



LAPORAN SKRIPSI

ANALISIS SENTIMEN MENGENAI PINJAMAN *ONLINE* DI MEDIA SOSIAL *TWITTER* MENGGUNAKAN METODE *NAIVE BAYES*

**ARY FAJAR ISKANDAR
NIM. 201851259**

DOSEN PEMBIMBING

**Muhammad Imam Ghozali, S.Kom., M.Kom
Wibowo Harry Sugiharto, S.Kom., M.Kom**

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MURIA KUDUS
FEBRUARI 2023**

HALAMAN PERSETUJUAN

ANALISIS SENTIMEN MENGENAI PINJAMAN ONLINE DI MEDIA SOSIAL TWITTER MENGGUNAKAN METODE NAIVE BAYES

ARY FAJAR ISKANDAR

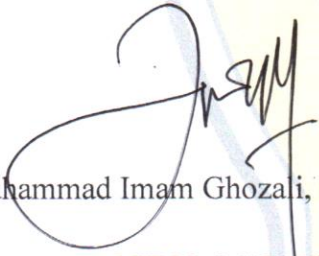
NIM. 201851259

Kudus, 1 Januari 2023

Menyetujui,

Pembimbing Utama,

Pembimbing Pendamping,


Muhammad Imam Ghozali, S.Kom., M.Kom


Wibowo Harry Sugiharto, S.Kom., M.Kom

NIDN. 0618058602

NIDN. 0619059101

HALAMAN PENGESAHAN

ANALISIS SENTIMEN MENGENAI PINJAMAN ONLINE DI MEDIA SOSIAL TWITTER MENGGUNAKAN METODE NAIVE BAYES

ARY FAJAR ISKANDAR

NIM. 201851259

Kudus, 20 Februari 2023

Menyetujui,

Ketua Penguji,



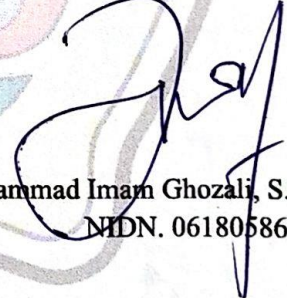
Endang Supriyati, S.Kom., M.Kom
NIDN. 0629077402

Anggota Penguji I,



Tri Listyorini, S.Kom., M.Kom
NIDN. 0616088502

Anggota Penguji II,



Muhammad Imam Ghozali, S.Kom., M.Kom
NIDN. 0618058602

Mengetahui

Dekan Fakultas Teknik



Muhammad Dahlan, ST., MT
NIS. 0610701000001141

Ketua Program Studi
Teknik Informatika

Mukhamad Nurkamid, S.Kom., M.Cs
NIS. 0610701000001212

PERNYATAAN KEASLIAN

Saya yang bertanda tangan dibawah ini :

Nama : Ary Fajar Iskandar
NIM : 201851259
Tempat & Tanggal Lahir : Kudus, 5 Februari 2001
Judul Skripsi : Analisis Sentimen Mengenai Pinjaman Online Di Media Sosial Twitter Menggunakan Metode Naive Bayes

Menyatakan dengan sebenarnya bahwa penulisan Skripsi ini berdasarkan hasil penelitian, pemikiran dan pemaparan asli dari saya sendiri, baik untuk naskah laporan maupun kegiatan lain yang tercantum sebagai bagian dari Skripsi ini. Seluruh ide, pendapat, atau materi dari sumber lain telah dikutip dalam Skripsi dengan cara penulisan referensi yang sesuai.

Demikian pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar dan sanksi lain sesuai dengan peraturan yang berlaku di Universitas Muria Kudus.

Demikian pernyataan ini saya buat dalam keadaan sadar tanpa paksaan dari pihak manapun.

