

**ANALISIS SENTIMEN TERHADAP ULASAN MIE
GACOAN JUANDA BEKASI PADA *GOOGLE REVIEW*
MENGUNAKAN METODE *NAÏVE BAYES* DAN C4.5**

SKRIPSI

Oleh:

Nur Fadila

201910225104



**PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS BHAYANGKARA JAKARTA RAYA
2023**

LEMBAR PERSETUJUAN PEMBIMBING
(DIGUNAKAN UNTUK TUGAS AKHIR)

Judul Tugas Akhir : Analisis Sentimen terhadap Ulasan Mie Gacoan
Juanda Bekasi pada *Google Review* Menggunakan
Metode *Naive Bayes* dan *C4.5*

Nama Mahasiswa : NUR FADILA
Nomor Pokok Mahasiswa : 201910225104
Program Studi/Fakultas : Informatika/Ilmu Komputer
Tanggal Lulus Ujian Skripsi : 04 Agustus 2023

Pembimbing I



Wowon Priatna, S.T., M.Ti.
NIDN: 0429118007

Pembimbing II



Siti Setiawati, S.Pd., M.Pd.
NIDN: 0313107904

LEMBAR PENGESAHAN

Judul Skripsi : ANALISIS SENTIMEN TERHADAP REVIEW
MIE GACOAN JUANDA BEKASI
MENGUNAKAN METODE NAÏVE BAYES
DAN C4.5

Nama Mahasiswa : NUR FADILA

Nomor Induk Mahasiswa : 201910225104

Program Studi/Fakultas : Informatika/Ilmu Komputer

Tanggal Lulus Ujian Skripsi : 04 Agustus 2023

Jakarta, 08 Agustus 2023

MENGESAHKAN

Ketua Tim Penguji : Dr. Dra. Tyastuti Sri Lestari, MM.
NIDN: 0327036701

Penguji II : Fata Nidaul Khasanah, S.Kom., M.Eng
NIDN: 0328068603

Penguji III : Wowon Priatna, S.T., M.Ti.
NIDN: 0429118007

MENGETAHUI,

Ketua
Program Studi Informatika



Ahmad Fathurrozi, S.E., M.M.S.I
NIP. 2012486

Dekan
Fakultas Ilmu Komputer



Dr. Dra. Tyastuti Sri Lestari, M.M.
NIP. 1408206



LEMBAR PERNYATAAN BUKAN PLAGIASI

Yang bertanda tangan dibawah ini :

Nama : Nur Fadila
NPM : 201910225104
Program Studi : Informatika
Fakultas : Ilmu Komputer
Judul Tugas Akhir : Analisis Sentimen Terhadap Ulasan Mie Gacoan Juanda
Bekasi Pada *Google Review* Menggunakan Metode
Naïve Bayes dan C4.5

Dengan ini menyatakan bahwa hasil penulisan skripsi yang telah saya buat ini merupakan **hasil karya saya sendiri dan benar keasliannya**. Apabila dikemudian hari penulisan skripsi ini merupakan plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggungjawabkan sekaligus bersedia menerima sanksi berdasarkan tata tertib di Universitas Bhayangkara Jakarta Raya.

Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan dari pihak manapun.

Jakarta, 10 Agustus 2023

Penulis



Nur Fadila

ABSTRACT

Nur Fadila. 201910225104. *Sentiment Analysis of Mie Gacoan Juanda Bekasi Reviews on Google Review Using Naïve Bayes and C4.5 Methods.* Faculty of Computer Science. Bekasi: Bhayangkara Jakarta Raya University. 2023.

