondr, meskipun masih terdapat ulasan negatif dan netral yang menunjukkan perlunya peningkatan pada beberapa aspek layanan.

Kata kunci. *Analisis Sentimen, Random Forest, Support Vector Machine, Hybrid RFSVM, Wondr*

ABSTRACT

RIFA KHAIRUNNISA PRATIWI. Sentiment Analysis of User Reviews of the Wondr Application Using a Hybrid RFSVM Model. Thesis, Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Jakarta State University. December 2025.

Wondr is BNI's mobile banking platform that helps customers manage their finances efficiently. To assess service quality and user satisfaction, sentiment analysis was conducted on user reviews. This study applied a hybrid method that combines the Random Forest (RF) and Support Vector Machine (SVM) algorithms, called hybrid RFSVM. The purpose of this study is to explain the model combination process and measure its performance in classifying Wondr app reviews. Data was collected from the Google Play Store from March 8 to April 30, 2025. The results show an accuracy of 95% with precision values for negative, neutral, and positive classes of 82%, 94%, and 97%, respectively. The recall values for the negative, positive, and neutral classes were 85%, 99%, and 58%, respectively, while the F1-scores were 84%, 98%, and 72%, respectively. Based on the classification results, there were 5,419 positive reviews, 749 negative reviews, and 137 neutral reviews. Overall, the majority of users had a positive perception of the Wondr app, although there were still negative and neutral reviews that indicated the need for improvement in several aspects of the service.

Keywords. *Sentiment Analysis, Random Forest, Support Vector Machine, hybrid RFSVM, Wondr*

KATA PENGANTAR

Puji syukur senantiasa penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat, karunia, dan petunjuk-Nya sehingga skripsi yang berjudul “Analisis Sentimen Ulasan Pengguna Aplikasi Wondr Menggunakan *Hybrid* RFSVM Model” dapat diselesaikan dengan baik. Skripsi ini disusun untuk memenuhi salah satu syarat dalam menyelesaikan tugas akhir pada Program Studi Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta.

Penyusunan skripsi ini tidak terlepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, disampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Ibu Qorry Meidianingsih, S.Si., M.Si. selaku dosen pembimbing I yang telah memberikan bimbingan, masukan, serta arahan dengan penuh kesabaran selama proses penyusunan skripsi ini.
2. Ibu Devi Eka Wardani Meganingtyas, S.Pd., M.Si. selaku dosen pembimbing II yang telah memberikan banyak saran, motivasi, dan koreksi yang sangat membantu dalam penyempurnaan penelitian ini.
3. Bapak/Ibu Dosen dan seluruh staf Program Studi Matematika FMIPA UNJ, atas ilmu dan bantuan yang telah diberikan selama masa perkuliahan.
4. Orang tua dan keluarga penulis yang selalu memberikan doa, dukungan moral, dan kasih sayang yang tiada henti.
5. Teman-teman seperjuangan Program Studi Matematika angkatan 2021, atas semangat, kebersamaan, serta bantuan selama proses perkuliahan hingga penyusunan skripsi ini.

6. Semua pihak yang tidak dapat disebutkan satu per satu, yang telah membantu dalam penyelesaian skripsi ini.

Disadari bahwa skripsi ini masih memiliki banyak kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi penyempurnaan karya ilmiah ini di masa yang akan datang. Akhir kata, semoga skripsi ini dapat memberikan manfaat bagi pembaca serta menjadi tambahan wawasan dalam pengembangan ilmu pengetahuan, khususnya di bidang analisis data dan pembelajaran mesin.

Jakarta, 9 Desember 2025



Rifa Khairunnisa Pratiwi



UNIVERSITAS NEGERI JAKARTA

DAFTAR ISI

LEMBAR PENGESAHAN	i
LEMBAR PERNYATAAN	ii
ABSTRAK	iii
ABSTRACT	iv
KATA PENGANTAR	v
DAFTAR ISI	vii
DAFTAR TABEL	ix
DAFTAR GAMBAR	xi
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	5
1.3 Tujuan Penelitian	6
1.4 Batasan Masalah	6
1.5 Manfaat Penelitian	6
BAB II LANDASAN TEORI	8
2.1 Analisis Sentimen	8
2.2 Ulasan Pengguna pada Google Play Store	9
2.3 Web Scraping pada Google Play Store	10
2.4 Text Mining	13
2.5 Data Preparation	15
2.6 Klasifikasi Teks Menggunakan Algoritma <i>Machine Learning</i> . .	23
2.6.1 <i>Decision Tree</i>	24

