

SKRIPSI

**ANALISIS SENTIMEN ULASAN APLIKASI DANA DI *PLAY*
STORE MENGGUNAKAN METODE *XGBOOST***



Oleh:

Siti Komariyah

42418026

**PROGRAM STUDI INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS PERADABAN**

2025

SKRIPSI

**ANALISIS SENTIMEN ULASAN APLIKASI DANA DI *PLAY*
STORE MENGGUNAKAN METODE *XGBOOST***



**Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana
Komputer (S.Kom)**

Oleh:

Siti Komariyah

42418026

**PROGRAM STUDI INFORMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS PERADABAN**

2025

PERNYATAAN ORISINALITAS

JUDUL : ANALISIS SENTIMEN ULASAN APLIKASI DANA DI *PLAY*
STORE MENGGUNAKAN METODE *XGBOOST*

NAMA : SITI KOMARIYAH

NIM : 42418026

“Saya menyatakan dan bertanggung jawab dengan sebenar-benarnya bahwa skripsi saya adalah hasil karya saya sendiri kecuali cuplikan dan ringkasan yang masing-masing telah saya jelaskan sumbernya. Jika pada waktu selanjutnya pihak lain ada yang mengklaim bahwa skripsi ini adalah karyanya yang disertai bukti-bukti yang cukup, maka saya bersedia dibatalkan gelar Sarjana Komputer saya beserta segala hak dan kewajiban yang melekat pada gelar tersebut”.

Bumiayu, 2 Agustus 2025



Siti Komariyah
NIM. 42418026

PERSETUJUAN SKRIPSI

JUDUL : ANALISIS SENTIMEN ULASAN APLIKASI
DANA DI *PLAY STORE* MENGGUNAKAN
METODE *XGBOOST*
NAMA : SITI KOMARIYAH
NIM : 42418026

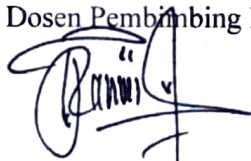
Skripsi ini telah diperiksa dan disetujui oleh dosen pembimbing tugas akhir guna
mencapai gelar Sarjana Komputer Strata Satu pada Fakultas Sains dan Teknologi,

Universitas Peradaban

Bumiayu, 02 Agustus 2025

Menyetujui

Dosen Pembimbing I



Tezhar Rayendra TPN, M.Kom
NIDN. 0619019201

Dosen Pembimbing II



Asep Saeful Millah, M.Kom
NIDN. 0613068803

Mengetahui

Ketua Program Studi Informatika



Nurul Mega Saraswati, M.Kom
NIDN. 0606069102

PENGESAHAN SKRIPSI

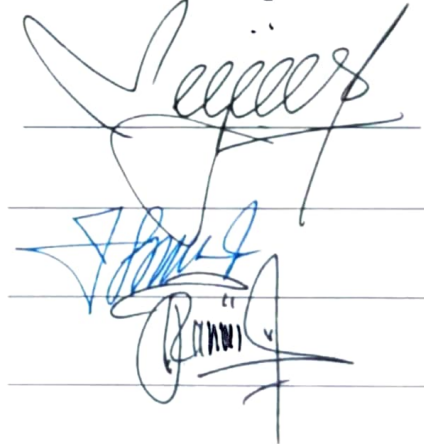
JUDUL : ANALISIS SENTIMEN ULASAN APLIKASI
DANA DI *PLAY STORE* MENGGUNAKAN
METODE *XGBOOST*
NAMA : SITI KOMARIYAH
NIM : 42418026

Skripsi ini telah diseminarkan dihadapan Dewan Penguji Skripsi, pada tanggal 04
Agustus 2025. Menurut pandangan kami, skripsi ini memadai dari segi kualitas
untuk memperoleh gelar Sarjana Komputer (S.Kom)
Bumiayu, 16 Agustus 2025

