

**ANALISIS SENTIMEN PINJAMAN ONLINE
DI INDONESIA MENGGUNAKAN METODE
RANDOM FOREST
(STUDI KASUS PADA APLIKASI *TWITTER*)**



**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS MERDEKA PASURUAN**

2024

**ANALISIS SENTIMEN PINJAMAN ONLINE
DI INDONESIA MENGGUNAKAN METODE *RANDOM
FOREST*
(STUDI KASUS PADA APLIKASI *TWITTER*)**

SKRIPSI

OLEH:

MUHAMMAD ALIF ARDIANSYAH

2055201001024



PROGRAM STUDI TEKNIK INFORMATIKA

FAKULTAS TEKNOLOGI INFORMASI

UNIVERSITAS MERDEKA PASURUAN

2024

**ANALISIS SENTIMEN PINJAMAN ONLINE
DI INDONESIA MENGGUNAKAN METODE *RANDOM
FOREST*
(STUDI KASUS PADA APLIKASI *TWITTER*)**

SKRIPSI

Diajukan sebagai salah satu persyaratan dalam menyelesaikan program

Sarjana Komputer (S.Kom)

Oleh:

Muhammad Alif Ardiansyah

2055201001024

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS MERDEKA PASURUAN**

2024

PERSETUJUAN PEMBIMBING SKRIPSI

ANALISIS SENTIMEN PINJAMAN ONLINE DI INDONESIA
MENGUNAKAN METODE *RANDOM FOREST*
(STUDI KASUS PADA APLIKASI TWITTER)

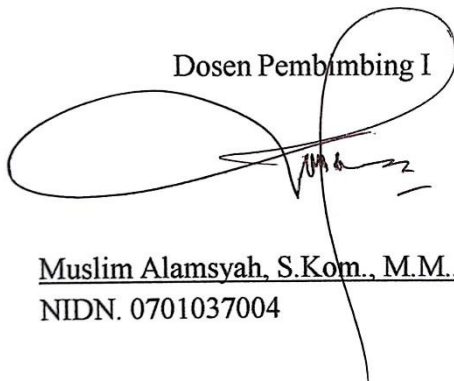
Nama : Muhammad Alif Ardiansyah
NPM : 2055201001024
Program Studi : Informatika



Pasuruan, 30 Juli 2024

Telah diperiksa dan disetujui oleh

Dosen Pembimbing I



Muslim Alamsyah, S.Kom., M.M., M.Kom.
NIDN. 0701037004

Dosen Pembimbing II



Mochammad Firman Arif, S.Kom., M.Kom.
NIDN. 0704029201

PERSETUJUAN PEMBIMBING SKRIPSI

Skripsi oleh Muhammad Alif A. ini telah dipertahankan didepan dewan penguji pada tanggal 30 Juli 2024

Ketua Penguji



(Dian Ahkam Sani, S.Kom., M.Kom.)

NIDN. 0724079202

Penguji I



(Rudi Hariyanto, S.Kom., M.T.)

NIDN. 0718018401

Penguji II



(Anang Aris Widodo, S.Kom., M.T.)

NIDN. 0702038102

Mengesahkan,

Mengetahui,

Dekan Fakultas Teknologi Informasi

Ketua Program Studi



Muhammad Misdrum, S.Kom., M.Kom.

NIDN. 0717046704



Dian Ahkam Sani, S.Kom., M.Kom.

NIDN. 0724079202

PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini :

Nama : Muhammad Alif A.
NPM : 2055201001024
Jurusan / Program Studi : Teknik Informatika
Fakultas : Teknologi Informasi

Menyatakan dengan sesungguhnya bahwa skripsi yang saya tulis ini benar- benar tulisan saya, dan bukan merupakan plagiasi baik Sebagian atau seluruhnya. Apabila di kemudian hari terbukti atau dapat dibuktikan bahwa skripsi ini hasil plagiasi, baik Sebagian atau seluruhnya, maka saya bersedia menerima sanksi atas perbuatan tersebut sesuai dengan ketentuan yang berlaku.

