

***EMOTION RECOGNITION* PADA KOMENTAR YOUTUBE  
MENGUNAKAN METODE *NAÏVE BAYES* (STUDI KASUS  
DITUTUPNYA TIKTOK SHOP)**

**SKRIPSI**

Diajukan untuk memenuhi sebagian persyaratan  
dalam memperoleh gelar Sarjana Komputer  
Program Studi Sistem Informasi



**Disusun Oleh:**

**RAMA ARIYA CANDRA**

**20082010135**

**PROGRAM STUDI SISTEM INFORMASI  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"  
JAWA TIMUR  
SURABAYA**

**2024**

**SKRIPSI**

**EMOTION RECOGNITION PADA KOMENTAR YOUTUBE MENGGUNAKAN  
METODE NAÏVE BAYES (STUDI KASUS DITUTUPNYA TIKTOK SHOP)**

**Disusun oleh :**

**RAMA ARIYA CANDRA**  
**20082010135**

**Telah dipertahankan di hadapan dan diterima oleh Tim Penguji Skripsi Program Studi  
Sistem Informasi Fakultas Ilmu Komputer Universitas Pembangunan Nasional  
"Veteran" Jawa Timur  
Pada Tanggal 26 April 2024**

**Pembimbing :**

1.



**Eka Dyar Wahyuni, S.Kom., M.Kom.**  
**NIP. 19841201 2021212 005**

2.



**Rizka Hadiwiyanti, S.Kom, M.Kom.MBA**  
**NIP.19860727 2018032 001**

**Tim Penguji :**

1.



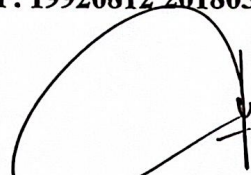
**Dr. Eng Agussalim, S.Pd., M.T.**  
**NIP. 19850811 2019031 005**

2.



**Amalia Anjani Arifiyanti, S.Kom., M.Kom.**  
**NIP. 19920812 2018032 001**

3.




**Dhian Satria Yudha Kartika, S.Kom., M.Kom.**  
**NPT. 201198 60 522249**

**Mengetahui,**

**Dekan Fakultas Ilmu Komputer**

**Universitas Pembangunan Nasional "Veteran" Jawa Timur**


**Prof. Dr. Ir. Novirina Hendrasarie, MT.**  
**NIP. 19681126 199403 2 001**

**LEMBAR PENGESAHAN**

**EMOTION RECOGNITION PADA KOMENTAR YOUTUBE MENGGUNAKAN  
METODE *NAÏVE BAYES* (STUDI KASUS DITUTUPNYA TIKTOK SHOP)**

**Disusun oleh :**

**RAMA ARIYA CANDRA**  
**20082010135**

**Telah disetujui mengikuti Ujian Negara Lisan Gelombang April Periode 2024 pada  
Tanggal 26 April 2024**

**Menyetujui,**

**Dosen Pembimbing 1**



**Eka Dyar Wahyuni, S.Kom., M.Kom.**  
**NIP. 19841201 2021212 005**

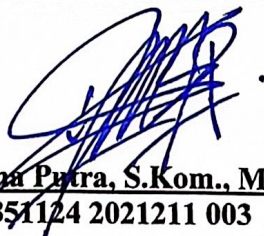
**Dosen Pembimbing 2**



**Rizka Hadiwiyanti, S.Kom, M.Kom.MBA**  
**NIP.19860727 2018032 001**

**Mengetahui,**

**Ketua Program Studi Sistem Informasi  
Fakultas Ilmu Komputer  
Universitas Pembangunan Nasional "Veteran" Jawa Timur**



**Agung Brastama Putra, S.Kom., M.Kom**  
**NIP. 19851124 2021211 003**



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"  
JAWA TIMUR  
FAKULTAS ILMU KOMPUTER

KETERANGAN REVISI

Kami yang bertanda tangan di bawah ini menyatakan bahwa mahasiswa berikut :

Nama : Rama Ariya Candra

NPM : 20082010135

Program Studi : Sistem Informasi

Telah mengerjakan revisi Ujian Negara Lisan Skripsi pada tanggal 26 April 2024 dengan judul:

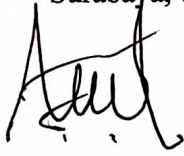
**EMOTION RECOGNITION PADA KOMENTAR YOUTUBE MENGGUNAKAN  
METODE NAÏVE BAYES (STUDI KASUS DITUTUPNYA TIKTOK SHOP)**


Oleh karenanya, mahasiswa tersebut dinyatakan bebas revisi Ujian Negara Lisan Skripsi dan diijinkan untuk membukukan laporan Skripsi dengan judul tersebut.