Nama Penguji

1. Nurul Mega Saraswati, M.Kom
2. Sorikhi, M.Kom
3. Asep Saeful Millah, M.Kom
4. Tezhar Rayendra TPN, M.Kom

Tanda Tangan



Menyetujui

Dekan Fakultas Sains dan Teknologi
Universitas Peradaban



Rizki Noor Prasetyono, M.Pd
NIDN. 0611099101

Ketua Program Studi
Informatika



Nurul Mega Saraswati, M.Kom
NIDN.0606069102

MOTTO DAN DEDIKASI

Motto :

- *Lakukanlah sesuatu yang membuatmu bahagia walaupun itu hal yang sederhana*
- *Allah SWT tidak menuntut dan menyuruhmu untuk sempurna dalam menjalani hidup ini, cukup kamu berusaha dan percaya pada segala ketetapan Nya itu yang terbaik dan terindah untukmu*
- *Ketakutan yang ada dipikiranmu belum tentu itu yang akan terjadi, bisa jadi hal – hal indah yang akan terjadi padamu, jadi stop berpikir buruk*
- *Berawal dari pikiran menjadi tindakan, menjadi kebiasaan, menjadi pola hidup dan menjadi kesuksesan. Maka tentukanlah pikiranmu dari sekarang*

Dedikasi :

Skripsi ini saya dedikasikan untuk setiap orang yang mencintai saya dengan sepenuh hati, dan untuk mereka yang saya kasihi dengan sepenuh jiwa. yaitu; bapak dan mama, adikku, kaki dan biyung, bu lik dan pak lik, saudara, teman, dan orang yang memperjuangkan saya untuk menjadi pasangannya. Alhamdulillah, atas doa dan dukungan yang tak henti mengalir skripsi ini akhirnya telah sampai di garis akhir.

ABSTRACT

This research implements the XGBoost algorithm for the DANA application sentiment analysis system. This research implements the XGBoost algorithm for the sentiment analysis system of the DANA application. This research aims to determine the accuracy value of the Extreme Gradient Boosting (XGBoost) method in classifying the sentiment (positive and negative) of DANA application users. Methods include dataset retrieval, preprocessing, training of test data and training data, implementation of the XGBoost algorithm, visualisation of positive and negative words, and model evaluation using the ROC (Receiver Operating Characteristic) curve. The data sources consist of Indonesian language reviews obtained using a web scraping approach. Data preprocessing stages such as tokenisation, stopwords, normalisation, filtering and casefolding. Next, text features are extracted using TF-IDF and classified into positive or negative sentiment with the XGBoost algorithm. The results show that the XGBoost model achieved 78% accuracy in classifying the sentiment of user reviews. In conclusion, the XGBoost algorithm can be used for sentiment analysis to produce effective results. This research is expected to help developers understand user needs and complaints through automated review analysis.

Keywords: *Sentiment Analysis, User Reviews, DANA App, XGBoost.*

ABSTRAK

Penelitian ini mengimplementasikan algoritma *XGBoost* untuk sistem analisis sentimen aplikasi DANA. Penelitian ini bertujuan untuk mengetahui nilai akurasi metode *Extreme Gradient Boosting (XGBoost)* dalam mengklasifikasikan sentimen (positif dan negatif) pengguna aplikasi DANA. Metode meliputi pengambilan dataset, proses *preprocessing*, pelatihan data uji dan data latih, implementasi algoritma *XGBoost*, visualisasi kata positif dan negatif, dan evaluasi model menggunakan kurva *ROC (Receiver Operating Characteristic)*. Sumber data terdiri dari ulasan berbahasa Indonesia yang diperoleh menggunakan pendekatan *web scraping*. Tahapan proses *preprocessing* data seperti *tokenisasi*, *stopword*, *normalisasi*, *filtering* dan *casefolding*. Selanjutnya, Fitur teks diekstraksi menggunakan *TF-IDF* dan diklasifikasikan ke dalam sentimen positif atau negatif dengan algoritma *XGBoost*. Hasil menunjukkan bahwa model *XGBoost* berhasil mencapai akurasi sebesar 78% dalam mengklasifikasikan sentimen ulasan pengguna. Kesimpulannya, algoritma *XGBoost* dapat digunakan untuk melakukan analisis sentimen dan mendapatkan hasil yang baik dan cukup efektif. Penelitian ini diharapkan membantu pengembang memahami kebutuhan dan keluhan pengguna lewat analisis ulasan otomatis.

