

**ANALISIS SENTIMEN IDIOM INDONESIA PADA TWITTER  
MENGUNAKAN ALGORITMA NAÏVE BAYES**

**TUGAS AKHIR**

**Disusun oleh:**

**Rendi Andre Aldi Pai**

**16013084**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS KATOLIK DE LA SALLE  
MANADO  
2020**

**ANALISIS SENTIMEN IDIOM INDONESIA PADA TWITTER  
MENGUNAKAN ALGORITMA NAÏVE BAYES**

**TUGAS AKHIR**

Ditulis untuk memenuhi sebagian persyaratan untuk mendapatkan  
Gelar Sarjana Teknik pada Program Studi Teknik Informatika

**Disusun oleh:**

**Rendi Andre Aldi Pai**

**16013084**



**PROGRAM STUDI TEKNIK INFORMATIKA  
FAKULTAS TEKNIK  
UNIVERSITAS KATOLIK DE LA SALLE  
MANADO**

**2020**

## LEMBAR PERNYATAAN

Yang bertandatangan di bawah ini:

Nama : Rendi Andre Aldi Pai  
NIM : 16013084  
Tempat/Tanggal Lahir : Tomohon/05 April 1998  
Fakultas/Program Studi : Teknik/Teknik Informatika

Menyatakan bahwa Laporan Tugas Akhir dan Aplikasi Analisis Sentimen Idiom Indonesia pada Twitter Menggunakan Algoritma Naïve Bayes yang telah dibuat merupakan hasil karya saya dan bukan karya tulis orang lain, baik sebagian atau seluruhnya kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya dan apabila pernyataan ini tidak benar maka saya bersedia untuk menerima sanksi akademik sesuai dengan peraturan Fakultas.

Manado, 16 September 2020

Menyatakan,  
  
Rendi Andre Aldi Pai

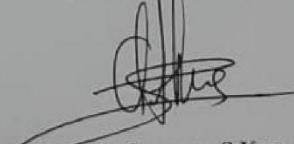
Menyetujui

Dosen Pembimbing I



Dr. Liza Wikarsa, B.C.S., M.Comp.

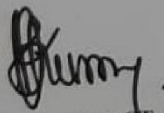
Dosen Pembimbing II



Thomas Christian Suwanto, S.Kom., MMm

Mengetahui

Ketua Program Studi



Vivie deyby Kumenap ST., M.Cs.

Fakultas Teknik  
  
Ronald Albert Rochmadi ST., MT



**UNIVERSITAS KATOLIK DE LA SALLE  
MANADO - INDONESIA**

Nama : Rendi Andre Aldi Pai  
NIM : 16013084  
Fakultas : Teknik  
Program Studi : Teknik Informatika  
Judul Tugas Akhir : Analisis Sentimen Idiom Indonesia pada Twitter  
Menggunakan Algoritma Naïve Bayes  
Pembimbing I : Dr. Liza Wikarsa, B.C.S., M.Comp.  
Pembimbing II : Thomas Christian Suwanto, S.Kom., MMm

Menyetujui

Dosen Pembimbing I

Dr. Liza Wikarsa, B.C.S., M.Comp.

Dosen Pembimbing II

Thomas Christian Suwanto, S.Kom., MMm

Mengetahui

Ketua Program Studi

Vivie deyby Kumenap ST., M.Cs.

Fakultas Teknik



Ronald Albert Rochmadi ST., MT

## **ABSTRACT**

*Idioms are often used in statements, opinions, and sentiments about something or someone in expressions whose meaning cannot be interpreted literally. Twitter is a microblogging social media where users can search for and discuss current issues or topics. In Twitter there is a collection of tweets sent by users from various parts of the world. However, there has been no research on idiom sentiment analysis using Indonesian.*

*This study aimed at identifying and classifying idioms using Indonesian in Twitter using Naïve Bayes Classifier. Sentiments contained within the idioms will be classified into seven, namely happy, surprise, neutral, angry, fear, sad, and disgusted. The four stages of this research are Data Collection, Pre-processing, Processing, and Validation. Experiments were carried out by dividing the results of the data by the percentage of 70% training data and 30% testing data. There are 3 keywords used, namely proverbs, satire, and taunts. From 5 times the experiment, the highest accuracy results were obtained in the first experiment with a percentage of 79.41%. For the second, third and up to fifth experiments the accuracy decreased. The results decreased due to the increasing number of testing data and had not entered the training data.*