Pasuruan, 30 Juli 2024
Yang membuat pernyataan



MUHAMMAD ALIF A.
2055201001024

MOTTO

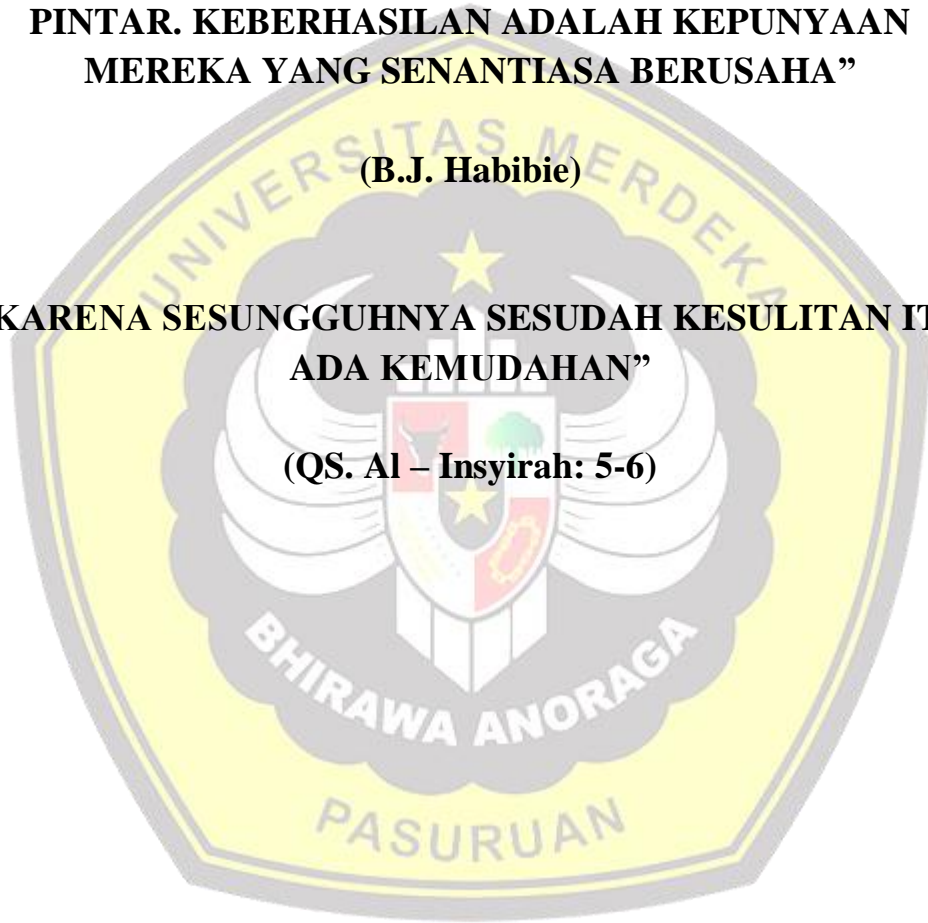
**“DISIPLIN ITU TIDAK MENGEKANG, SEBALIKNYA
JUSTRU MEMBEBAHKAN”**

**“KEBERHASILAN BUKANLAH MILIK ORANG YANG
PINTAR. KEBERHASILAN ADALAH KEPUNYAAN
MEREKA YANG SENANTIASA BERUSAHA”**

(B.J. Habibie)

**“KARENA SESUNGGUHNYA SESUDAH KESULITAN ITU
ADA KEMUDAHAN”**

(QS. Al – Insyirah: 5-6)



PERSEMBAHAN

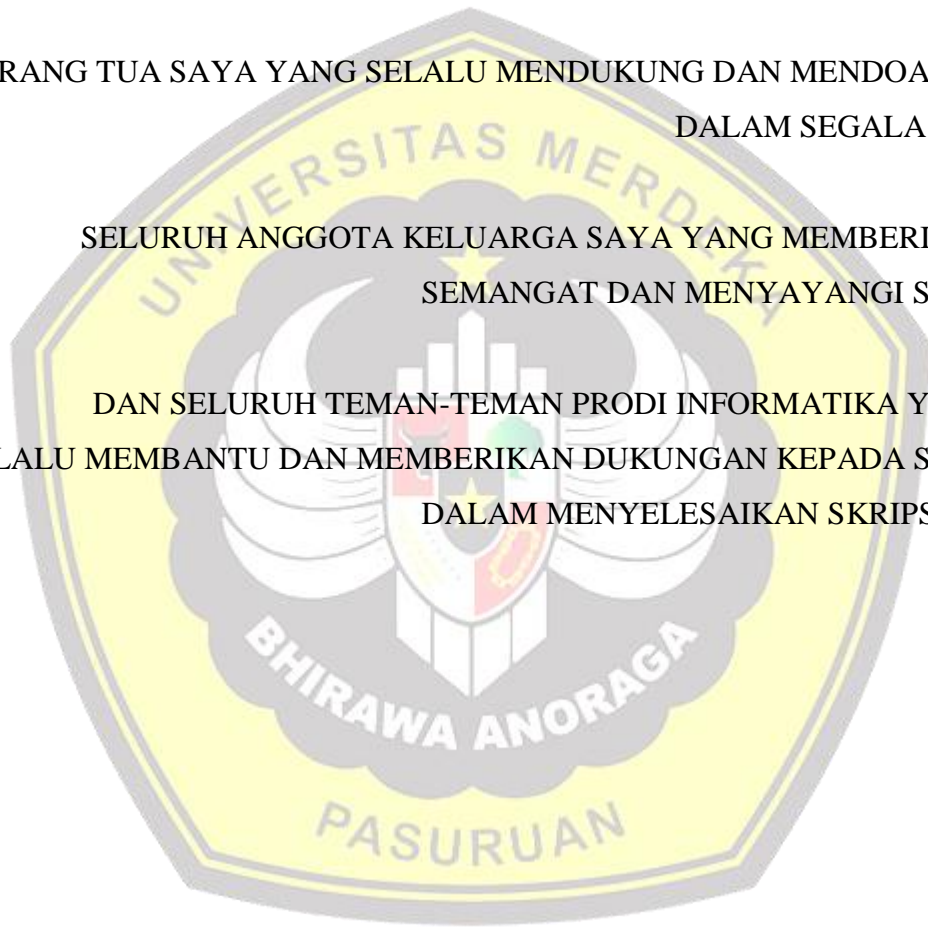
SKRIPSI INI SAYA PERSEMBAHKAN KEPADA :

