



**ANALISIS SENTIMEN TERHADAP ULASAN PENGGUNA PADA *TOOLS*  
MANAJEMEN PROYEK MENGGUNAKAN *RANDOM FOREST* DAN  
*LOGISTIC REGRESSION***

**SKRIPSI**

Untuk memenuhi persyaratan  
melakukan penelitian dalam rangka penyusunan skripsi

Oleh

**NAJELA QUMAIRA**

**NIM 2011016120006**

**PROGRAM STUDI S-1 ILMU KOMPUTER  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMBUNG MANGKURAT  
BANJARBARU**

**JANUARI 2026**

# SKRIPSI

## ANALISIS SENTIMEN TERHADAP ULASAN PENGGUNA PADA TOOLS MANAJEMEN PROYEK MENGGUNAKAN *RANDOM FOREST* DAN *LOGISTIC REGRESSION*

Oleh:

**NAJELA QUMAIRA**

**NIM. 2011016120006**

Telah dipertahankan di depan Dosen Penguji pada tanggal 10 November 2025.

Susunan Dosen Penguji:

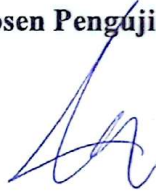
### **Pembimbing I**



Rudy Herteno, S.Kom., M.Kom.

NIP. 198809252022031003

### **Dosen Penguji I**



Muhammad Itqan Mazdadi, S.Kom., M.Kom.

NIP. 199006122019031013

### **Pembimbing II**



Fatma Indriani, S. T., M. I. T., Ph. D.

NIP. 198404202008122004

### **Dosen Penguji II**



Dwi Kartini, S.Kom., M.Kom.

NIP. 198704212012122003

Banjarbaru, 15 Januari 2026

**Koordinator Program Studi Ilmu Komputer**



Dwi Kartini, S.Kom., M.Kom.

NIP. 198704212012122003

## PERNYATAAN

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Banjarbaru, 15 Januari 2026

Yang Menyatakan,



Najela Qumaira

---

2011016120006

## ABSTRAK

### **ANALISIS SENTIMEN TERHADAP ULASAN PENGGUNA PADA *TOOLS* MANAJEMEN PROYEK MENGGUNAKAN *RANDOM FOREST* DAN *LOGISTIC REGRESSION***

(Oleh: Najela Qumaira; Pembimbing: Rudy Herteno S.Kom M.Kom; Fatma Indriani S.T M.I.T Ph.D; 61 halaman)

Manajemen proyek telah mengalami perubahan signifikan dengan hadirnya teknologi digital. Penerapan teknologi dalam manajemen proyek dapat menjadi solusi yang efektif. Berbagai alat yang tersedia di Playstore menawarkan fitur-fitur yang mendukung pengelolaan proyek dengan *rating* yang tinggi. Namun, angka tersebut belum memberikan gambaran menyeluruh terhadap sehingga perlu dilakukan analisis sentimen. Penelitian ini bertujuan untuk memahami persepsi pengguna terhadap alat-alat tersebut dan mengetahui kinerja dari metode klasifikasi yang digunakan. Metode yang digunakan dalam penelitian ini meliputi proses *Labeling*, *Preprocessing*, TF-IDF, *Cross Validation*, serta klasifikasi dengan algoritma Regresi Logistik dan *Random forest*. Hasil penelitian menunjukkan bahwa algoritma *Random forest* memberikan performa terbaik dengan akurasi 89,00%, diikuti oleh Regresi Logistik 70,80%. Dari seluruh algoritma yang diuji pada semua dataset, *Random forest* secara konsisten memberikan hasil yang paling unggul. Tingginya tingkat akurasi menjadikan *Random forest* sebagai pilihan terbaik untuk analisis sentimen pada alat manajemen proyek. Penelitian ini memiliki implikasi yang signifikan dalam meningkatkan produktivitas dan dapat diterapkan di berbagai industri serta organisasi.

