

**PERBANDINGAN METODE *NAÏVE BAYES* DAN *K-NEAREST
NEIGHBOR* DALAM ANALISIS SENTIMEN TERHADAP
PROGRAM MAKAN BERGIZI GRATIS**

TUGAS AKHIR

NO. 1279/WM.FT.H6/T.ILKOM/TA/2025

**Diajukan Sebagai Salah Satu Syarat Untuk Memperoleh
Gelar Sarjana Komputer**



Disusun oleh :

CHRISTY INDA WATTIMENA

23122072

PROGRAM STUDI ILMU KOMPUTER

FAKULTAS TEKNIK

UNIVERSITAS KATOLIK WIDYA MANDIRA

KUPANG

2026

HALAMAN PERSETUJUAN

TUGAS AKHIR

NO. 1279/WM.FT.H6/T.ILKOM/TA/2025

PERBANDINGAN METODE *NAÏVE BAYES* DAN *K-NEAREST NEIGHBOR* DALAM ANALISIS SENTIMEN TERHADAP
PROGRAM MAKAN BERGIZI GRATIS

OLEH:

CHRISTY INDA WATTIMENA

23122072

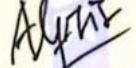
TELAH DIPERIKSA/ DISETUJUI OLEH DOSEN

DI : KUPANG
PADA TANGGAL : JANUARI 2026

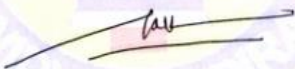
DOSEN PENGUJI I


Sisilia Daeng B. Mau, S.Kom., M.T.
NIDN : 0807098502

DOSEN PENGUJI II


Alfry Aristo J. Sinla, S.Kom., M.Cs
NIDN : 0807078704

DOSEN PENGUJI III


Donatus J. Manehat, S.Si, M.Kom
NIDN : 0828126601

KETUA PELAKSANA


Donatus J. Manehat, S.Si, M.Kom
NIDN : 0828126601

SEKRETARIS PELAKSANA


Paul F. M. Tengangatu, S.Kom., M.T.I
NUPTK : 4955769670130332

HALAMAN PENGESAHAN

TUGAS AKHIR

NO. 1279/WM.FT.H6/T.ILKOM/TA/2025

PERBANDINGAN METODE *NAÏVE BAYES* DAN *K-NEAREST
NEIGHBOR* DALAM ANALISIS SENTIMEN TERHADAP
PROGRAM MAKAN BERGIZI GRATIS

Oleh:

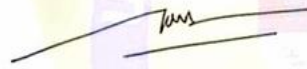
CHRISTY INDA WATTIMENA

23122072

TELAH DIPERTAHANKAN DIDEPAN PEMBIMBING

DOSEN PEMBIMBING I

DOSEN PEMBIMBING II



Donatus J. Manehat, S.Si, M.Kom

Paul F. M. Tengangatu, S.Kom., M.T.I

NIDN : 0828126601

NUPTK : 4955769670130332

MENGETAHUI

MENGETAHUI,

KETUA PROGRAM STUDI

DEKAN FAKULTAS TEKNIK

ILMU KOMPUTER

UNIKA WIDYA MANDIRA

UNIKA WIDYA MANDIRA

UNIVERSITAS KATOLIK



Yovina C. Hoar Siki, ST., MT



Dr. Don Gaspar N. da Costa, S.T., M.T

NIDN : 0805058803

NIDN : 0820036801

MOTTO

“Belum ada alasan untuk berhenti, besok mungkin kita sampai”

(Baskara Putra - Hindia)

“Hidup bukan saling mendahului bermimpilah sendiri-sendiri”

(Baskara Putra - Hindia)

PERNYATAAN DAN KEASLIAN HASIL KARYA

Yang bertanda tangan di bawah ini:

Nama : Christy Inda Wattimena
Nim : 231 22 072
Fakultas : Teknik
Program studi : Ilmu Komputer

Dengan ini menyatakan bahwa Tugas Akhir dengan judul “**Perbandingan Metode Naïve Bayes Dan K-Nearest Neighbor Dalam Analisis Sentimen Terhadap Program Makan Bergizi Gratis**” adalah benar-benar karya saya sendiri. Apabila dikemudian hari ditemukan penyimpangan, maka saya bersedia dituntut secara hukum.

Disahkan/Diketahui

Pembimbing I



Donatus J. Manchat, S.Si, M.Kom

Kupang, Januari 2026

Mahasiswa



Christy Inda Wattimena

HALAMAN PERSEMBAHAN

Karya ini secara khusus saya persembahkan untuk:

Bapak, Mama, Kakak, Adik dan seluruh keluarga yang selalu mendoakan yang terbaik untuk penulis serta semua teman, sahabat yang selalu membantu, dan teman-teman seperjuangan Mata Merah 22.

Almamater UNWIRA tercinta

KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas berkat dan penyertaan-Nya sehingga penulis dapat menyelesaikan tugas akhir dengan judul: “Perbandingan Metode *Naïve Bayes* dan *K-Nearest Neighbor* Dalam Analisis Sentimen Terhadap Program Makan Bergizi Gratis”. Proposal ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Ilmu Komputer, Universitas Katolik Widya Mandira Kupang.