2.6.2	<i>Random Forest</i>	38
2.6.3	<i>Support Vector Machine (SVM)</i>	50
2.6.4	<i>Random Forest Support Vector Machine (RFSVM)</i>	73
2.7	Evaluasi Kinerja Model RFSVM	77
BAB III METODOLOGI PENELITIAN		82
3.1	Jenis Penelitian	82
3.2	Sumber dan Teknik Pengumpulan Data	82
3.3	Alur Penelitian	82
3.4	Diagram Alir	84
BAB IV HASIL DAN PEMBAHASAN		86
4.1	<i>Web Scraping</i>	86
4.2	<i>Labelling</i>	87
4.3	<i>Text Preprocessing</i>	89
4.3.1	<i>Case Folding</i>	89
4.3.2	<i>Cleaning</i>	89
4.3.3	<i>Tokenizing</i>	90
4.3.4	<i>Normalize</i>	91
4.3.5	<i>Stopword Removal</i>	91
4.3.6	<i>Stemming</i>	92
4.4	Data Split	92
4.5	Pembobotan Kata	93
4.6	Klasifikasi RFSVM	101
4.6.1	<i>Random Forest</i>	101
4.6.2	<i>Support Vector Machine (SVM)</i>	114
4.7	Evaluasi Kinerja	124
4.8	Interpretasi	127
BAB V KESIMPULAN DAN SARAN		132
5.1	Kesimpulan	132
5.2	Saran	133
DAFTAR PUSTAKA		135
LAMPIRAN		140
RIWAYAT HIDUP		161

DAFTAR TABEL

Tabel 2.1	Data Sebelum dan Sesudah <i>Case Folding</i>	16
Tabel 2.2	Data Sebelum dan Sesudah <i>Cleaning</i>	16
Tabel 2.3	Data Sebelum dan Sesudah <i>Tokenizing</i>	17
Tabel 2.4	Data Sebelum dan Sesudah <i>Normalize</i>	17
Tabel 2.5	Data Sebelum dan Sesudah <i>Stopword Removal</i>	17
Tabel 2.6	Data Sebelum dan Sesudah <i>Stemming</i>	18
Tabel 2.7	Contoh Data Ulasan untuk Pembobotan Kata TF-IDF	19
Tabel 2.8	Hasil Perhitungan TF	20
Tabel 2.9	Hasil Perhitungan IDF	21
Tabel 2.10	Hasil Perhitungan TF-IDF	22
Tabel 2.11	Contoh Data Ulasan untuk Klasifikasi dengan <i>Decision Tree</i>	32
Tabel 2.12	Nilai TF-IDF Data Ulasan untuk Klasifikasi <i>Decision Tree</i>	33
Tabel 2.13	Hasil Evaluasi Kandidat <i>Threshold</i>	35
Tabel 2.14	<i>Threshold</i> Terbaik pada Setiap Fitur Kata	36
Tabel 2.15	Contoh Data Ulasan untuk Klasifikasi dengan <i>Random Forest</i>	44
Tabel 2.16	Contoh Data Ulasan untuk Prediksi dengan <i>Random Forest</i>	47
Tabel 2.17	Nilai TF-IDF Data Ulasan untuk Prediksi	47
Tabel 2.18	Contoh Data Ulasan untuk Klasifikasi dengan SVM	68
Tabel 2.19	Nilai TF-IDF Data Ulasan untuk Klasifikasi SVM	68
Tabel 2.20	Perhitungan Jarak Kuadrat pada Vektor x_1 dan x_2	70
Tabel 2.21	<i>Confusion Matrix</i>	78
Tabel 2.22	Kriteria Akurasi	78
Tabel 2.23	Skala Interpretasi Nilai Presisi	79
Tabel 2.24	Skala Interpretasi Nilai <i>Recall</i>	80
Tabel 2.25	Skala Interpretasi Nilai <i>F1-score</i>	81