ALLAH SWT YANG SELALU SAYA HARAPKAN RIDHO DAN
HIDAYAHNYA

ORANG TUA SAYA YANG SELALU MENDUKUNG DAN MENDOAKAN
DALAM SEGALA HAL

SELURUH ANGGOTA KELUARGA SAYA YANG MEMBERIKAN
SEMANGAT DAN MENYAYANGI SAYA

DAN SELURUH TEMAN-TEMAN PRODI INFORMATIKA YANG
SELALU MEMBANTU DAN MEMBERIKAN DUKUNGAN KEPADA SAYA
DALAM MENYELESAIKAN SKRIPSI INI



ABSTRAK

Muhammad Alif A, 2024. Analisis Sentimen Pinjaman Online di Indonesia Menggunakan Metode *Random Forest*: Studi Kasus Pada Aplikasi Twitter

Pembimbing : (I) Muslim Alamsyah, S.Kom., M.M., M.Kom. (II) Mochammad Firman Arif, S.Kom., M.Kom

Kata kunci : analisis sentimen, pinjaman online, *Random Forest*, *SMOTE*, Twitter

Di tengah kemajuan pesat era digital, sektor pinjaman online di Indonesia mengalami pertumbuhan signifikan. Namun, regulasi yang belum memadai dan meningkatnya tindakan tidak etis oleh penyedia layanan telah menimbulkan beragam pandangan di kalangan masyarakat, terutama pada platform media sosial seperti Twitter. Tujuan dari penelitian ini adalah menganalisis persepsi publik tentang pinjaman online dengan menerapkan metode *Random Forest* pada teks *tweet* di Twitter. Metode ini dipilih karena kemampuannya dalam mengatasi *overfitting* dan memberikan hasil klasifikasi yang akurat. Data penelitian terdiri dari 1000 *tweet* yang dikategorikan sebagai sentimen positif, negatif, atau netral. Sebelum klasifikasi, dilakukan *preprocessing* data termasuk pembersihan teks, tokenisasi, dan stemming serta dilakukan perhitungan TF-IDF. Hasil penelitian menunjukkan bahwa kombinasi metode *Random Forest* dan teknik *SMOTE* memberikan kontribusi signifikan dalam analisis sentimen. Model yang dikembangkan mencapai akurasi 80%, dengan *precision* untuk sentimen negatif 83%, netral 65%, dan positif 83%. Nilai *recall* untuk sentimen negatif adalah 93%, netral 55%, dan positif 45%, sedangkan *F1-score* masing-masing adalah 88%, 59%, dan 59%. Kesimpulan dapat disimpulkan bahwa meskipun model ini menunjukkan kinerja yang optimal, terutama untuk sentimen negatif, peningkatan lebih lanjut diperlukan untuk analisis yang lebih komprehensif, terutama pada sentimen netral.