The development of the culinary business is growing rapidly which has resulted in many culinary businessmen opening restaurants with attractive menus. One of the restaurants that is in great demand by the public is Mie Gacoan. Mie Gacoan has opened many branches, one of which is located in Bekasi, especially on Jalan Juanda. The platform that is widely used by business people as a means of receiving feedback is Google Review. Google Review is one of them, the best platform to collect reviews from customers regarding products or services. The reviews on Google Reviews serve as a reference for visitors and also serve as evaluation material for culinary business owners to make improvements to increase visitor comfort and satisfaction, both in terms of taste and facilities. In this study, researchers used the Naïve Bayes and C4.5 methods to solve these problems by focusing on the performance of the Naïve Bayes method in classifying sentiment analysis. Data is collected by fetching reviews from Google Reviews using the Instant Data Scraper tool on the Google Chrome platform. A total of 1,291 reviews were retrieved, which includes both positive and negative reviews. The test results show that the system succeeded in classifying sentiment using the Naïve Bayes and C4.5 methods, with the highest accuracy rate of 83% in the Naïve Bayes classification, while the accuracy of C4.5 classification is 74%. With the results of the naïve Bayes algorithm 83% and C4.5 74%, the algorithm is able to classify data properly.

Keywords: Sentiment Analysis, Mie Gacoan, Google Review, Naïve Bayes, and C4.5.

ABSTRAK

Nur Fadila. 201910225104. Analisis Sentimen terhadap Ulasan Mie Gacoan Juanda Bekasi pada *Google Review* Menggunakan Metode *Naïve Bayes* dan C4.5. Fakultas Ilmu Komputer. Bekasi: Universitas Bhayangkara Jakarta Raya. 2023

Perkembangan bisnis kuliner semakin berkembang pesat yang mengakibatkan banyak para pembisnis kuliner yang membuka restoran dengan menu yang menarik. Salah satu restoran yang banyak diminati oleh masyarakat adalah Mie Gacoan. Mie Gacoan telah banyak membuka cabang, salah satunya yang berada di Bekasi khususnya di jalan Juanda. *Platform* yang banyak dipakai oleh para pembisnis sebagai sarana penerima *feedback* yaitu *Google Review*. *Google Review* merupakan salah satu platform terbaik untuk mengumpulkan ulasan dari pelanggan mengenai produk atau layanan. Ulasan yang terdapat di *Google Review* menjadi referensi bagi pengunjung dan juga menjadi bahan evaluasi bagi pemilik usaha kuliner untuk melakukan perbaikan guna meningkatkan kenyamanan dan kepuasan pengunjung, baik dari segi rasa maupun fasilitasnya. Dalam penelitian ini, peneliti menggunakan metode *Naïve Bayes* dan C4.5 untuk menyelesaikan permasalahan tersebut dengan fokus pada kinerja metode *Naïve Bayes* dalam mengklasifikasikan analisis sentimen. Data dikumpulkan dengan mengambil ulasan dari *Google Review* menggunakan alat bantu *Instant Data Scraper* di *platform Google Chrome*. Total terdapat 1.291 ulasan yang diambil, yang mencakup ulasan positif dan negatif. Hasil pengujian menunjukkan bahwa sistem berhasil melakukan klasifikasi sentimen menggunakan metode *Naïve Bayes* dan C4.5, dengan tingkat akurasi tertinggi sebesar 83% pada klasifikasi *Naïve Bayes*, sedangkan akurasi klasifikasi C4.5 sebesar 74%. Dengan hasil algoritma *naïve bayes* 83% dan C4.5 74% maka algoritma mampu mengklasifikasikan data dengan baik.

Kata kunci: Analisis Sentimen, Mie Gacoan, *Google Review*, *Naïve Bayes*, dan C.45

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIK**

Sebagai sivitas akademik Universitas Bhayangkara Jakarta Raya, saya yang bertanda tangan di bawah ini :

Nama : Nur Fadila
NPM : 201910225104
Program Studi : Informatika
Fakultas : Ilmu Komputer
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Bhayangkara Jakarta Raya **Hak Bebas Royalti Non-Eksklusif (*Non-Exclusive Royalty-Free Right*)**, atas karya ilmiah saya yang berjudul :

Analisis Sentimen Terhadap Ulasan Mie Gacoan Juanda Bekasi Pada *Google Review* Menggunakan Metode *Naïve Bayes* dan *C4.5*

beserta perangkat yang ada (bila diperlukan). Dengan hak bebas royalti non-eksklusif ini, Universitas Bhayangkara Jakarta Raya berhak menyimpan, mengalihmediakan, mengelolanya dalam bentuk pangkalan data (*database*), mendistribusikannya dan mempublikasikannya di Internet atau media lain untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis dan sebagai pemilik hak cipta.