Tabel 4.1	Hasil <i>Labelling</i> pada Data Ulasan	88
Tabel 4.2	Contoh Proses <i>Case Folding</i>	89
Tabel 4.3	Contoh Proses <i>Cleaning</i>	90
Tabel 4.4	Contoh Proses <i>Tokenizing</i>	90
Tabel 4.5	Contoh Proses <i>Normalize</i>	91
Tabel 4.6	Contoh Proses <i>Stopword Removal</i>	91
Tabel 4.7	Contoh Proses <i>Stemming</i>	92
Tabel 4.8	Pembagian Data Latih dan Data Uji	93
Tabel 4.9	Contoh Data Latih untuk Pembobotan Kata TF-IDF	94
Tabel 4.10	Hasil Perhitungan Nilai TF pada Contoh Data Latih	95
Tabel 4.11	Hasil Perhitungan IDF pada Contoh Data Latih	96
Tabel 4.12	Hasil Perhitungan TF-IDF pada Contoh Data Latih	97
Tabel 4.13	Hasil Perhitungan TF-IDF Seluruh Data Latih	97
Tabel 4.14	Contoh Data Uji untuk Pembobotan Kata TF-IDF	98
Tabel 4.15	Hasil Perhitungan TF pada Contoh Data Uji	98
Tabel 4.16	Hasil Perhitungan IDF pada Contoh Data Uji	99
Tabel 4.17	Hasil Perhitungan TF-IDF pada Contoh Data Uji	100
Tabel 4.18	Hasil Perhitungan TF-IDF Seluruh Data Uji	100
Tabel 4.19	Contoh Data Latih untuk Klasifikasi dengan <i>Random Forest</i>	101
Tabel 4.20	Nilai TF-IDF Data Latih untuk Klasifikasi <i>Random Forest</i>	101
Tabel 4.21	Kandidat <i>Threshold</i> untuk Kata "bagus"	104
Tabel 4.22	Kandidat <i>Threshold</i> Seluruh Fitur Kata	105
Tabel 4.23	<i>Threshold</i> Terbaik pada Setiap Fitur Kata	106
Tabel 4.24	Contoh Data Uji untuk Prediksi dengan <i>Random Forest</i>	111
Tabel 4.25	Nilai TF-IDF Data Uji untuk Prediksi	111
Tabel 4.26	Contoh Data Latih untuk Klasifikasi dengan SVM	114
Tabel 4.27	Nilai TF-IDF Data Latih untuk Klasifikasi SVM	114
Tabel 4.28	Perhitungan Jarak Kuadrat	115
Tabel 4.29	Perhitungan Jarak Kuadrat Python	116
Tabel 4.30	Hasil Perhitungan α_i dengan Python	118
Tabel 4.31	Contoh Data Uji untuk Prediksi dengan SVM	122
Tabel 4.32	Nilai TF-IDF Data Uji untuk Prediksi dengan SVM	122
Tabel 4.33	Hasil <i>Confusion Matrix</i>	124

DAFTAR GAMBAR

Gambar 1.1	Jumlah Pengguna BNI <i>Mobile Banking</i>	2
Gambar 2.1	Contoh Ulasan Pengguna pada Google Play Store . . .	10
Gambar 2.2	<i>Install Library</i> Google-Play-Scraper	12
Gambar 2.3	<i>Import Packages</i>	12
Gambar 2.4	URL Aplikasi Wondr	12
Gambar 2.5	<i>Scraping Review</i>	13
Gambar 2.6	Membuat <i>Data Frame</i>	13
Gambar 2.7	Tahapan <i>Text Mining</i>	14
Gambar 2.8	Tahapan <i>Text Preprocessing</i>	16
Gambar 2.9	<i>Decision Tree</i>	25
Gambar 2.10	Contoh Hasil Pemisahan Pertama <i>Decision Tree</i>	37
Gambar 2.11	Contoh Hasil Pemisahan Lanjutan <i>Decision Tree</i>	38
Gambar 2.12	<i>Random Forest</i>	41
Gambar 2.13	Hasil Akhir <i>Decision Tree</i> DT_1	45
Gambar 2.14	Hasil Akhir <i>Decision Tree</i> DT_2	46
Gambar 2.15	Hasil Akhir <i>Decision Tree</i> DT_3	46
Gambar 2.16	Prediksi pada <i>Decision Tree</i> DT_1	48
Gambar 2.17	Prediksi pada <i>Decision Tree</i> DT_2	48
Gambar 2.18	Prediksi pada <i>Decision Tree</i> DT_3	49
Gambar 2.19	<i>Hyperplane</i> dan <i>Hard Margin</i> pada SVM	50
Gambar 2.20	<i>Soft Margin</i> pada SVM	53
Gambar 2.21	Metode <i>One Versus Rest</i> pada 3 Kelas	59
Gambar 2.22	Metode <i>One Versus One</i> pada 3 Kelas	60
Gambar 2.23	Klasifikasi Non-Linear	61
Gambar 2.24	<i>Hybrid</i> RFSVM	75
Gambar 3.1	Diagram Alir Penelitian	85
Gambar 4.1	Hasil <i>Data Scraping</i>	86
Gambar 4.2	Hasil <i>Labelling</i> Manual	88

Gambar 4.3	Contoh Hasil Pemisahan Pertama DT_1	107
Gambar 4.4	Contoh Hasil Pemisahan Lanjutan D_1	107
Gambar 4.5	Hasil Akhir DT_1 Python	108
Gambar 4.6	Hasil Akhir DT_2 Manual	109
Gambar 4.7	Hasil Akhir DT_2 Python	109
Gambar 4.8	Hasil Akhir DT_3 Manual	110
Gambar 4.9	Hasil Akhir DT_3 Python	110
Gambar 4.10	Prediksi Data Uji pada DT_1	112
Gambar 4.11	Prediksi Data Uji pada DT_2	112
Gambar 4.12	Prediksi Data Uji pada DT_3	112
Gambar 4.13	Hasil Klasifikasi dengan RFSVM	128
Gambar 4.14	Wordcloud Positif	128
Gambar 4.15	Wordcloud Netral	129
Gambar 4.16	Wordcloud Negatif	130
Gambar 4.17	Top 10 Gabungan Kata di Sentimen Negatif	130