Kudus, 17 Januari 2023

Yang memberi pernyataan,



Ary Fajar Iskandar
NIM. 201851259

KATA PENGANTAR

Segala puji syukur bagi Allah Yang Maha Pengasih dan Penyayang yang telah melimpahkan rahmat dan kasih sayang-Nya, sehingga akhirnya penulis berhasil menyelesaikan penulis laporan skripsi berjudul : "Analisis Sentimen Mengenai Pinjaman Online di Media Sosial Twitter Menggunakan Metode Naive Bayes". Penyusunan Laporan Skripsi ini merupakan salah satu syarat untuk menyelesaikan Program Studi Teknik Informatika pada Fakultas Teknik Universitas Muria Kudus. Dalam proses penyelesaian tidak luput dari berbagai macam bantuan, dukungan, saran, dan kritik yang telah penulis terima, oleh karena itu dalam kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Prof. Dr. Ir. Darsono, M.Si. selaku Rektor Universitas Muria Kudus.
2. Mohammad Dahlan, ST., MT, selaku Dekan Fakultas Teknik Universitas Muria Kudus.
3. Mukhamad Nurkamid, S.Kom, M.Cs, selaku Ketua Program Studi Teknik Informatika Fakultas Teknik Universitas Muria Kudus.
4. Ratih Nindyasari, S.Kom., M.Kom., selaku Koordinator Skripsi.
5. Muhammad Imam Ghozali, S.Kom., M.Kom., selaku pembimbing utama, dan Wibowo Harry Sugiharto, S.Kom., M.Kom., selaku pembimbing pendamping yang telah memberikan bimbingan, ilmu, saran, motivasi, nasihat, pengarahan, dan bimbingan selama penyusunan skripsi ini.
6. Kedua orang tua penulis, yang selalu mendoakan dan selalu memberikan segala dukungan kepada penulis dalam menyelesaikan skripsi ini.

Akhir kata, penulis menyadari adanya kekurangan dan kesalahan dalam penulisan skripsi ini, karena itu penulis menerima dengan baik kritik, saran yang membangun dari pembaca sekalian. Akhirnya penulis berharap semoga skripsi ini bisa bermanfaat khususnya bagi penulis dan bagi para pembaca.

Kudus, Januari 2023

Penulis

ANALISIS SENTIMEN MENGENAI PINJAMAN ONLINE DI MEDIA SOSIAL TWITTER MENGGUNAKAN METODE NAIVE BAYES

Nama mahasiswa : Ary Fajar Iskandar

NIM : 201851259

Pembimbing :

1. Muhammad Imam Ghozali, S.Kom., M.Kom
2. Wibowo Harry Sugiharto, S.Kom., M.Kom

RINGKASAN

Analisis sentimen dilakukan untuk mengukur kecenderungan opini masyarakat terhadap suatu kejadian yang sedang atau telah terjadi. Salah satu kasus yang diangkat untuk dianalisis adalah Pinjaman *Online* atau biasa dikenal dengan singkatan Pinjol. Data penelitian mengenai Pinjaman *Online* di media sosial *Twitter* dengan kata kunci Pinjaman *Online*. Metode analisis yang digunakan yaitu *Naïve Bayes*. Sebelum proses analisis sentimen, dilakukan terlebih dahulu pengambilan data berupa teks dengan metode *crawling* dari API *Twitter* dengan aplikasi Rapidminer, kemudian data tersebut kemudian diolah dengan melakukan *text pre-processing* pada data teks tersebut. Data akan melalui pembobotan TF IDF. Hasil dari penelitian ini yaitu menunjukkan bagaimana kecenderungan konflik sentimen positif dan negatif pada tiap *tweet* oleh para pengguna *Twitter* yang membahas tentang Pinjaman *Online*.

Kata kunci : Pinjaman Online, *Naive Bayes*, Analisis Sentimen, *Twitter*.

