



**ANALISIS SENTIMEN MASYARAKAT TERHADAP DATA
TWEET TRAVELOKA SELAMA *RAPID TEST* ANTIGEN
MENGUNAKAN ALGORITMA *NAIVE BAYES***

SKRIPSI

FENNY NOVIANTI

18141009P

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS TEKNOLOGI
UNIVERSITAS BINA DARMA
PALEMBANG**

2023



**ANALISIS SENTIMEN MASYARAKAT TERHADAP DATA
TWEET TRAVELOKA SELAMA *RAPID TEST* ANTIGEN
MENGUNAKAN ALGORITMA *NAIVE BAYES***

**FENNY NOVIANTI
18141009P**

**Skripsi ini diajukan sebagai salah satu syarat memperoleh
gelar Sarjana Komputer**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS TEKNOLOGI
UNIVERSITAS BINA DARMA
PALEMBANG
2023**

HALAMAN PENGESAHAN

**ANALISIS SENTIMEN MASYARAKAT TERHADAP DATA
TWEET TRAVELOKA SELAMA RAPID TEST ANTIGEN
MENGUNAKAN ALGORITMA NAIVE BAYES**

FENNY NOVIANTI

18141009P

**Telah diterima sebagai salah satu syarat untuk memperoleh gelar
Sarjana Komputer pada Program Studi Sistem Informasi**

Pembimbing



Kiky Rizky Nova Wardani, M.Kom

Palembang, 02 Maret 2023

Fakultas Sains Teknologi

Universitas Bina Darma

Dekan,



Fakultas Sains Teknologi

Dr. Tata Sutabri, S.Kom., MMSI., MKM.

HALAMAN PERSETUJUAN

Skripsi Berjudul "Analisis Sentimen Masyarakat Terhadap Data Tweet Traveloka Selama Rapid Test Antigen Menggunakan Algoritma Naive Bayes" Oleh "Fenny Novianti", telah dipertahankan di depan komisi penguji pada hari Kamis tanggal 02 Maret 2023.

Komisi Penguji

1. Ketua : Kiky Rizky Nova Wardani, M.Kom

(.....)

2. Anggota : Susan Dian Purnamasari, M.Kom

(.....)

3. Anggota : Ferdiansyah, M.Kom., CEH

(.....)

Mengetahui,
Program Studi Sistem Informasi
Fakultas Ilmu Komputer
Universitas Bina Darma
Ketua,

Universitas Bina Darma
Fakultas Sains Teknologi

Nita Rosa Damayanti, M.Kom.,Ph.D.

HALAMAN PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Fenny Novianti

NIM : 18141009P

Dengan ini menyatakan bahwa:

1. Karya tulis saya (Skripsi) adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana) di Universitas Bina Darma atau perguruan tinggi lainnya ;
2. Karya tulis ini murni gagasan, rumusan dan penelitian saya dengan arahan dari tim pembimbing ;
3. Di dalam karya tulis ini tidak terdapat karya atau pendapat yang telah ditulis atau di publikasikan orang lain, kecuali secara tertulis dengan jelas dikutip dengan mencantumkan nama pengarang dan memasukkan ke dalam daftar rujukan ;
4. Saya bersedia tugas skripsi, di cek keasliannya menggunakan plagiarism checker serta di unggah ke internet, sehingga dapat diakses secara daring ;
5. Surat pernyataan ini saya tulis dengan sungguh-sungguh dan apabila terbukti melakukan penyimpangan atau ketidakbenaran dalam pernyataan ini maka saya bersedia menerima sanksi dengan peraturan dan perundang-undangan yang berlaku ;

Demikian surat pernyataan ini saya buat agar dapat dipergunakan sebagaimana mestinya.