Kata Kunci: Analisis Sentimen, Ulasan Pengguna, Aplikasi DANA, *XGBoost*.

KATA PENGANTAR



Dengan memanjatkan puji syukur kehadiran Tuhan Yang Maha Esa yang telah memberikan nikmat dan hidayah-Nya, sehingga penyusunan Skripsi ini dapat terselesaikan dengan baik. Penyusunan penelitian ini dilakukan untuk mencapai salah satu syarat gelar sarjana program studi Informatika pada Fakultas Sains dan Teknologi Universitas Peradaban. Proses penyelesaian penelitian ini tidak lepas dari berbagai bantuan moral dan material, dukungan saran hingga kritik. Oleh karena itu, peneliti ingin mengucapkan rasa terimakasih sebesar-besarnya kepada:

1. Dr. Muh Kadarisman, S.H., M.Si selaku Rektor Universitas Peradaban Bumiayu.
2. Bapak Rizki Noor Prasetyono, M.Pd, selaku Dekan Fakultas Sains dan Teknologi Universitas Peradaban Bumiayu.
3. Ibu Nurul Mega Saraswati, M.Kom, selaku Ketua Program Studi Informatika Universitas Peradaban Bumiayu.
4. Bapak Tezhar Rayendra TPN, M.Kom, selaku Dosen Pembimbing I yang telah membimbing serta memotivasi penulis dalam menyelesaikan penyusunan skripsi ini.
5. Bapak Asep Saeful Millah, M.Kom, selaku Dosen Pembimbing II yang telah membimbing serta memotivasi penulis dalam menyelesaikan skripsi ini.

6. Kedua orang tua saya, Bapak Salimin dan Ibu Mudarojah yang selalu memberikan do'a dan dorongan selama peneliti menyusun skripsi.
7. Teman-teman seperjuangan yang selalu memberikan semangat dan do'a.
8. Saudara, kerabat serta seorang yang telah memberikan do'a tulusnya, menasehati penulis agar tetap semangat sehingga skripsi ini selesai dengan baik.
9. Peneliti yang selalu berusaha dan berjuang hingga akhir

Bumiayu, 16 Agustus 2025

A handwritten signature in black ink, appearing to be 'Siti Komariyah', written in a cursive style with a large loop at the end.

Siti Komariyah
NIM. 42418026

DAFTAR ISI

PERNYATAAN ORISINALITAS	iii
PERSETUJUAN SKRIPSI	iv
PENGESAHAN SKRIPSI	v
MOTTO DAN DEDIKASI	vi
ABSTRACT	vii
ABSTRAK	viii
KATA PENGANTAR	ix
DAFTAR ISI	xi
DAFTAR GAMBAR	xiv
DAFTAR TABEL	xvi
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	5
1.6 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	7
2.1 Penelitian Terkait.....	7
2.2 Landasan Teori	13
2.2.1 Analisis Sentimen.....	13