**Kata Kunci :** *Machine Learning, Random Forest, Sentiment Analysis, Google Play Store*

## **ABSTRACT**

### ***SENTIMENT ANALYSIS OF PROJECT MANAGEMENT TOOLS REVIEWS USING RANDOM FOREST AND LOGISTIC REGRESSION***

*(By: Najela Qumaira; Supervisor: Rudy Herteno S.Kom M.Kom; Fatma Indriani S.T M.I.T Ph.D; 61 Pages)*

*Project management has changed significantly due to the use of digital technology.. While numerous high rated project management tools are available on the Play Store, user ratings alone do not provide a comprehensive picture of user experience. Therefore, sentiment analysis is needed. This study aims to evaluate user perceptions of these tools and compare the performance of different classification methods. The methodology involves data labeling, preprocessing, TF-IDF vectorization, and cross-validation, with classification performed using Logistic Regression and Random Forest algorithms. The results indicate that Random Forest outperformed Logistic Regression, achieving an accuracy of 89.00% compared to 70.80%. Given its consistent high performance across all datasets, Random Forest is the most effective choice for sentiment analysis in this project. These findings offer significant implications for enhancing productivity and are applicable across various industries and organizations.*

**Keywords :** *Machine Learning, Random forest, Sentiment Analysis, Google Play Store*

## PRAKATA

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya sehingga saya dapat menyelesaikan skripsi yang berjudul *"Analisis Sentimen Terhadap Ulasan Pengguna pada Tools Manajemen Proyek Menggunakan Random Forest dan Logistic Regression"*. Skripsi ini disusun sebagai salah satu syarat untuk menyelesaikan program studi pada jurusan Ilmu Komputer, Universitas Lambung Mangkurat.

Selama proses penyusunan skripsi ini, saya menyadari bahwa tidak sedikit bantuan, bimbingan, serta dukungan dari berbagai pihak yang sangat berharga. Oleh karena itu, pada kesempatan ini, saya ingin mengucapkan terima kasih kepada:

1. Allah SWT, karena atas karunia-Nya lah penulis dapat menyelesaikan skripsi dengan tepat waktu.
2. Bapak Rudy Herteno selaku dosen pembimbing utama, yang telah memberikan arahan, bimbingan, dan waktu berharga selama proses penyusunan skripsi ini.
3. Ibu Fatma Indriani selaku dosen pembimbing pendamping, yang telah memberikan arahan, bimbingan, dan waktu berharga selama proses penyusunan skripsi ini.
4. Ibu Dwi Kartini selaku Ketua Program Studi Ilmu Komputer FMIPA ULM yang telah memberikan bantuan dan izin beliau hingga skripsi ini dapat diselesaikan.
5. Seluruh Dosen dan Staff Program Studi Ilmu Komputer FMIPA ULM atas ilmu dan bantuan yang diberikan selama ini.
6. Keluarga yang selalu memberikan doa, dukungan moral, serta semangat.
7. Noor Zalekha Alhabesyah dan Rahmina Ulfah Aflaha yang selalu memberikan semangat dalam segala situasi hingga akhir perjalanan kuliah dan membantu dalam memberikan solusi dalam pengerjaan skripsi.
8. Azizah, Nurhaliza, Meida, Rara, Nurul, Dinda dan Risa sebagai teman yang selalu memberikan informasi, hiburan dan dukungan hingga akhir.
9. Rafayel sebagai ikan kesayangan yang selalu menghibur disaat saya lelah dalam menulis naskah dan akan selalu menghibur saya walau dari jauh.

10. Teman-teman dan rekan mahasiswa di Universitas Lambung Mangkurat yang telah memberikan bantuan serta dukungan selama masa studi dan penyusunan skripsi ini.

Saya menyadari bahwa skripsi ini masih memiliki banyak kekurangan. Oleh karena itu, kritik dan saran yang membangun sangat saya harapkan demi kesempurnaan karya ini di masa mendatang.