Palembang, 01 Maret 2023

Yang membuat pernyataan,



Fenny Novianti

NIM 18141009P

MOTTO DAN PERSEMBAHAN

MOTTO

“Belajarlah dari sebuah kegagalan tanpa sadar dengan kegagalan yang terdahulu sekarang kesuksesan sudah sampai” (Fenny Novianti, 2023)

PERSEMBAHAN

Skripsi ini kupersembahkan teruntuk:

1. Allah SWT, terima kasih atas segala rahmat dan hidayah-Mu, laporan skripsi ini dapat terselesaikan dengan baik ;
2. Kedua orangtua bapak H. Hasan Sumiadi dan ibu Hj. Husna yang senantiasa selalu memberikan do'a dan dukungan yang diberikan ;
3. Almarhum kakak kandungku H. Soehendra Wiratama, kakak iparku Mba Dewi dan Abang Azlan serta Adek Naeema yang selalu memberikan nasehat, semangat dan memberikan kasih sayang kepadaku ;
4. Orang terdekat adek ike sapitri dan teman-teman
5. Almamater Universitas Bina Darma

ABSTRAK

Traveloka melalui akun twitter selama pandemi, banyak mengeluarkan pendapat berupa komentar-komentar tentang syarat untuk melakukan setiap perjalanan baik di luar maupun di dalam negeri yaitu surat dari hasil *rapid test* antigen maupun PCR dengan dinyatakan negatif Covid-19. Rumusan masalah dari penelitian ini adalah bagaimana cara mendapatkan data analisis sentimen dari media sosial Twitter pada data *tweet* Traveloka selama *rapid test* antigen dengan algoritma *naïve bayes*. Penelitian ini merupakan penelitian kuantitatif dengan metode yang digunakan dalam penelitian ini adalah analisis sentimen masyarakat terhadap data *tweet* Traveloka selama *rapid test* antigen berdasarkan dari opini masyarakat atau netizen pada Twitter. Sementara itu, analisis pada penelitian ini menggunakan *Naïve Bayes*. Dalam melakukan pengukuran pada analisis sentimen didapatkan hasil dari pembobotan *TF-IDF* serta hasil akurasi algoritma *Naïve Bayes* sebesar 75% dengan membagi data *training* dan data *testing* dengan rasio 80:20 serta mendapatkan nilai *recall* sebesar 67%, nilai *precision* sebesar 86%, dan nilai *f1-score* sebesar 76%.

Kata Kunci : Traveloka, Twitter, Pandemi, Rapid Test, Naïve Bayes

ABSTRACT

Traveloka, through its Twitter account, during the pandemic, issued many opinions in the form of comments regarding the requirements for each trip both outside and within the country, namely letters from the results of rapid antigen and PCR tests which stated negative for Covid-19. The problem formulation of this research is how to obtain sentiment analysis data from social media Twitter on Traveloka tweet data during the antigen rapid test with the Naïve Bayes algorithm. This research is a quantitative research with the method used in this study is the analysis of public sentiment towards Traveloka tweet data during the antigen rapid test based on public opinion or netizens on Twitter. Meanwhile, the analysis in this study uses Naïve Bayes. In measuring the sentiment analysis, the results obtained from the TF-IDF weighting and the accuracy of the Naïve Bayes algorithm were 75% by dividing the training data and data testing with a ratio of 80:20 and obtaining a recall value of 67%, a precision value of 86%, and a value f1-score of 76%.

Keywords : *Traveloka, Twitter, Pandemic, Rapid Test, Naïve Bayes*

KATA PENGANTAR

Assalamu'alaikum Wr.Wb

Puji Syukur penulis panjatkan kepada Allah SWT yang telah memberikan rahmat, kekuatan serta hidayah-Nya untuk menyelesaikan skripsi dengan judul **“ANALISIS SENTIMEN MASYARAKAT TERHADAP DATA *TWEET* TRAVELOKA SELAMA *RAPID TEST* ANTIGEN MENGGUNAKAN ALGORITMA NAÏVE BAYES”**. Penulis menyadari bahwa dalam penyusunan skripsi ini masih banyak kekurangan baik pengetahuan maupun pengalaman yang dimiliki. Penulis berharap skripsi ini bisa menjadi referensi bagi pembaca dan dapat memenuhi salah satu syarat wajib untuk mendapatkan gelar Sarjana (Strata Satu) jurusan Sistem Informasi Fakultas Ilmu Komputer di Universitas Bina Darma Palembang.