SENTIMENT ANALYSIS OF ONLINE LOANS ON SOCIAL MEDIA TWITTER USING NAIVE BAYES METHOD

Student Name : Ary Fajar Iskandar

Student Identity Number : 201851259

Supervisor :

1. Muhammad Imam Ghozali, S.Kom., M.Kom
2. Wibowo Harry Sugiharto, S.Kom., M.Kom

ABSTRACT

Sentiment analysis is carried out to measure the trend of public opinions towards an ongoing or past event. One case that is taken for analysis is Online Loans or commonly known as Pinjol. Research data on Online Loans in social media Twitter with the keyword Online Loans. The analysis method used is Naive Bayes. Before the sentiment analysis process, data collection was first carried out in the form of text using the crawling method from the Twitter API with the Rapidminer application, then the data was processed by performing text pre-processing on the text data. Data will go through TF-IDF weighting. The result of this research is to show how the trend of positive and negative sentiment conflicts in each tweet by Twitter users discussing Online Loans.

Keywords: Online Loans, Naive Bayes, Sentiment Analysis, Twitter.

DAFTAR ISI

HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR	v
RINGKASAN	vi
<i>ABSTRACT</i>	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xii
DAFTAR LAMPIRAN.....	xiii
DAFTAR ISTILAH DAN SINGKATAN	xiv
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah.....	2
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian.....	3
1.5. Manfaat Penelitian.....	3
1.6. Sistematika penulisan.....	3
BAB II TINJAUAN PUSTAKA.....	5
2.1. Tinjauan Terkait	5
2.2. Landasan Teori	6
2.2.1. Pinjaman Online.....	6
2.2.2. <i>Crawling</i>	7
2.2.3. Twitter	7
2.2.4. Diagram Alir	7
2.2.5. Analisis Sentimen	8
2.2.6. <i>Text Preprocessing</i>	9
2.2.7. TF-IDF	9
2.2.8. Teorama Bayes.....	10
2.2.8. Naïve Bayes Classifier	11
2.2.9. Evaluasi	13
2.3. Kerangka Pikir.....	14
BAB III METODOLOGI.....	15

3.1. Tahap Awal	15
3.1.1. Studi Literatur	16
3.1.2. Pengumpulan data	16
3.2. Tahap Pengembangan Penelitian.....	19
3.2.1. Analisis.....	21
3.2.2. Penambangan Data.....	23
3.2.3. Data <i>Training</i> dan Data <i>Testing</i>	24
3.2.4. <i>Text Pre-processing</i>	25
3.2.5. Pembobotan TF-IDF	32
3.2.6. Klasifikasi Naive Bayes	33
3.2.7. Evaluasi dan Validasi	34
3.3. Tahap akhir penelitian	35
3.3.1. Kesimpulan	35
3.3.2. Saran.....	35
BAB IV HASIL DAN PEMBAHASAN.....	37
4.1. Analisis	37
4.2. Penambangan Data	37
4.3. Data <i>Training</i> dan Data <i>Testing</i>	38
4.4. <i>Text Pre-Processing</i>	39
4.5. Pembobotan TF-IDF.....	41
4.5.1 Perhitungan manual TF-IDF	43
4.6. Klasifikasi Dengan Algoritma Naive Bayes.....	44
4.6.1. Perhitungan manual Naive Bayes	48
4.7. Evaluasi dan Validasi	49
BAB V PENUTUP.....	53
5.1. Kesimpulan.....	53
5.2. Saran.....	53
DAFTAR PUSTAKA	54
LAMPIRAN 1	56
LAMPIRAN 2	66
BIODATA PENULIS	70