Semoga skripsi ini dapat memberikan manfaat bagi pembaca dan perkembangan ilmu pengetahuan, khususnya dalam bidang ilmu komputer.

Banjarbaru, 15 Januari 2026



Najela Qumaira

## DAFTAR ISI

	<b>Halaman</b>
HALAMAN JUDUL HALAMAN JUDUL .....	i
HALAMAN PENGESAHAN.....	ii
PERNYATAAN.....	iii
ABSTRAK .....	iv
<i>ABSTRACT</i> .....	v
PRAKATA.....	vi
DAFTAR ISI.....	viii
Halaman .....	viii
DAFTAR TABEL.....	x
DAFTAR GAMBAR .....	xi
DAFTAR LAMPIRAN .....	xiii
BAB I PENDAHULUAN.....	1
1.1    Latar Belakang .....	1
1.2    Rumusan Masalah .....	2
1.3    Tujuan Penelitian .....	3
1.4    Manfaat Penelitian .....	3
1.5    Batasan Masalah.....	3
BAB II TINJAUAN PUSTAKA.....	5
2.1    Kajian Terdahulu.....	5
2.2    Landasan teori .....	10
2.2.1    Manajemen Proyek .....	10
2.2.2    Google Play Store .....	10
2.2.3    Analisis Sentimen .....	10

2.2.4	TF-IDF .....	11
2.2.5	<i>Cross Validation</i> .....	12
2.2.6	<i>Machine Learning</i> .....	12
2.2.7	Random Forest.....	14
2.2.8	Regresi Logistik.....	16
2.2.9	<i>Confusion Matrix</i> .....	17
BAB III METODE PENELITIAN.....		20
3.1	Bahan Penelitian.....	20
3.2	Alat Penelitian.....	20
3.3	Prosedur Penelitian.....	21
BAB IV HASIL DAN PEMBAHASAN .....		24
4.1	Dataset.....	24
4.1.1	Crawling data.....	24
4.1.2	Labeling .....	26
4.1.3	<i>Preprocessing</i> .....	29
4.1.4	TF-IDF ( <i>Term Frequency- Inverse Document Frequency</i> ) .....	38
4.1.5	<i>K- Cross Validation</i> .....	43
4.1.6	Klasifikasi <i>Random Forest</i> .....	44
4.1.7	Klasifikasi Regresi Logistik.....	45
4.1.8	Evaluasi.....	46
4.1.8.1	Evaluasi Algoritma <i>Random Forest</i> .....	46
4.1.8.2	Evaluasi Algoritma Regresi Logistik.....	49
4.2	Pembahasan.....	53
BAB V KESIMPULAN.....		60
5.1	Kesimpulan .....	60
5.2	Saran.....	60
DAFTAR PUSTAKA .....		61
LAMPIRAN.....		64

## DAFTAR TABEL

### Tabel

Tabel 1. Keaslian Penelitian.....	7
Tabel 2. Perencanaan Penelitian .....	9
Tabel 3. <i>Confusion Matrix</i> .....	18
Tabel 4. Rumus Evaluasi Performa Metode .....	18
Tabel 5. Data penelitian .....	20
Tabel 6. Contoh Ulasan untuk Sentimen Positif dan Negatif .....	29
Tabel 7. Hasil <i>Cleaning Data</i> .....	29
Tabel 8. Hasil <i>Transform Case</i> .....	30
Tabel 9. Hasil <i>Tokenize</i> dan <i>Filter Token by Length</i> .....	32
Tabel 10. Hasil <i>Stopwords</i> dan <i>Stemming</i> .....	33
Table 11. Contoh Dokumen yang Digunakan Untuk Perhitungan TF-IDF .....	38
Tabel 12. Contoh Perhitungan <i>Term Frequency</i> (TF).....	39
Tabel 13. Contoh Perhitungan <i>Invers Document Frequency</i> (IDF).....	40
Tabel 14. Contoh Perhitungan <i>Term Frequency- Invers Document Frequency</i> (TF-IDF ).....	42
Tabel 15. Hasil <i>Confusion Matrix</i> Algoritma <i>Random Forest</i> .....	45
Tabel 16. Hasil <i>Confusion Matrix</i> Algoritma Regresi Logistik.....	46
Tabel 17. Hasil Akurasi Algoritma <i>Random Forest</i> .....	47
Tabel 18. Hasil Akurasi Algoritma Regresi Logistik.....	50