Pada kesempatan yang baik ini, penulis tidak dapat lepas dari banyak pihak yang telah ikut berperan dalam memberikan bimbingan, dukungan serta semangatnya agar skripsi ini dapat selesai dengan baik dan tepat waktu. Maka pada kesempatan kali ini penulis ingin mengucapkan terima kasih kepada :

1. Ibu Dr.Sunda Ariana M.Pd., M.M selaku Rektor Universitas Bina Darma Palembang.
2. Bapak Dr. Tata Sutabari, S.Kom, MMSI, MKM selaku Dekan Fakultas Ilmu Komputer.
3. Ibu Nita Rosa Damayanti, M.Kom., Ph.D selaku Ketua Program Studi Sistem Informasi dan Ibu Eka Puji Agustin S.Kom., M.M sebagai Sekretaris Program Studi Sistem Informasi.
4. Ibu Linda Atika, M.Kom selaku Dosen Pembimbing Akademik (PA) yang telah memberikan arahan dan bimbingannya.
5. Ibu Kiky Rizky Nova Wardani, M.Kom. selaku Dosen Pembimbing yang telah memberikan bimbingan kepada penulis dan terimakasih atas bantuan yang telah diberikan selama menjalani masa studi.

6. Ibu Susan Dian Purnamasari, M.Kom selaku Dosen Penguji I dan Bapak Ferdiansyah, M.Kom., CEH selaku Dosen Penguji II yang telah memberikan saran dan arahan bagi penulis agar dapat menyelesaikan perbaikan skripsi ini dengan baik.
7. Kedua orang tua tercintaku, serta seluruh keluarga besarku yang selalu memberikan kasih sayang, doa, dukungan dan semangat serta motivasi kepada penulis sehingga dapat menyelesaikan skripsi dan berhasil menyelesaikan pendidikan dengan baik.
8. Semua pihak yang telah membantu hingga terselesaikannya pembuatan skripsi ini yang tidak dapat disebutkan satu persatu.

Penulis menyadari pembuatan skripsi ini masih banyak kekurangan dan keterbatasan yang dimiliki serta masih jauh dari kata sempurna. Oleh karena itu, segala kritik dan saran yang membangun akan menyempurnakan penulisan skripsi ini serta dapat bermanfaat bagi penulis dan pembaca. Atas segala bantuan do'a yang telah diberikan penulis mengucapkan terima kasih.

Wassalamu'alaikum Wr.Wb.