Segala bentuk tuntutan hukum yang timbul atas pelanggaran hak cipta dalam karya ilmiah ini menjadi tanggung jawab saya pribadi

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di : Jakarta
Pada tanggal : 10 Agustus 2023
Yang Menyatakan



Nur Fadila

KATA PENGANTAR

Puji syukur peneliti panjatkan kepada Tuhan Yang Maha Esa Allah SWT yang telah melimpahkan karunia-Nya sehingga peneliti dapat menyelesaikan skripsi yang disusun sebagai salah satu syarat untuk mencapai gelar sarjana Strata satu (S1) pada program studi Informatika Universitas Bhayangkara Jakarta Raya.

Penelitian ini tidaklah dapat terwujud dengan baik tanpa adanya bimbingan, dorongan, saran-saran dan bantuan dari berbagai pihak. Pada kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Allah S.W.T yang telah meridhai penulis selaku hamba-Nya untuk menulis proposal studi independen ini hingga selesai.
2. Bapak Inspektur Jenderal Polisi Irjen Pol. (Purn) Dr, Drs. H. Bambang Karsono, S.H.,M.M. selaku Rektor Universitas Bhayangkara Jakarta Raya.
3. Ibu Dr. Tyastuti Sri Lestari, S.Si., MM selaku Dekan Fakultas Ilmu Komputer Universitas Bhayangkara Jakarta Raya.
4. Bapak Ahmad Fathurrozi, S.E., M.M.S.I. selaku Ketua Program Studi Informatika.
5. Bapak Wowon Priatna, S.T., M.TI selaku dosen pembimbing satu penyusunan laporan skripsi.
6. Ibu Siti Setiawati, S.Pd., M.Pd. selaku dosen pembimbing dua penyusunan laporan skripsi.

7. Ibu Hafizah, S.S., M.Pd. selaku validator instrumen dalam penentuan palabelan data dalam pembuatan laporan skripsi.
8. Kedua orang tua saya yang selalu memberikan dukungan dan semangat dan selalu medoakan saya setiap harinya dalam agar diberikan kemudahan dalam menyelesaikan skripsi ini.
9. Teman-teman ANGK yang selalu memberikan dukungan dan semangat dalam menyelesaikan skripsi ini.
10. Teman-teman seperjuangan pada program studi Informatika Universitas Bhayangkara Jakarta Raya yang selalu memberikan dukungan dan selalu membantu dalam menyelesaikan skripsi ini.
11. Eka Nur A'ini yang telah bersama-sama membantu dalam proses penyelesaian skripsi ini.
12. Dhea Putri Aprilyana yang telah bersama-sama membantu dalam proses penyelesaian skripsi ini.

Peneliti menyadari bahwa penelitian ini jauh dari kesempurnaan, untuk itu peneliti mohon maaf apabila terdapat kekurangan dan kesalahan baik dari isi maupun tulisan. Oleh karena itu, peneliti mengharapkan kritik dan saran yang membangun