DAFTAR GAMBAR

Gambar 2. 1 Kerangka Pikir.....	14
Gambar 3. 1 Metodologi Penelitian	15
Gambar 3. 2 Tahap Awal Penelitian	16
Gambar 3. 3 Contoh <i>tweet</i> dari perusahaan pinjol	17
Gambar 3. 4 Contoh <i>tweet</i> dari perusahaan pinjol	17
Gambar 3. 5 Studi Dokumentasi 1	18
Gambar 3. 6 Studi Dokumentasi 2	18
Gambar 3. 7 Diagram Alir Teknik Pengolahan Data.....	19
Gambar 3. 8 Tahap Pengembangan Penelitian	20
Gambar 3. 9 Proses Crawling dengan RapidMiner.....	23
Gambar 3. 10 Data <i>Tweet</i> Hasil <i>Crawling</i>	23
Gambar 3. 11 Diagram Alir Pembagian Data Training dan testing	24
Gambar 3. 12 Pemodelan Data.....	25
Gambar 3. 13 Diagram Alir Proses Text Pre-processing	26
Gambar 3. 14 Diagram Alir Pembobotan TF-IDF	32
Gambar 3. 15 Diagram Alir Proses Klasifikasi.....	33
Gambar 3. 16 Tahap Akhir Penelitian.....	35
Gambar 4. 1 Hasil <i>Crawling</i> di Rapidminer	38
Gambar 4. 2 Data Latih atau Data <i>Training</i>	39
Gambar 4. 3 Data Uji atau Data <i>Testing</i>	39
Gambar 4. 4 <i>Script Cleansing</i>	40
Gambar 4. 5 <i>Script Case Folding</i>	40
Gambar 4. 6 <i>Script Tokenizing</i>	41
Gambar 4. 7 <i>Script Stemming</i>	41
Gambar 4. 8 <i>Script Stopword</i>	41
Gambar 4. 9 <i>Script TfIdf</i>	42
Gambar 4. 10 Data nilai IDF tiap kata	42
Gambar 4. 11 Data TF dan TFIDF Kata Positif dan Negatif	43

Gambar 4. 12 Tampilan <i>input file</i> Data Sentimen.....	44
Gambar 4. 13 Gambar Proses Perhitungan Bobot tiap Data Sentimen.....	45
Gambar 4. 14 Hasil Perhitungan Kalimat Sentimen	46
Gambar 4. 15 hasil analisis sentimen pinjaman <i>online</i>	47
Gambar 4. 16 Input File Data Uji	50
Gambar 4. 17 Hasil Data Uji.....	50
Gambar 4. 18 Perhitungan Kalimat Data Uji	51
Gambar 4. 19 Hasil Perhitungan Data Uji	51



DAFTAR TABEL

Tabel 2. 1. Tabel Diagram Alir	7
Tabel 3. 1. Analisis Kebutuhan Pengguna	21
Tabel 3. 2. Contoh <i>cleansing</i> pada <i>tweet</i>	26
Tabel 3. 3. Contoh <i>case folding</i> pada <i>tweet</i>	28
Tabel 3. 4. Contoh <i>tokenizing</i> pada <i>tweet</i>	29
Tabel 3. 5. Contoh <i>stopword removal</i> pada <i>tweet</i>	29
Tabel 3. 6. Contoh <i>stemming</i> pada <i>tweet</i>	31
Tabel 3. 7. Contoh pembobotan TF-IDF	32
Tabel 4. 1 Nilai IDF	43
Tabel 4. 2 Nilai TFIDF Positif	44
Tabel 4. 3 Nilai TFIDF Positif	44
Tabel 4. 4 Hasil analisis tiap waktu	47
Tabel 4. 5 Data probabilitas posterior	48

DAFTAR LAMPIRAN

Lampiran 1	Hasil Pengujian Sentimen.....	56
Lampiran 2	Lampiran Konsultasi Penelitian	66



DAFTAR ISTILAH DAN SINGKATAN

TF	: Term Frequency
IDF	: Inverse Document Frequency
Pinjol	: Pinjaman Online
Sentimen	: Sentimen (atau dalam bahasa Indonesia dikenal dengan istilah sentimen) adalah pandangan atau perasaan seseorang terhadap suatu hal, seperti produk, merek, layanan, atau kejadian.
Term	: kata-kata atau frasa
Dokumen	: kumpulan teks atau informasi
Prior	: Prior adalah distribusi probabilitas yang digunakan untuk menggambarkan pengetahuan atau keyakinan kita tentang suatu kejadian sebelum melihat data baru.
Posterior	: Posterior adalah distribusi probabilitas hasil dari menggabungkan data baru dengan prior dalam metode Bayes.
Conditional	: Conditional probability (probabilitas kondisional) adalah probabilitas suatu kejadian terjadi, diberikan informasi tentang kejadian lain.