*Keywords: Idioms, Sentiment Analysis, Naïve Bayes, Twitter Indonesia.*

## ABSTRAK

Idiom sering digunakan dalam pernyataan, pendapat dan sentimen tentang sesuatu atau seseorang dalam ungkapan yang maknanya tidak dapat dipahami secara harfiah. Twitter adalah mikroblog media sosial tempat pengguna dapat mencari dan mendiskusikan masalah atau topik terkini. Di Twitter terdapat kumpulan tweet yang dikirim oleh pengguna dari seluruh dunia. Namun belum ada penelitian tentang analisis sentimen idiom yang menggunakan bahasa Indonesia

Penelitian ini bertujuan untuk mengidentifikasi dan mengklasifikasikan idiom berbahasa Indonesia di Twitter menggunakan *Naïve Bayes Classifier*. Sentimen yang terkandung dalam idiom akan diklasifikasikan menjadi tujuh, yaitu senang, terkejut, netral, marah, takut, sedih, dan jijik. Empat tahapan dalam penelitian ini adalah Pengumpulan Data, *Pre-processing*, Pengolahan, dan Validasi. Eksperimen yang dilakukan dengan membagi hasil data dengan persentasi 70% data training dan 30% data testing. kata kunci ada 3 yang digunakan yaitu peribahasa, sindiran, dan ejekan. Dari percobaan sebanyak 5x, hasil akurasi tertinggi di dapat pada percobaan pertama dengan persentasi 79.41%. Untuk percobaan yang kedua, ketiga dan sampai kelima hasil akurasinya menurun. Hasil menurun disebabkan karena data testing semakin banyak dan belum masuk pada data *training*.

Kata kunci: Idiom, Analisis Sentimen, Naïve Bayes, Twitter, Indonesia.

## KATA PENGANTAR

Puji dan syukur kepada Tuhan Yang Maha Esa karena atas berkat dan kasih-Nya, penulis dapat menyelesaikan Tugas Akhir ini untuk memenuhi persyaratan dalam mendapatkan gelar Sarjana Teknik pada program studi Teknik Informatika Universitas Katolik De La Salle Manado.

Dalam penyusunan Laporan Tugas Akhir ini, penulis banyak memperoleh bantuan, bimbingan dan dorongan dari berbagai pihak. Penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Bapak Pastor Prof. Dr. Johanis Ohoitumur selaku Rektor Universitas Katolik De La Salle Manado.
2. Bapak Ronald Albert Rachmadi, S.T., M.T. selaku Dekan Fakultas Teknik.
3. Bapak Thomas Christian Suwanto, S.Kom., M.Mm. selaku Wakil Dekan Fakultas Teknik dan Pembimbing 2 yang telah memberikan bantuan dalam penulisan laporan.
4. Ibu Vivie Deyby Kumenap, S.T., M.Cs. selaku Ketua Program Studi Teknik Informatika dan dosen Pembimbing Akademik.
5. Ibu Dr. Liza Wikarsa, BCS., M.Comp. selaku dosen pembimbing 1 yang telah memberikan bimbingan dan arahan yang sangat bermanfaat bagi penulis.
6. Mama, Papa dan Kakak yang telah memberikan banyak doa, dukungan dan semangat yang sangat berguna.
7. Teman-teman angkatan 2016.

Penulis menyadari bahwa dalam penyusunan Tugas Akhir ini, tidak terlepas dari kekurangan. Saran dan kritik dari para pembaca sangat penulis harapkan.

Manado, September 2020

Penulis

## DAFTAR ISI

HALAMAN JUDUL.....	i
LEMBAR PERNYATAAN .....	ii
UNIVERSITAS KATOLIK DE LA SALLE MANADO - INDONESIA.....	iii
<i>ABSTRACT</i> .....	iv
ABSTRAK .....	v
KATA PENGANTAR .....	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR .....	xi
DAFTAR LAMPIRAN.....	xii
DAFTAR RUMUS .....	xiii
BAB I PENDAHULUAN.....	<b>Error! Bookmark not defined.</b>
1.1. Latar Belakang.....	<b>Error! Bookmark not defined.</b>
1.2. Rumusan Masalah.....	<b>Error! Bookmark not defined.</b>
1.3. Tujuan Tugas Akhir.....	<b>Error! Bookmark not defined.</b>
1.4. Manfaat Tugas Akhir.....	<b>Error! Bookmark not defined.</b>
1.5. Batasan Masalah .....	<b>Error! Bookmark not defined.</b>
1.6. Metodologi Penelitian.....	<b>Error! Bookmark not defined.</b>
1.7. Sistematika Penulisan .....	<b>Error! Bookmark not defined.</b>
BAB II STUDI PUSTAKA.....	<b>Error! Bookmark not defined.</b>
2.1. Teori Pendukung.....	<b>Error! Bookmark not defined.</b>
2.1.1. Idiom .....	<b>Error! Bookmark not defined.</b>
2.1.2. Analisis Sentimen.....	<b>Error! Bookmark not defined.</b>
2.1.3. Klasifikasi.....	<b>Error! Bookmark not defined.</b>
2.1.4. Metode Naïve Bayes .....	<b>Error! Bookmark not defined.</b>
2.1.5. <i>One Vs All</i> .....	<b>Error! Bookmark not defined.</b>
2.1.6. <i>Confusion Matrix</i> .....	<b>Error! Bookmark not defined.</b>
2.1.7. Twitter .....	<b>Error! Bookmark not defined.</b>
2.1.8. <i>Data Collection</i> .....	<b>Error! Bookmark not defined.</b>
2.2. Teknologi Pengembangan Perangkat Lunak.....	<b>Error! Bookmark not defined.</b>
2.2.1. Python.....	<b>Error! Bookmark not defined.</b>
2.2.2. MongoDB.....	<b>Error! Bookmark not defined.</b>