Penulis menyampaikan terima kasih kepada:

1. Tuhan Yesus Kristus, terima kasih atas kasih setia dan anugerah-Mu yang tidak pernah berkesudahan. Dalam setiap proses penyusunan skripsi ini, dalam perjalanan studi yang penuh dinamika, serta dalam setiap langkah kehidupan penulis, Engkau selalu menjadi sumber kekuatan, pengharapan, dan penghiburan. Di saat menghadapi tantangan, keraguan, dan kelelahan, Engkau tetap setia menyertai dan menuntun penulis hingga mampu menyelesaikan skripsi ini dengan baik. Segala sesuatu yang telah tercapai sampai saat ini semata-mata terjadi karena penyertaan, hikmat, dan kebaikan-Mu.
2. Pater Dr. Stefanus Lio, SVD., S.Fil., MA., selaku rektor Universitas Katolik Widya Mandira Kupang.
3. Bapak Dr. Don Gaspar N. Da Costa, ST., MT. selaku Dekan Fakultas Teknik Universitas Katolik Widya Mandira Kupang.
4. Ibu Yovina Carmeneja Hoar Siki, S.T., M.T. selaku Ketua Program Studi Ilmu Komputer Universitas Katolik Widya Mandira Kupang.
5. Bapak Donatus Joseph Manehat, S.Si., M.Kom. selaku Dosen Pembimbing I serta Bapak Paul Filson Mite Tengangatu, S.Kom., M.T.I. selaku Dosen Pembimbing II, yang dengan penuh kesabaran membimbing penulis, meluangkan waktu untuk berdiskusi, serta memberikan arahan dan masukan yang sangat berarti sampai skripsi ini terselesaikan.
6. Ibu Sisilia Daeng B. Mau, S.Kom., M.T. selaku dosen penguji I dan Bapak Alfry Aristo J. SinlaE, S.Kom., M.Cs selaku dosen penguji II, yang telah

memberikan tanggapan, kritik membangun, dan saran perbaikan yang membantu penulis menyempurnakan skripsi ini.

7. Segenap Bapak dan Ibu Dosen Program Studi Ilmu Komputer serta Staf Tata Usaha yang telah memberikan bekal ilmu pengetahuan dan bantuan administratif selama masa perkuliahan.
8. Bapak Juventinho J. M. de Araujo, S.Kom., selaku Laboran yang telah memberikan dukungan moral dan bantuan kepada penulis selama proses penyusunan skripsi ini.
9. Penghormatan dan terima kasih yang paling utama penulis persembahkan kepada kedua pahlawan hidup penulis, Bapak Godlief Rudolf Wattimena dan Mama Nelcie Kapioru. Terima kasih atas cinta yang tak bertepi, doa-doa tulus yang senantiasa dilangitkan demi keberhasilan penulis, serta segala pengorbanan dan dukungan penuh, baik secara moral maupun materi, hingga penulis dapat menyelesaikan pendidikan ini. Serta kepada adik-adik tersayang, Demico Indel Wattimena dan Helena Indira Wattimena, terima kasih atas semangat, keceriaan, dan kasih sayang yang menjadi kekuatan bagi penulis.
10. Sahabat-sahabat terdekat penulis selama masa perkuliahan: Cory Selan, Jessica Setiawati, Prilly Seran, dan Elga Tabala, terima kasih telah menjadi teman seperjuangan yang luar biasa dalam setiap suka dan duka. Kebersamaan kalian adalah kekuatan bagi penulis dalam melewati setiap tantangan akademik.
11. Rekan-rekan mahasiswa angkatan Mata Merah, khususnya teman-teman Kelas A Program Studi Ilmu Komputer Universitas Katolik Widya Mandira Kupang. Terima kasih atas kebersamaan, rasa solidaritas, dan perjuangan bersama yang saling menguatkan selama kita menimba ilmu.
12. Seluruh pihak yang telah memberikan sumbangan dalam masa perkuliahan dan penyelesaian Tugas Akhir ini yang tidak dapat disebutkan namanya satu persatu, semoga Tuhan Yang Maha Esa membalas budi baik saudara-saudari sekalian. Tiada yang penulis berikan, selain ucapan terima kasih dan doa tulus, semoga segala bantuan dan dukungan yang diberikan mendapat balasan berkat yang setimpal dari Tuhan.

Penulis menyadari sepenuhnya bahwa penyusunan proposal tugas akhir ini masih jauh dari kata sempurna, baik dari segi materi, metodologi, maupun teknik penulisannya. Hal ini dikarenakan keterbatasan pengetahuan dan pengalaman yang penulis miliki. Oleh karena itu, kritik serta saran yang membangun dari Bapak/Ibu dosen penguji, pembimbing, dan para pembaca sangat penulis harapkan sebagai bahan evaluasi dan perbaikan di masa mendatang.