Surabaya, 15 Mei 2024

Dosen penguji yang memeriksa revisi :

1. Dr. Eng Agussalim, S.Pd., M.T.  
NIP. 19850811 2019031 005
2. Amalia Anjani Arifiyanti, S.Kom., M.Kom.  
NIP. 19920812 2018032 001
3. Dhian Satria Yudha Kartika, S.Kom., M.Kom.  
NPT. 201198 60 522249

{  }

{  }

{  }

Mengetahui,

Dosen Pembimbing 1

Dosen Pembimbing 2

  
Eka Dyar Wahyuni, S.Kom., M.Kom.  
NIP. 19841201 2021212 005

  
Rizka Hadiwiyanti, S.Kom, M.Kom.MBA  
NIP.19860727 2018032 001



**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN  
UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN"  
JAWA TIMUR  
FAKULTAS ILMU KOMPUTER**

**SURAT PERNYATAAN**

Saya yang bertanda tangan di bawah ini :

Nama : Rama Ariya Candra

NPM : 20082010135

Program Studi : Sistem Informasi

Menyatakan bahwa judul/ Tugas Akhir berikut :

***EMOTION RECOGNITION* PADA KOMENTAR YOUTUBE MENGGUNAKAN  
METODE *NAÏVE BAYES* (STUDI KASUS DITUTUPNYA TIKTOK SHOP)**

Bukan merupakan plagiat dari Skripsi/ Tugas Akhir/ Penelitian orang lain dan juga bukan Produk/ Hasil Karya yang saya beli dari orang lain.

Saya juga menyatakan bahwa Skripsi/ Tugas Akhir ini adalah pekerjaan saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional "Veteran" Jawa Timur maupun Instansi Pendidikan lain. Jika dinyatakan dikemudian hari pernyataan tersebut terbukti benar, maka saya bertanggung jawab penuh dan siap menerima segala konsekuensi, termasuk pembatalan ijazah dikemudian hari.

Surabaya, 15 Mei 2024

Hormat Saya,



**Rama Ariya Candra**  
**NPM. 20082010135**

**Judul** : *Emotion Recognition* Pada Komentar Youtube Menggunakan Metode *Naive Bayes* (Studi Kasus Ditutupnya Tiktok Shop)

**Penulis** : Rama Ariya Candra

**Pembimbing 1** : Eka Dyar Wahyuni, S.Kom, M.Kom

**Pembimbing 2** : Rizka Hadiwiyanti, S.Kom, M.Kom.MBA

---

### ABSTRAK

TikTok Shop adalah *fitur* di *platform* TikTok yang memungkinkan pengguna melihat dan membeli produk langsung dari aplikasi tanpa keluar dari aplikasi. Ditutupnya TikTok Shop dikarenakan adanya kecurigaan hadirnya *Project S* yang merugikan UMKM serta melanggar Permendag nomor 31 tahun 2023 karena melakukan transaksi jual beli pada platform media sosial Tiktok. Masyarakat memberikan tanggapan dan pendapat yang berbeda di komentar YouTube tentang ditutupnya TikTok Shop. Ada yang memberikan komentar marah, sedih, senang, takut, dan terkejut. Dari sinilah diperlukannya klasifikasi terhadap kebijakan pemerintah ditutupnya tiktok shop untuk melihat tanggapan atau komentar dari masyarakat mengenai kebijakan tersebut. Untuk melakukannya diperlukannya suatu proses untuk mengklasifikasikan atau mengkategorikan komentar tersebut sesuai emosi, proses tersebut di sebut *Emotion Recognition*. Dengan begini pemerintah dapat mengetahui secara cepat dan tepat emosional masyarakat melalui data komentar sehingga dapat mengambil keputusan untuk kepentingan masyarakat. *Emotion Recognition* adalah proses mengklasifikasikan atau mengkategorikan komentar tersebut sesuai emosi. Oleh karena itu, *Emotion Recognition* dapat digunakan sebagai refrensi untuk pengambilan keputusan oleh *stakeholder*. *Emotion Recognition* dilakukan menggunakan metode *Naive Bayes*