Jakarta,.....2023

Nur Fadila

DAFTAR ISI

LEMBAR PERSETUJUAN PEMBIMBING	i
LEMBAR PERNYATAAN BUKAN PLAGIASI	iii
ABSTRAK	iv
ABSTRACT	v
KATA PENGANTAR.....	vii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN.....	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Identifikasi Masalah.....	6
1.3 Batasan Masalah	6
1.4 Rumusan Masalah.....	6
1.5 Tujuan dan Manfaat Penelitian.....	7
1.5.1 Tujuan Penelitian	7
1.5.2 Manfaat Penelitian	7
1.6 Sistematika Penulisan	8
BAB II TINJAUAN PUSTAKA.....	10
2.1 <i>Table Relate of State</i>	10
2.2 Penelitian Terdahulu	12
2.3 <i>Data Science</i>	15
2.4 <i>Data Mining</i>	15
2.4.1 <i>Text Mining</i>	16
2.4.2 <i>KDD (Knowledge Discovery Database)</i>	19
2.4.3 <i>Confusion Matrix</i>	20
2.5 Sentimen	23
2.6 Analisis Sentimen/ <i>Sentiment Analysis</i>	23

2.7	<i>Review</i>	24
2.8	Mie Gacoan.....	25
2.9	<i>Google Review</i>	25
2.10	<i>Naïve Bayes</i>	26
2.11	Algoritma C4.5	28
2.12	Visualisasi.....	30
BAB III METODOLOGI PENELITIAN		31
3.1	Kerangka Penelitian.....	31
3.2	Pengumpulan Data.....	33
3.3	Pelabelan.....	33
3.4	<i>Pre-Processing</i>	34
3.4.1	<i>Case Folding</i>	35
3.4.2	<i>Tokenizing</i>	36
3.4.3	<i>Filtering</i>	38
3.4.4	<i>Stemming</i>	40
3.5	TF-IDF (<i>Term Frequency-Inverse Document Frequency</i>).....	42
3.6	<i>Confusion Matrix</i>	42
3.7	Algoritma <i>Naïve Bayes</i>	43
3.8	Algoritma C4.5	44
BAB IV HASIL DAN PEMBAHASAN		45
4.1	Hasil.....	45
4.1.1	<i>Data Selection</i>	45
4.1.2	<i>Pre-Processing</i>	51
4.1.3	<i>Transformation</i>	58
4.1.4	<i>Data Mining</i>	68
4.1.5	<i>Evaluation</i>	86
4.1.6	<i>Knowledge Presentation</i>	91
4.2	Pembahasan.....	93
BAB V PENUTUP.....		94
5.1	Kesimpulan	94
5.2	Saran	95
DAFTAR PUSTAKA.....		96



DAFTAR TABEL

Table 2. 1 Table Relate of State	10
Table 2. 2 Table Confusion Matrix	21
Table 3. 1 Contoh Tahap Case Folding.....	36
Table 3. 2 Contoh Tahap Tokenizing.....	38
Table 3. 3 Contoh Tahap Filtering	40
Table 3. 4 Contoh Tahap Stemming	42
Table 4. 1 Tabel Contoh Dokumen	59
Table 4. 2 Tabel Perhitungan TF-IDF	61
Table 4. 3 Tabel Perhitungan Bobot Kata	64
Table 4. 4 Tabel Data Training Naive Bayes.....	68
Table 4. 5 Tabel Prediksi Data Training Naive Bayes.....	69
Table 4. 6 <i>Prior Probability</i>	71
Table 4. 7 Data Uji Naive Bayes.....	72
Table 4. 8 Contoh <i>Data Training C4.5</i>	75
Table 4. 9 Nilai Data Training	75
Table 4. 10 Perhitungan Gain untuk Simpul Akar.....	78
Table 4. 11 Nilai Data Training	79
Table 4. 12 Perhitungan Gain untuk Simpul Akar.....	79
Table 4. 13 Nilai Data Training	80
Table 4. 14 Perhitungan Gain Simpul Gacoan.....	81
Table 4. 15 Tabel Confusion Matrix Naive Bayes.....	87
Table 4. 16 Tabel Confision Matrix C4.5	90
Table 4. 17 Hasil Confusion Matrix.....	93