2.3.	Penelitian terdahulu .....	<b>Error! Bookmark not defined.</b>
2.4.	Metodologi Pengembangan Perangkat Lunak .....	<b>Error! Bookmark not defined.</b>
2.5.	Kakas Pemodelan yang Digunakan .....	<b>Error! Bookmark not defined.</b>
2.5.1.	Data Flow Diagram (DFD) .....	<b>Error! Bookmark not defined.</b>
2.5.2.	<i>Entity Relationship Diagram</i> (ERD) ...	<b>Error! Bookmark not defined.</b>
2.6.	Prosedur Pengumpulan dan Pengolahan Data .....	<b>Error! Bookmark not defined.</b>
2.6.1	Cara Pengumpulan Data.....	<b>Error! Bookmark not defined.</b>
2.6.2	Pengolahan Data.....	<b>Error! Bookmark not defined.</b>
BAB III ANALISIS .....		<b>Error! Bookmark not defined.</b>
3.1.	Pengumpulan Data.....	<b>Error! Bookmark not defined.</b>
3.1.1.	Data <i>Tweets</i> .....	<b>Error! Bookmark not defined.</b>
3.1.2.	Idiom .....	<b>Error! Bookmark not defined.</b>
3.2.	Daftar Spesifikasi Persyaratan .....	<b>Error! Bookmark not defined.</b>
3.3.	Kerangka Kerja.....	<b>Error! Bookmark not defined.</b>
3.3.1.	<i>Data Collection Stage</i> .....	<b>Error! Bookmark not defined.</b>
3.3.2.	<i>Pre-Processing Stage</i> .....	<b>Error! Bookmark not defined.</b>
3.3.3.	<i>Processing Stage</i> .....	<b>Error! Bookmark not defined.</b>
3.3.4.	<i>Validation Stage</i> .....	<b>Error! Bookmark not defined.</b>
BAB IV PERANCANGAN .....		<b>Error! Bookmark not defined.</b>
4.1.	Desain Sistem Secara Keseluruhan .....	<b>Error! Bookmark not defined.</b>
4.2.	Desain Basis Data .....	<b>Error! Bookmark not defined.</b>
4.3.	Desain Antarmuka .....	<b>Error! Bookmark not defined.</b>
4.3.1.	Beranda.....	<b>Error! Bookmark not defined.</b>
4.3.2.	Daftar Idiom .....	<b>Error! Bookmark not defined.</b>
4.3.3.	Daftar <i>Tweet</i> .....	<b>Error! Bookmark not defined.</b>
4.3.4.	Daftar <i>Jargon</i> .....	<b>Error! Bookmark not defined.</b>
4.3.5.	<i>Tweet</i> .....	<b>Error! Bookmark not defined.</b>
4.3.6.	<i>Pre-processing</i> .....	<b>Error! Bookmark not defined.</b>
4.3.7.	Data <i>testing</i> .....	<b>Error! Bookmark not defined.</b>
4.3.8.	Data <i>training</i> .....	<b>Error! Bookmark not defined.</b>
4.3.9.	<i>Validation</i> .....	<b>Error! Bookmark not defined.</b>
4.4.	Desain Modul Program.....	<b>Error! Bookmark not defined.</b>