dan pembobotan kata *Term Frequency – Inverse Document Frequency* (TF-IDF). Dalam *Emotion recognition* untuk mencari akurasi terbaik dilakukan 5 skenario yaitu, skenario 1 dengan *training* model menggunakan hyperparameter ( $\alpha=0.1$ ,  $\text{fit\_prior}=\text{True}$ ) menggunakan model multinomial Naive Bayes. Skenario 2 dengan dilakukan *training* model menggunakan hyperparameter ( $\alpha=1.0$ ,  $\text{binarize}=0.0$ ,  $\text{fit\_prior}=\text{True}$ ) menggunakan model bernoulli Naive Bayes. Skenario 3 dengan *training* model menggunakan Bernaoulli Naive Bayes tanpa hyperparameter. Skenario 4 dengan dilakukan *training* model menggunakan Bernaoulli Naive Bayes dengan hyperparameter ( $\alpha=1.0$ ,  $\text{binarize}=0.5$ ,  $\text{fit\_prior}=\text{False}$ ). Skenario 5 dengan dilakukan *training* model menggunakan Multinomial Naive Bayes tanpa hyperparameter. Model dengan accuracy terbaik yaitu pada skenario 1 dengan *accuracy* 82 %. Hasil dari uji validasi dilakukan dengan menggunakan 115 data baru, 17 emosi marah dari 38 data yang bernilai benar, 15 emosi sedih dari 16 data bernilai benar, 13 emosi senang dari 26 data bernilai benar, 14 emosi takut dari 17 data bernilai benar, dan 9 emosi terkejut dari 18 data yang bernilai benar. Sehingga *presentase accuracy* model adalah 0.78 atau 78%.

**Kata Kunci:** Naive Bayes, Emotion Recognition, Ditutupnya Tiktok Shop, TF-IDF

## KATA PENGANTAR

Segala puji syukur penulis panjatkan atas kehadiran Allah SWT karena atas berkat dan hidayah dari Allah SWT penulis mampu menyelesaikan skripsi yang berjudul “*Emotion Recognition* Pada Komentar Youtube Menggunakan Metode *Naive Bayes* (Studi Kasus ditutupnya TikTok Shop)” sebagai syarat untuk menyelesaikan program studi Sistem Informasi di Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Dalam upaya menyelesaikan skripsi ini, tidak dapat dilepaskan dari dukungan yang diterima dari berbagai pihak, dan dalam kesempatan ini penulis ingin menyampaikan ucapan terima kasih kepada:

1. Kedua orang tua yang senantiasa doa dan restunya tidak pernah terputus, sehingga penulis terus termotivasi dalam menyelesaikan skripsi ini.
2. Bapak Asif Faroqi, S.Kom, M.Kom. selaku dosen wali yang selalu memberikan motivasi dan arahan selama menjalani pendidikan di Sistem Informasi UPN “Veteran” Jawa Timur
3. Bu Eka Dyar Wahyuni, S.Kom, M.Kom. dan Bu Rizka Hadiwiyanti, S.Kom, M.Kom, MBA. Selaku dosen pembimbing skripsi yang telah membimbing dengan sabar, memberikan ilmu berharga serta motivasi yang sangat bermanfaat sehingga penulis dapat menyelesaikan skripsi ini.
4. Seluruh bapak dan ibu dosen program studi Sistem Informasi UPN “Veteran” Jawa Timur yang telah memberikan banyak ilmu pengetahuan yang bermanfaat selama masa perkuliahan.



5. Terima kasih kepada yanti, wildan, dan orang-orang baik yang senantiasa meluangkan waktunya untuk membantu saya dalam pengerjaan skripsi ini.
6. Terima kasih kepada Sistem Informasi Angkatan 2020 (ATENSI) yang telah memberikan dukungan dan berjuang bersama mulai dari masa orientasi maba hingga saat ini.
7. Terima Kasih kepada Igo dan Hafid yang selalu memberikan dukungan, motivasi, hiburan, dan candaan “Tempat Singgah” sehingga saya selalu merasa bahagia dan nyaman ketika mengerjakan skripsi ini.
8. Terima kasih kepada seluruh teman-teman yang selalu mendukung dan memberikan masukan dan semangat hingga saat ini.

Semoga Allah SWT membalas kebaikan pihak-pihak diatas. Dalam pengerjaan skripsi ini penulis menyadari masih banyak sekali kekurangan. Semoga skripsi ini bermanfaat bagi semua pihak.