Palembang, 01 Maret 2023

Fenny Novianti
18141009P

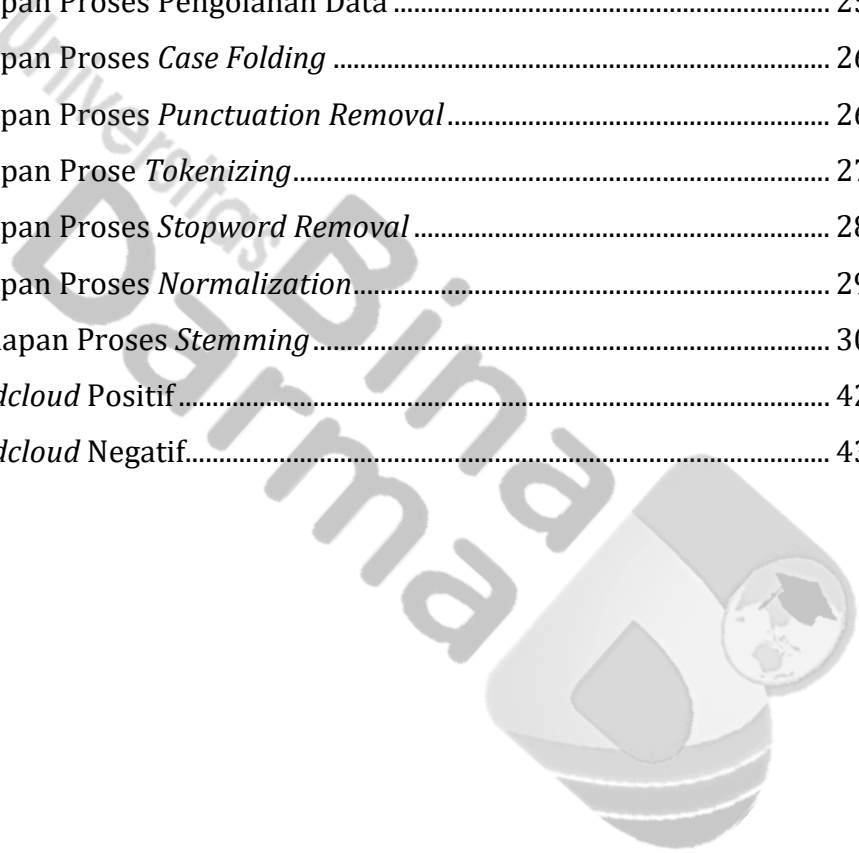
DAFTAR ISI

| | |
|--|------|
| HALAMAN JUDUL..... | i |
| HALAMAN PENGESAHAN..... | ii |
| HALAMAN PERSETUJUAN..... | iii |
| HALAMAN PERNYATAAN..... | iv |
| HALAMAN MOTTO DAN PERSEMBAHAN..... | v |
| ABSTRAK..... | vi |
| ABSTRACT..... | vii |
| KATA PENGANTAR..... | viii |
| DAFTAR ISI..... | x |
| DAFTAR GAMBAR..... | xii |
| DAFTAR TABEL..... | xiii |
| BAB I PENDAHULUAN | 1 |
| 1.1 Latar Belakang..... | 1 |
| 1.2 Perumusan Masalah..... | 4 |
| 1.2.1 Identifikasi Masalah..... | 4 |
| 1.2.2 Rumusan Masalah..... | 4 |
| 1.3 Batasan Masalah..... | 4 |
| 1.4 Tujuan dan Manfaat Penelitian..... | 5 |
| 1.4.1 Tujuan Penelitian..... | 5 |
| 1.4.2 Manfaat Penelitian..... | 5 |
| 1.5 Sistematika Penulisan..... | 6 |
| BAB II TINJAUAN PUSTAKA | 7 |
| 2.1 Media Sosial..... | 7 |
| 2.2 <i>Twitter</i> | 7 |
| 2.3 Traveloka..... | 8 |
| 2.4 <i>Rapid Test</i> Antigen..... | 8 |
| 2.5 <i>Data Mining</i> | 9 |
| 2.6 <i>Text Mining</i> | 10 |
| 2.7 Analisis Sentimen..... | 11 |
| 2.8 <i>Natural Language Processing</i> | 11 |
| 2.9 <i>Text Processing</i> | 11 |
| 2.10 Algoritma <i>Naïve Bayes</i> | 13 |
| 2.11 Evaluasi Performansi..... | 13 |
| 2.11.1 <i>Accuracy</i> | 14 |
| 2.11.2 <i>Precision</i> | 14 |
| 2.11.3 <i>Recall</i> | 15 |
| 2.11.4 <i>F1-Score</i> | 15 |
| 2.12 <i>Term Frequency-Inverse Document Frequency (TF-IDF)</i> | 15 |
| 2.13 <i>Python</i> | 16 |
| 2.14 <i>Jupyter Notebook</i> | 17 |
| 2.15 Penelitian Terdahulu..... | 17 |