DAFTAR GAMBAR

Gambar 1. 1 Mie Gacoan Juanda Bekasi	2
Gambar 2. 1 KDD (<i>Knowledge Discovery Database</i>)	19
Gambar 3. 1 Diagram Alir Kerangka Penelitian KDD	31
Gambar 3. 2 Diagram Alir Pelabelan Data	34
Gambar 3. 3 Diagram Alir <i>Case Folding</i>	35
Gambar 3. 4 Diagram Alir <i>Tokenizing</i>	37
Gambar 3. 5 Diagram Alir <i>Filtering</i>	39
Gambar 3. 6 Diagram Alir <i>Stemming</i>	41
Gambar 3. 7 Diagram Alir Algoritma <i>Naive Bayes</i>	43
Gambar 3. 8 Diagram Alir Algoritma C4.5	44
Gambar 4. 1 <i>Platform Instant Data Scraper</i>	46
Gambar 4. 2 Ulasan Pelanggan Mie Gacoan Juanda Bekasi	47
Gambar 4. 3 <i>Scrapping Data</i>	48
Gambar 4. 4 Pengambilan Data	48
Gambar 4. 5 Menyimpan Hasil <i>Scrapping Data</i>	49
Gambar 4. 6 Hasil <i>Scrapping Data</i>	49
Gambar 4. 7 Pelabelan Manual	50
Gambar 4. 8 <i>Script Import Data</i>	52
Gambar 4. 9 Dataset	52
Gambar 4. 10 <i>Script</i> Melihat Informasi Data	52
Gambar 4. 11 Informasi Data	53
Gambar 4. 12 <i>Script</i> Proses <i>Case Folding</i>	53
Gambar 4. 13 Hasil Proses <i>Case Folding</i>	54
Gambar 4. 14 <i>Script</i> Proses <i>Tokenizing</i>	54
Gambar 4. 15 Hasil Proses <i>Tokenizing</i>	55
Gambar 4. 16 <i>Script</i> Proses <i>Filtering</i>	56
Gambar 4. 17 Hasil Proses <i>Filtering</i>	56
Gambar 4. 18 <i>Script</i> Proses <i>Stemming</i>	57
Gambar 4. 19 Hasil Proses <i>Stemming</i>	57
Gambar 4. 20 Tahapan <i>Splitting Data</i> dan Hasil <i>Splitting Data</i>	58
Gambar 4. 21 <i>Script</i> TF-IDF	67
Gambar 4. 22 Hasil TF-IDF	67
Gambar 4. 23 <i>Script</i> Model Klasifikasi <i>Naive Bayes</i>	74
Gambar 4. 24 Hasil Perhitungan Nilai Gain Simpul Akar	78
Gambar 4. 25 Hasil Perhitungan Nilai Gain Simpul Gacoan	80
Gambar 4. 26 Hasil Perhitungan Gain Simpul Gacoan	81
Gambar 4. 27 <i>Script</i> C4.5	82

Gambar 4. 28 <i>Script</i> Menampilkan Pohon Keputusan.....	83
Gambar 4. 29 Hasil Awal Pohon Keputusan	84
Gambar 4. 30 Hasil Akhir Pohon Keputusan.....	85
Gambar 4. 31 Tabel <i>Confusion Matrix Naive Bayes</i>	86
Gambar 4. 32 Hasil <i>Confusion Matrix Naive Bayes</i>	87
Gambar 4. 33 Tabel <i>Confusion Matrix C4.5</i>	89
Gambar 4. 34 Hasil <i>Confusion Matrix C4.5</i>	89
Gambar 4. 35 <i>Wordcloud</i>	92