Surabaya, 26 April 2024

Penulis

## DAFTAR ISI

ABSTRAK .....	i
KATA PENGANTAR .....	vii
DAFTAR ISI.....	ix
DAFTAR GAMBAR .....	xii
DAFTAR TABEL.....	xv
BAB I PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	9
1.3 Batasan Masalah.....	9
1.4 Tujuan.....	9
1.5 Sistematika Penulisan.....	10
BAB II TINJAUAN PUSTAKA.....	12
2.1 Dasar Teori .....	12
2.1.1 Youtube .....	12
2.1.2 <i>Text Mining</i> .....	12
2.1.3 Deteksi Emosi .....	13
2.1.4 Pelabelan .....	13
2.1.5 <i>Exploratory Analysis Data (EDA)</i> .....	14
2.1.6 <i>Preprocessing Text</i> .....	14
2.1.7 Oversampling .....	14
2.1.8 TF-IDF .....	15
2.1.9 <i>Confusion Matrix</i> .....	16

2.1.10	<i>Naive Bayes</i> .....	17
2.2	Penelitian Terdahulu .....	19
2.2.1	Penelitian I .....	19
2.2.2	Penelitian II .....	20
2.2.3	Penelitian III .....	21
2.2.4	Penelitian IV .....	22
2.2.5	Penelitian V .....	23
BAB III METODOLOGI PENELITIAN .....		24
3.1	Studi Literatur .....	25
3.2	Analisis Kebutuhan Sistem .....	25
3.2.1	Kebutuhan Data .....	25
3.2.2	Kebutuhan <i>Software</i> dan <i>Hardware</i> .....	25
3.3	Pengumpulan Data .....	26
3.4	Pelabelan Data .....	27
3.5	Pembagian Data .....	30
3.6	Perancangan Model .....	31
3.6.1	<i>Preprocessing Data</i> .....	32
3.6.2	Pembagian Dataset untuk <i>Training</i> dan <i>Testing Model Naive Bayes</i> 34	
3.6.3	<i>TF-IDF</i> .....	35
3.6.4	Klasifikasi Emosi <i>Model Naive Bayes</i> .....	35
3.6.5	<i>Evaluasi Model</i> .....	35
3.6.6	Penyimpanan Model .....	35
3.7	Penerapan Model .....	36

3.8	<i>Validasi Model</i> .....	36
BAB IV HASIL DAN PEMBAHASAN .....		37
4.1	Analisa Data .....	37
4.1.1	Data .....	37
4.1.2	<i>Hardware &amp; Software</i> .....	37
4.2	Persiapan Data .....	37
4.2.1	Pengambilan Data .....	37
4.3	Pelabelan Data .....	40
4.4	Pembagian Data .....	43
4.5	Perancangan Model .....	43
4.5.1	<i>Preprocessing Data</i> .....	44
4.5.2	Pembagian Dataset Untuk <i>Training</i> dan <i>Testing Model Naive Bayes</i> 59	
4.5.3	<i>TF-IDF</i> .....	60
4.5.4	Klasifikasi Emosi Model Naive Bayes .....	60
4.5.5	Evaluasi Model .....	69
4.5.6	Penyimpanan Model .....	95
4.6	Penerapan Model .....	96
4.7	Validasi Model .....	99
BAB V KESIMPULAN DAN SARAN .....		101
5.1	Kesimpulan .....	101
5.2	Saran .....	102
DAFTAR PUSTAKA .....		103
LAMPIRAN .....		107

## DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian.....	24
Gambar 3. 2 Data Pencarian Vidio menggunakan Keyword “Ditutupnya Tiktok Shop” .....	26
Gambar 3. 3 Data Komentari Youtube.....	27
Gambar 3. 4 Perancangan Model .....	32
Gambar 4. 1 Sourcecode Crawling data Youtube .....	38
Gambar 4. 2 Total Dataset Tiktok Shop.....	42
Gambar 4. 3 Sourcecode pembagian data validasi dan training .....	43
Gambar 4. 4 Data Persebaran Tiap Emosi .....	44
Gambar 4. 5 Sourcecode Top 30 Kata yang paling sering muncul pada emosi sedih .....	45
Gambar 4. 6 Sourcecode Top 30 Kata yang paling sering muncul pada emosi Takut .....	45
Gambar 4. 7 Sourcecode Top 30 Kata yang paling sering muncul pada emosi Marah .....	46
Gambar 4. 8 Sourcecode Top 30 Kata yang paling sering muncul pada emosi Terkejut .....	46
Gambar 4. 9 Sourcecode Top 30 Kata yang paling sering muncul pada emosi Senang .....	47
Gambar 4. 10 Persebaran Jumlah Emosi pada jangka waktu 04 Oktober – 11 November 2023 .....	47
Gambar 4. 11 Sourcecode Jumlah duplicate dan grafik.....	48
Gambar 4. 12 Baris Yang Terdapat Duplicate .....	48