2.2.2	Aplikasi DANA.....	14
2.2.3	<i>Google Play Store</i>	15
2.2.4	<i>XGBoost (Extreme Gradient Boosting)</i>	17
2.2.5	<i>TF-IDF (Term Frequency – Inverse Document Frequency)</i>	21
2.2.6	Klasifikasi	21
2.2.7	<i>Text Mining</i>	22
2.2.8	<i>Kurva ROC (Receiver Operating Characteristic)</i>	22
2.2.9	<i>Streamlit</i>	22
2.2.10	<i>Python</i>	23
2.2.11	<i>Visual Studio Code</i>	24
2.2.12	<i>Google Chrome</i>	25
2.2.13	<i>UML (Unified Modeling Language)</i>	26
2.2.14	<i>Draw.io</i>	29
2.3	Kerangka Berpikir	29
BAB III METODE PENELITIAN		31
3.1	Tahapan Penelitian	31
3.1.1	Studi Literatur	32
3.1.2	Pengumpulan Data	32
3.1.3	Pengolahan Data.....	33
3.1.4	Perancangan Sistem	33
3.1.5	Evaluasi	35
3.1.6	Hasil	35
3.2	Perangkat Penelitian	36

3.3	Metode Yang Diusulkan.....	37
BAB IV HASIL DAN PEMBAHASAN.....		39
4.1	Hasil Pengambilan Data	39
4.2	Hasil <i>Preprocessing</i>	42
4.3	Implementasi	43
4.3.1	<i>Upload File</i> Dataset	43
4.3.2	Proses <i>Preprocessing</i>	44
4.3.3	Pelatihan Data Uji dan Data Latih	47
4.3.4	Visualisasi	48
4.3.5	Evaluasi Model.....	50
BAB V KESIMPULAN.....		51
5.1	Kesimpulan.....	51
5.2	Saran.....	51
DAFTAR PUSTAKA		53
LAMPIRAN.....		58

DAFTAR GAMBAR

Gambar 2. 1 Logo Aplikasi DANA	15
Gambar 2. 2 Grafik Penjualan <i>Smartphone</i> Secara Global.....	16
Gambar 2. 3 Logo <i>Google Play Store</i>	17
Gambar 2. 4 Logo <i>Visual Studio Code</i>	24
Gambar 2. 5 Logo <i>Google Chrome</i>	25
Gambar 3. 1 Tahapan Penelitian	31
Gambar 3. 2 <i>Usecase Diagram</i> Sistem Klasifikasi	34
Gambar 3. 3 Metode yang Diusungkan.....	37
Gambar 4. 1 Cara pengambilan data menggunakan <i>google colaboratory</i>	39
Gamba 4. 2 Lanjutan cara pengambilan data	40
Gambar 4. 3 Lanjutan cara pengambilan data.....	40
Gambar 4. 4 Tampilan Awal Sistem	43
Gambar 4. 5 <i>Input Dataset</i>	43
Gambar 4. 6 Tampilan Dataset.....	44
Gambar 4. 7 Tampilan <i>Case Folding</i>	44
Gambar 4. 8 Tampilan <i>Filtering</i>	45
Gambar 4. 9 Tampilan Normalisasi	45
Gambar 4. 10 Tampilan <i>Stopword</i>	46
Gambar 4. 11 Tampilan <i>Tokenizing</i>	46
Gambar 4. 12 Tampilan Pelatihan.....	47
Gambar 4. 13 Lanjutan Tampilan Pelatihan	47

Gambar 4. 14 Tampilan Visualisasi	48
Gambar 4. 15 Tampilan <i>Positive Wordcloud</i>	49
Gambar 4. 16 Tampilan <i>Negative Wordcloud</i>	49
Gambar 4. 17 Tampilan Kurva <i>ROC</i>	50

DAFTAR TABEL

Tabel 2. 1 Penelitian Terkait	10
Tabel 2. 2 Penelitian Terkait (Lanjutan)	11
Tabel 2. 3 Penelitian Terkait (Lanjutan)	12
Tabel 2. 4 <i>Usecase Diagram</i>	27
Tabel 2. 5 <i>Activity Diagram</i>	28
Tabel 3. 1 Perangkat Penelitian.....	36
Tabel 4. 1 Ulasan Hasil <i>Scraping</i>	41
Tabel 4. 2 Hasil <i>Preprocessing</i>	42