BAB V IMPLEMENTASI.....	<b>Error! Bookmark not defined.</b>
5.1. <i>Rule-Based</i> .....	<b>Error! Bookmark not defined.</b>
5.2. Basis Data .....	<b>Error! Bookmark not defined.</b>
5.3. Antarmuka Aplikasi.....	<b>Error! Bookmark not defined.</b>
5.3.1. <i>Tweet</i> .....	<b>Error! Bookmark not defined.</b>
5.3.2. <i>Data Testing</i> .....	<b>Error! Bookmark not defined.</b>
5.3.3. <i>Pre-Processing</i> .....	<b>Error! Bookmark not defined.</b>
5.4. Implementasi Modul Program .....	<b>Error! Bookmark not defined.</b>
 BAB VI PENGUJIAN .....	<b>Error! Bookmark not defined.</b>
6.1. Tujuan Pengujian .....	<b>Error! Bookmark not defined.</b>
6.2. Kriteria Pengujian .....	<b>Error! Bookmark not defined.</b>
6.3. Kasus Pengujian.....	<b>Error! Bookmark not defined.</b>
6.4. Pelaksanaan Pengujian.....	<b>Error! Bookmark not defined.</b>
6.5. Analisis Hasil Pengujian.....	<b>Error! Bookmark not defined.</b>
 BAB VII KESIMPULAN DAN SARAN .....	<b>Error! Bookmark not defined.</b>
7.1 Kesimpulan.....	<b>Error! Bookmark not defined.</b>
7.2 Saran .....	<b>Error! Bookmark not defined.</b>
DAFTAR PUSTAKA .....	<b>Error! Bookmark not defined.</b>
LAMPIRAN	

## DAFTAR TABEL

Tabel 2.1 <i>Confusion Matrix</i> .....	<b>Error! Bookmark not defined.</b>
Tabel 2.2. Notasi DFD Yourdon dan DeMarco ....	<b>Error! Bookmark not defined.</b>
Tabel 2.3. Notasi-notasi ERD Crow .....	<b>Error! Bookmark not defined.</b>
Tabel 2.4. Kardinalitas relasi .....	<b>Error! Bookmark not defined.</b>
Tabel 6.1 Kasus Pengujian.....	<b>Error! Bookmark not defined.</b>
Tabel 6.2 Pelaksanaan Pengujian.....	<b>Error! Bookmark not defined.</b>
Tabel 6.3 Analisis Hasil Pengujian .....	<b>Error! Bookmark not defined.</b>

## DAFTAR GAMBAR

Gambar 2.1	Skema penarikan data <i>tweet</i> pada Twitter.....	<b>Error! Bookmark not defined.</b>
Gambar 3.1	Kerangka Kerja .....	<b>Error! Bookmark not defined.</b>
Gambar 4.1	DFD Level 0.....	<b>Error! Bookmark not defined.</b>
Gambar 4.2	DFD Level 1.....	<b>Error! Bookmark not defined.</b>
Gambar 4.3	Basis Data dengan Notasi <i>Crow</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.4	Tampilan Beranda .....	<b>Error! Bookmark not defined.</b>
Gambar 4.5	Tampilan Daftar Idiom.....	<b>Error! Bookmark not defined.</b>
Gambar 4.6	Tampilan Daftar <i>Tweet</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.7	Tampilan Daftar <i>Jargon</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.8	Tampilan <i>Tweet</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.9	Tampilan <i>Pre-Processing</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.10	Tampilan Data <i>Testing</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.11	Tampilan Data <i>Training</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 4.12	Tampilan Validation.....	<b>Error! Bookmark not defined.</b>
Gambar 4.13	<i>Flowchart</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 5.1	Basis Data Klasifikasi Idiom.....	<b>Error! Bookmark not defined.</b>
Gambar 5.2	<i>Collection Idiom_Collection</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 5.3	<i>Collection Jargon</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 5.4	<i>Collection Pre-processing</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 5.5	<i>Collection Result</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 5.6	<i>Collection Training</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 5.7	<i>Collection Testing</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 5.8	<i>Collection Tweet_Collection</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 5.9	<i>Collection Tweets</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 5.10	Tampilan <i>Tweet</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 5.11	Tampilan Data <i>Testing</i> .....	<b>Error! Bookmark not defined.</b>
Gambar 5.12	Tampilan <i>Pre-Processing</i> .....	<b>Error! Bookmark not defined.</b>

## DAFTAR LAMPIRAN

LAMPIRAN A gambar-gambar pemodelan DFD ..... **A-Error! Bookmark not defined.**

LAMPIRAN B *User Acceptance Testing* (UAT) ..... **B-Error! Bookmark not defined.**

## DAFTAR RUMUS

Rumus 1 <i>Conditional Probabilities</i> .....	7
Rumus 2 Peluang Kelas .....	7
Rumus 3 Peluang Kata .....	7
Rumus 4 <i>Accuracy</i> .....	<b>Error! Bookmark not defined.</b>
Rumus 5 <i>Precision</i> .....	<b>Error! Bookmark not defined.</b>
Rumus 6 <i>Recall</i> .....	<b>Error! Bookmark not defined.</b>
Rumus 7 <i>F-Measure</i> .....	13