| | |
|--|----|
| BAB III METODE PENELITIAN | 20 |
| 3.1 Tahapan Penelitian..... | 20 |
| 3.2 Pengumpulan Data..... | 23 |
| 3.2.1 Tahapan Mendapatkan API Key Twitter..... | 23 |
| 3.2.2 Tahapan Pengambilan Data di Twitter..... | 24 |
| 3.3 Pengelolaan Pada Data..... | 25 |
| 3.3.1 Tahapan <i>Case Folding</i> | 26 |
| 3.3.2 Tahapan <i>Punctuation Removal</i> | 26 |
| 3.3.3 Tahapan <i>Tokenizing</i> | 27 |
| 3.3.4 Tahapan <i>Stopword Removal</i> | 27 |
| 3.3.5 Tahapan <i>Normalization</i> | 28 |
| 3.3.6 Tahapan <i>Stemming</i> | 29 |
| BAB IV HASIL DAN PEMBAHASAN | 31 |
| 4.1 Analisa Studi Kasus..... | 31 |
| 4.2 Pengumpulan Data..... | 31 |
| 4.3 Pelabelan Data..... | 32 |
| 4.4 <i>Preprocessing Data</i> | 32 |
| 4.4.1 <i>Case Folding</i> | 33 |
| 4.4.2 <i>Punctuation Removal</i> | 33 |
| 4.4.3 <i>Tokenizing</i> | 34 |
| 4.4.4 <i>Stopword Removal</i> | 35 |
| 4.4.5 <i>Normalization</i> | 35 |
| 4.4.6 <i>Stemming</i> | 36 |
| 4.5 Pembobotan TF-IDF..... | 37 |
| 4.6 Hasil Klasifikasi Menggunakan Algoritma <i>Naïve Bayes</i> | 39 |
| 4.7 Analisis Sentimen..... | 40 |
| 4.8 Hasil Evaluasi Performansi..... | 40 |
| 4.9 <i>Wordcloud</i> | 42 |
| BAB V KESIMPULAN DAN SARAN | 44 |
| 5.1 Kesimpulan..... | 44 |
| 5.2 Saran..... | 45 |
| DAFTAR PUSTAKA | 46 |
| LAMPIRAN | |

DAFTAR GAMBAR

| | |
|--|----|
| Gambar 3.1 Tahapan Penelitian..... | 20 |
| Gambar 3.2 Tahapan Proses API Key Twitter..... | 24 |
| Gambar 3.3 Proses Pengambilan Data Twitter..... | 24 |
| Gambar 3.4 Tahapan Proses Pengolahan Data..... | 25 |
| Gambar 3.5 Tahapan Proses <i>Case Folding</i> | 26 |
| Gambar 3.6 Tahapan Proses <i>Punctuation Removal</i> | 26 |
| Gambar 3.7 Tahapan Proses <i>Tokenizing</i> | 27 |
| Gambar 3.8 Tahapan Proses <i>Stopword Removal</i> | 28 |
| Gambar 3.9 Tahapan Proses <i>Normalization</i> | 29 |
| Gambar 3.10 Tahapan Proses <i>Stemming</i> | 30 |
| Gambar 4.1 <i>Wordcloud</i> Positif..... | 42 |
| Gambar 4.2 <i>Wordcloud</i> Negatif..... | 43 |



DAFTAR TABEL

| | |
|---|----|
| Tabel 2.1 <i>Confusion Matrix</i> | 14 |
| Tabel 2.2 Penelitian Terdahulu | 17 |
| Tabel 4.1 Data <i>Tweets</i> yang sudah di label | 32 |
| Tabel 4.2 <i>Case Folding</i> | 33 |
| Tabel 4.3 <i>Punctuation Removal</i> | 34 |
| Tabel 4.4 <i>Tokenizing</i> | 34 |
| Tabel 4.5 <i>Stopword Removal</i> | 35 |
| Tabel 4.6 <i>Normalization</i> | 36 |
| Tabel 4.7 <i>Stemming</i> | 36 |
| Tabel 4.8 Perhitungan TF (<i>Term Frequency</i>) | 38 |
| Tabel 4.9 Perhitungan DF (<i>Document Frequency</i>) | 38 |
| Tabel 4.10 Perhitungan IDF (<i>Inverse Document Frequency</i>)..... | 39 |
| Tabel 4.11 Perhitungan TF-IDF | 39 |
| Tabel 4.12 <i>Ration Testing</i> | 40 |
| Tabel 4.13 Hasil Evaluasi Performansi | 41 |