## DAFTAR GAMBAR

### Gambar

Gambar 1. <i>Flowchart Random Forest</i> .....	16
Gambar 2. Contoh Alur Pembagian Data dalam <i>Cross- Validation (CV)</i>	12
Gambar 3. <i>Flowchart</i> Penelitian .....	21
Gambar 4. Hasil <i>Crawling Data</i> pada setiap Aplikasi .....	25
Gambar 5. Hasil <i>Crawling Data</i> yang Sudah <i>Diremove</i> Duplikat dan Diseimbangkan.....	25
Gambar 6. Jumlah Label Positif pada Data <i>Raw</i> .....	27
Gambar 7. Jumlah Label Negatif pada Data <i>Raw</i> .....	27
Gambar 8. Jumlah Label Positif pada Data yang Sudah <i>Diremove</i> Duplikat dan Diseimbangkan .....	28
Gambar 9. Jumlah Label Negatif pada Data yang Sudah <i>Diremove</i> Duplikat dan Diseimbangkan .....	28
Gambar 10. <i>Preprocessing</i> Tahap Pertama .....	30
Gambar 11. Parameter <i>Tokenize Regular Expression</i> .....	31
Gambar 12. Parameter <i>Tokenize Non-Letters</i> .....	31
Gambar 13. Total Atribut Sebelum Dilakukan <i>Preprocessing</i> .....	32
Gambar 14. Total Atribut Setelah Dilakukan <i>Preprocessing</i> .....	34
Gambar 15. <i>Preprocessing</i> Tahap ke Dua .....	34
Gambar 16. <i>Wordcloud</i> Kelas Positif pada Dataset Google Task .....	35
Gambar 17. <i>Wordcloud</i> Kelas Negatif pada Dataset Google Task.....	35
Gambar 18. <i>Wordcloud</i> Kelas Positif pada Dataset Microsoft Team .....	36
Gambar 19. <i>Wordcloud</i> Kelas Negatif pada Dataset Microsoft Team.....	36
Gambar 20. <i>Wordcloud</i> Kelas Positif Pada Dataset Microsoft To Do.....	37
Gambar 21. <i>Wordcloud</i> Kelas Negatif Pada Dataset Microsoft To Do ...	37
Gambar 22. Proses <i>10K- Fold Cross Validation</i> .....	44
Gambar 23. Alur Penelitian pada Rapidminer .....	46
Gambar 24. Hasil Akurasi, Presisi dan <i>Recall</i> pada Data Google Task ..	47
Gambar 25. Hasil Akurasi, Presisi dan <i>Recall</i> pada Data Microsoft Team .....	48

Gambar 26. Hasil Akurasi, Presisi dan <i>Recall</i> pada Data Microsoft Team .....	49
Gambar 27. Hasil Akurasi, Presisi dan <i>Recall</i> pada Data Google Task ..	51
Gambar 28. Hasil Akurasi, Presisi dan <i>Recall</i> pada Data Microsoft Team .....	52
Gambar 29. Hasil Akurasi, Presisi dan <i>Recall</i> pada Data Microsoft Team .....	53
Gambar 30. Frekuensi Kemunculan Kata Terbanyak pada Data Google Task .....	54
Gambar 31. Frekuensi Kemunculan Kata Terbanyak pada Data Microsoft Team.....	55
Gambar 32. Frekuensi Kemunculan Kata Terbanyak pada Data Microsoft To Do.....	55
Gambar 33. Perbandingan Hasil Akurasi Algoritma pada Setiap Dataset .....	.57