Gambar 4. 13 Sourcecode Cleaning Teks .....	49
Gambar 4. 14 Sourcecode case folding.....	50
Gambar 4. 15 Sourcecode Tokenization .....	51
Gambar 4. 16 Sourcecode Stopword Removal .....	53
Gambar 4. 17 Sourcecode lemmatization .....	54
Gambar 4. 18 Kata yang paling umum di kelas Marah.....	55
Gambar 4. 19 Kata Paling Umum Untuk Emosi Takut .....	55
Gambar 4. 20 Kata Paling Umum Untuk Emosi Sedih.....	56
Gambar 4. 21 Kata Paling Umum Untuk Emosi Senang.....	56
Gambar 4. 22 Kata Paling Umum Untuk Emosi Terkejut .....	56
Gambar 4. 23 Sourcecode imbalance data .....	57
Gambar 4. 24 Hasil Sebelum oversampling.....	58
Gambar 4. 25 Hasil sesudah oversampling .....	58
Gambar 4. 26 Sourcecode pembagian dataset untuk training dan test model naive bayes.....	59
Gambar 4. 27 Sourcecode TF-IDF.....	60
Gambar 4. 28 Sourcecode klasifikasi emosi model naive bayes .....	61
Gambar 4. 29 Accuracy comparison skenario 1 .....	62
Gambar 4. 30 Accuracy comparison skenario 2 .....	64
Gambar 4. 31 Accuracy comparison skenario 3 .....	65
Gambar 4. 32 Accuracy comparison skenario 4 .....	67
Gambar 4. 33 Accuracy comparison skenario 5 .....	68
Gambar 4. 34 Confusion Matrix dan Classication Report Skenario 1.....	70
Gambar 4. 35 Confusion Matrix dan Classication Report Skenario 2.....	75

Gambar 4. 36 Confusion Matrix dan Classification Report Skenario 3.....	80
Gambar 4. 37 Confusion Matrix dan Classification Report Skenario 4.....	85
Gambar 4. 38 Confusion Matrix dan Classification Report Skenario 5.....	90
Gambar 4. 39 Sourcecode Save vectorization.....	96
Gambar 4. 40 Sourcecode Save Model pickle .....	96
Gambar 4. 41 Sourcecode Penerapan Model .....	97
Gambar 4. 42 Tampilan Pengembangan Model.....	97
Gambar 4. 43 Hasil Prediksi Teks.....	98
Gambar 4. 44 Hasil Prediksi Dataset .....	98
Gambar 4. 45 Hasil Prediksi Dataset (1).....	99
Gambar 4. 46 Hasil Prediksi Dataset (2).....	99

## DAFTAR TABEL

Tabel 2. 1 Confusion Matrix .....	16
Tabel 2. 2 Penelitian Terdahulu I.....	19
Tabel 2. 3 Penelitian Terdahulu II.....	20
Tabel 2. 4 Penelitian Terdahulu III .....	21
Tabel 2. 5 Penelitian Terdahulu IV .....	22
Tabel 2. 6 Penelitian Terdahulu V .....	23
Tabel 3. 1 Data Yang sudah diberi Label.....	27
Tabel 3. 2 Jumlah Data Yang Sudah Diberi Label.....	30
Tabel 3. 3 Jumlah Data Model .....	30
Tabel 3. 4 Jumlah Data Validasi .....	31
Tabel 4. 1 Hasil Crawling Data Komentar Youtube .....	39
Tabel 4. 2 Data Yang Sudah diLabeling .....	41
Tabel 4. 3 Hasil Pelabelan yang digunakan Klasifikasi.....	42
Tabel 4. 4 Hasil cleaning teks .....	49
Tabel 4. 5 Hasil case folding.....	50
Tabel 4. 6 Hasil Tokenization .....	52
Tabel 4. 7 Hasil Stopword Removal .....	53
Tabel 4. 8 Hasil Lemmatization .....	54
Tabel 4. 9 Rangkuman Skenario .....	94