ABSTRACT

Muhammad Alif A, 2024. *Analysis of Online Loan Sentiment in Indonesia Using the Random Forest Method: Case Study on the Twitter Application*

Pembimbing : (I) Muslim Alamsyah, S.Kom., M.M., M.Kom. (II) Mochammad Firman Arif, S.Kom., M.Kom

Kata kunci : *online lending, Random Forest, sentiment analysis, SMOTE, twitter*

The rapid growth of digital technology has led to a significant surge in online lending in Indonesia. However, inadequate regulations and unethical practices by service providers have generated diverse public opinions, particularly on social media platforms such as Twitter. This study aims to analyze public sentiment towards online lending by applying the Random Forest algorithm to Twitter data. Random Forest was chosen for its ability to handle overfitting and provide accurate classification results. The dataset consisted of 1,000 tweets categorized into positive, negative, and neutral sentiments. Before classification, data preprocessing was conducted, including text cleaning, tokenization, stemming, and TF-IDF calculation. Results demonstrated that the combination of Random Forest and the SMOTE technique significantly contributed to sentiment analysis. The developed model achieved an accuracy of 80%, with precision scores of 83%, 65%, and 83% for negative, neutral, and positive sentiments, respectively. Recall scores were 93%, 55%, and 45% for negative, neutral, and positive sentiments, respectively, while F1-scores were 88%, 59%, and 59%. In conclusion, although the model exhibited optimal performance, especially for negative sentiment, further improvements are needed for a more comprehensive analysis, particularly for neutral sentiment.

KATA PENGANTAR

Puji syukur penulis panjatkan khadirat Allah SWT, yang telah memberikan rahmat dan hidayahNya kepada kita sekalian, sehingga penulis dapat menyelesaikan penulisan skripsi yang berjudul “Analisis Sentimen Pinjaman Online di Indonesia Menggunakan Metode *Random Forest*: Studi Kasus Pada Aplikasi Twitter”. Skripsi ini merupakan tugas akhir yang diajukan untuk memenuhi sebagian persyaratan dalam memperoleh gelar Sarjana Komputer pada Fakultas Teknologi Informasi Universitas Merdeka Pasuruan.

Penulis menyadari bahwa penyusunan skripsi ini tidak akan terwujud tanpa adanya bantuan dan dorongan dari berbagai pihak. Oleh karenanya pada kesempatan ini, penulis mengucapkan rasa terima kasih kepada:

1. Allah SWT yang selalu memberikan kekuatan kepada penulis, khususnya pada saat penyelesaian skripsi ini yang walaupun dalam diri ini banyak sekali berbuat salah dan dosa kepada-Mu.
2. Bapak Muhammad Misdram, S.Kom., M.Kom. selaku Dekan Fakultas Teknologi Informasi Universitas Merdeka Pasuruan.
3. Bapak Dian Ahkam Sani, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Merdeka Pasuruan.
4. Bapak Muslim Alamsyah, S.Kom., M.M., M.Kom. selaku Dosen Pembimbing I dalam penyusunan skripsi ini telah memberikan bimbingan, nasihat, dan ilmunya kepada penulis.

5. Bapak Mochammad Firman Arif, S.Kom., M.Kom. selaku Dosen Pembimbing II dalam penyusunan skripsi ini telah memberikan bimbingan, nasihat, dan ilmunya kepada penulis.
6. Dosen dan Staf Fakultas Teknologi Informasi Universitas Merdeka Pasuruan yang telah membekali penulisan dengan berbagai ilmu selama perkuliahan.
7. Kedua orang tua dan saudaraku tercinta yang selalu memberikan kasih sayang dan dukungan penuh terhadap penulis.
8. Teman-teman Fakultas Teknologi Informasi Angkatan 2020 atas kebersamaan, bantuan dan kerjasamanya selama ini.

Demi kesempurnaan skripsi ini, saran dan kritik yang sifatnya membangun sangat diharapkan penulis. Semoga Skripsi ini dapat bermanfaat bagi semua pihak khususnya Penulis.

Pasuruan, 30 Juli 2024

Muhammad Alif A.
NPM. 2055201001024

DAFTAR ISI

PERSETUJUAN PEMBIMBING SKRIPSI.....	iv
PERSETUJUAN DAN PENGESAHAN SKRIPSI.....	v
PERNYATAAN KEASLIAN TULISAN.....	vi
MOTTO.....	vii
PERSEMBAHAN.....	viii
ABSTRAK	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR TABEL	xvi
DAFTAR GAMBAR	xviii
DAFTAR LAMPIRAN.....	xix
BAB I PENDAHULUAN	
1.1. Latar Belakang	1
1.2. Rumusan Masalah	4
1.3. Batasan Masalah	4
1.4. Tujuan Penelitian	5
1.5. Manfaat Penelitian	5
1.6. Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	
2.1. Penelitian Terkait	7
2.2. Landasan Teori.....	9