DAFTAR GAMBAR

Gambar 1. 1 Mie Gacoan Juanda Bekasi	2
Gambar 2. 1 KDD (<i>Knowledge Discovery Database</i>)	19
Gambar 3. 1 Diagram Alir Kerangka Penelitian KDD	31
Gambar 3. 2 Diagram Alir Pelabelan Data	34
Gambar 3. 3 Diagram Alir <i>Case Folding</i>	35
Gambar 3. 4 Diagram Alir <i>Tokenizing</i>	37
Gambar 3. 5 Diagram Alir <i>Filtering</i>	39
Gambar 3. 6 Diagram Alir <i>Stemming</i>	41
Gambar 3. 7 Diagram Alir Algoritma <i>Naive Bayes</i>	43
Gambar 3. 8 Diagram Alir Algoritma C4.5	44
Gambar 4. 1 <i>Platform Instant Data Scraper</i>	46
Gambar 4. 2 Ulasan Pelanggan Mie Gacoan Juanda Bekasi	47
Gambar 4. 3 <i>Scrapping Data</i>	48
Gambar 4. 4 Pengambilan Data	48
Gambar 4. 5 Menyimpan Hasil <i>Scraping Data</i>	49
Gambar 4. 6 Hasil <i>Scraping Data</i>	49
Gambar 4. 7 Pelabelan Manual	50
Gambar 4. 8 <i>Script Import Data</i>	52
Gambar 4. 9 Dataset	52
Gambar 4. 10 <i>Script</i> Melihat Informasi Data	52
Gambar 4. 11 Informasi Data	53
Gambar 4. 12 <i>Script</i> Proses <i>Case Folding</i>	53
Gambar 4. 13 Hasil Proses <i>Case Folding</i>	54
Gambar 4. 14 <i>Script</i> Proses <i>Tokenizing</i>	54
Gambar 4. 15 Hasil Proses <i>Tokenizing</i>	55
Gambar 4. 16 <i>Script</i> Proses <i>Filtering</i>	56
Gambar 4. 17 Hasil Proses <i>Filtering</i>	56
Gambar 4. 18 <i>Script</i> Proses <i>Stemming</i>	57
Gambar 4. 19 Hasil Proses <i>Stemming</i>	57
Gambar 4. 20 Tahapan <i>Splitting Data</i> dan Hasil <i>Splitting Data</i>	58
Gambar 4. 21 <i>Script</i> TF-IDF	67
Gambar 4. 22 Hasil TF-IDF	67
Gambar 4. 23 <i>Script</i> Model Klasifikasi <i>Naive Bayes</i>	74
Gambar 4. 24 Hasil Perhitungan Nilai Gain Simpul Akar	78
Gambar 4. 25 Hasil Perhitungan Nilai Gain Simpul Gacoan	80
Gambar 4. 26 Hasil Perhitungan Gain Simpul Gacoan	81
Gambar 4. 27 <i>Script</i> C4.5	82

Gambar 4. 28 <i>Script</i> Menampilkan Pohon Keputusan.....	83
Gambar 4. 29 Hasil Awal Pohon Keputusan	84
Gambar 4. 30 Hasil Akhir Pohon Keputusan.....	85
Gambar 4. 31 Tabel <i>Confusion Matrix Naive Bayes</i>	86
Gambar 4. 32 Hasil <i>Confusion Matrix Naive Bayes</i>	87
Gambar 4. 33 Tabel <i>Confusion Matrix C4.5</i>	89
Gambar 4. 34 Hasil <i>Confusion Matrix C4.5</i>	89
Gambar 4. 35 <i>Wordcloud</i>	92



DAFTAR LAMPIRAN

Lampiran 1 Uji Plagiarisme	101
Lampiran 2 Biodata Peneliti	102
Lampiran 3 Kartu Bimbingan Tugas Akhir Dosen Pembimbing I	104
Lampiran 4 Kartu Bimbingan Tugas Akhir Dosen Pembimbing II	105
Lampiran 5 Surat Rekomendasi Dosen Pembimbing untuk Tugas Akhir	106
Lampiran 6 Surat Pernyataan Validasi Instrumen	107
Lampiran 7 Lembar Validasi Instrumen	108