2.2.1. Analisis Sentimen.....	9
2.2.2. Pinjaman Online	10
2.2.3. <i>Twitter</i>	10
2.2.4. <i>Text Mining</i>	10
2.2.5. <i>Python</i>	11
2.2.6. <i>Random Forest</i>	12
2.2.7. <i>Text Preprocessing</i>	12
2.2.8. <i>TF – IDF</i>	13
2.2.9. <i>Confusion Matrix</i>	13
2.2.10. <i>Word Cloud</i>	14
BAB III METODOLOGI PENELITIAN	
3.1. Alur Penelitian	15
3.2. Identifikasi Masalah	16
3.3. Studi Literatur	16
3.4. Pengumpulan Data	16
3.5. Analisa Data.....	17
3.6. Perancangan Sistem.....	21
3.7. Perhitungan Manual Metode <i>Random Forest</i>	24
3.8. Evaluasi	39
BAB IV HASIL DAN PEMBAHASAN	
4.1. Dataset	40
4.2. <i>Import Data</i>	40
4.3. <i>Preprocessing Data</i>	41

4.4. Model <i>Random Forest</i>	42
4.5. Visualisasi.....	44
BAB V PENUTUP	
4.1. Kesimpulan.....	46
4.2. Saran.....	46
DAFTAR PUSTAKA	48
LAMPIRAN	51



DAFTAR TABEL

Tabel 2.1. <i>Confusion Matrix</i>	13
Tabel 3.1 Data sentimen <i>Twitter</i> tentang pinjaman online di Indonesia.....	17
Tabel 3.2 Proses <i>Case Folding</i>	18
Tabel 3.3 Proses <i>Remove Punctuation</i>	19
Tabel 3.4 Proses <i>Tokenize</i>	20
Tabel 3.5 Proses <i>Stopword Removal</i>	20
Tabel 3.6 Proses <i>Stemming</i>	21
Tabel 3.7 <i>Sampling</i> data.....	25
Tabel 3.8 <i>Splitting</i> Data.....	26
Tabel 3.9 Distribusi label.....	27
Tabel 3.10 Dokumen hasil <i>TF – IDF</i>	28
Tabel 3.11 Data “pinjol”.....	28
Tabel 3.12 Data “judi”.....	29
Tabel 3.13 Data “online”.....	30
Tabel 3.14 Data “bantu”.....	31
Tabel 3.15 Hasil perhitungan <i>Gini Gain</i>	32
Tabel 3.16 Data “pinjol” subset non “bantu”.....	33
Tabel 3.17 Data “online” subset non “bantu”.....	34
Tabel 3.18 Data “judi” subset non “bantu”.....	35
Tabel 3.19 Hasil <i>gini gain</i> subset non “bantu”.....	36
Tabel 3.20 Hasil <i>gini gain</i> subset “pinjol”.....	37

Tabel 3.21 Hasil <i>gini gain</i> subset non “judi”	37
Tabel 3.22 Hasil dan <i>Voting</i>	39
Tabel 3.23 Evaluasi.....	39



DAFTAR GAMBAR

Gambar 3.1. Alur Penelitian.....	15
Gambar 3.2. Alur <i>Preprocessing Data</i>	18
Gambar 3.3. Perancangan Sistem.....	22
Gambar 3.4. Pohon Keputusan.....	32
Gambar 4.1 Proses <i>Crawling Data Tweets</i>	40
Gambar 4.2 <i>Import Data</i>	40
Gambar 4.3 <i>Import library Sastrawi</i>	41
Gambar 4.4 <i>Import library Pandas, NLTK</i>	41
Gambar 4.5 Implementasi <i>Preprocessing</i> menggunakan <i>Streamlit</i>	42
Gambar 4.6 <i>Library</i> Klasifikasi <i>Random Forest</i>	42
Gambar 4.7 Distribusi data latih menggunakan SMOTE	43
Gambar 4.8 Hasil klasifikasi <i>Random Forest</i>	43
Gambar 4.9 Implementasi <i>library</i> visualisasi.....	44
Gambar 4.10 Implementasi <i>WordCloud</i> menggunakan <i>Streamlit</i>	44
Gambar 4.11 Tabel <i>Confusion Matrix</i>	45

DAFTAR LAMPIRAN

Lampiran 1 <i>Source Code</i>	51
Lampiran 2 <i>Letter Of Acceptance</i>	55
Lampiran 3 Daftar Riwayat Hidup	56