Penulis berharap hasil penelitian mengenai perbandingan metode *Naïve Bayes* dan *K-Nearest Neighbor* pada program Makan Bergizi Gratis ini tidak hanya sekadar memenuhi syarat kelulusan, namun juga dapat memberikan kontribusi nyata dalam pengembangan keilmuan di bidang *Natural Language Processing* (NLP) dan analisis sentimen. Semoga proposal ini dapat menjadi referensi yang bermanfaat bagi penelitian-penelitian selanjutnya, khususnya dalam upaya mengukur persepsi publik terhadap kebijakan pemerintah melalui media sosial.

Kupang, Januari 2026

Penulis

ABSTRAK

Program Makan Bergizi Gratis (MBG) merupakan kebijakan sosial pemerintah yang memunculkan beragam respons masyarakat di media sosial. Penelitian ini bertujuan untuk menganalisis sentimen publik terhadap Program MBG serta membandingkan kinerja algoritma *Naïve Bayes* (NB) dan *K-Nearest Neighbor* (KNN) dalam klasifikasi sentimen berbasis teks. Data penelitian dikumpulkan dari platform X menggunakan kata kunci terkait MBG pada periode 1 Agustus 2025 sampai 30 September 2025, dengan total 15.854 *tweet* berbahasa Indonesia. Setelah melalui proses *cleaning* dan *preprocessing*, data valid yang digunakan untuk pemodelan berjumlah 14.169 *tweet*. Tahapan *preprocessing* meliputi *case folding*, *remove punctuation*, *stopword removal* (dengan normalisasi), *tokenizing*, dan *stemming*. Selanjutnya, fitur teks dibentuk menggunakan pembobotan TF-IDF dan data dibagi menjadi data latih dan data uji dengan rasio 70:30 menggunakan teknik *Stratify*. Klasifikasi dilakukan ke dalam tiga kelas sentimen, yaitu negatif, netral, dan positif. Hasil analisis menunjukkan distribusi sentimen didominasi oleh sentimen negatif sebesar 55,09% (7.805 data), diikuti sentimen netral 23,25% (3.294 data) dan sentimen positif 21,67% (3.070 data). Evaluasi kinerja model menggunakan metrik *accuracy*, *precision*, *recall*, dan *F1-score* menunjukkan bahwa *Naïve Bayes* memperoleh akurasi 72%, sedangkan KNN memperoleh akurasi 71%. Dengan mempertimbangkan akurasi keseluruhan dan karakteristik performa per kelas, *Naïve Bayes* memberikan kinerja yang lebih baik pada dataset penelitian ini.

Kata Kunci: Analisis Sentimen, *Naïve Bayes*, *K-Nearest Neighbor*, Program Makan Bergizi Gratis, Media Sosial X, *Text Mining*, *TF-IDF*

ABSTRACT

The Free Nutritious Meal (MBG) Program is a government social policy that has elicited various public responses on social media. This study aims to analyze public sentiment toward the MBG Program and compare the performance of Naïve Bayes (NB) and K-Nearest Neighbor (KNN) algorithms in text-based sentiment classification. Research data was collected from the X platform using keywords related to MBG during the period of August 1, 2025, to September 30, 2025, with a total of 15,854 Indonesian-language tweets. After cleaning and preprocessing, 14,169 valid tweets were used for modeling. Preprocessing stages included case folding, punctuation removal, stopword removal (with normalization), tokenizing, and stemming. Furthermore, text features were formed using TF-IDF weighting, and the data was split into training and testing sets with a 70:30 ratio using the Stratify technique. Classification was conducted into three sentiment classes: negative, neutral, and positive. The results showed that the sentiment distribution was dominated by negative sentiment at 55.09% (7,805 data), followed by neutral sentiment at 23.25% (3,294 data) and positive sentiment at 21.67% (3,070 data). Performance evaluation using accuracy, precision, recall, and F1-score metrics showed that Naïve Bayes achieved an accuracy of 72%, while KNN achieved an accuracy of 71%. Considering the overall accuracy and performance characteristics per class, Naïve Bayes provided better performance on this research dataset.

Keywords: Sentiment Analysis, Naïve Bayes, K-Nearest Neighbor, Free Nutritious Meals Program, X Social Media, Text Mining, TF-IDF

DAFTAR ISI

HALAMAN PERSETUJUAN	ii
HALAMAN PENGESAHAN.....	iii
MOTTO	iv
PERNYATAAN DAN KEASLIAN HASIL KARYA.....	v
HALAMAN PERSEMBAHAN	vi
KATA PENGANTAR.....	vii
ABSTRAK	x
ABSTRACT	xi
DAFTAR ISI.....	xii
DAFTAR TABEL	xvi
DAFTAR GAMBAR.....	xvii
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Tujuan Penelitian	3
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian.....	5
1.6 Sistematika Penulisan	5
BAB1 PENDAHULUAN	5
BAB II TINJAUAN PUSTAKA	5
BAB III METODOLOGI PENELITIAN	6
BAB IV HASIL DAN PEMBAHASAN.....	6
BAB V PENUTUP	6
BAB II TINJAUAN PUSTAKA.....	8
2.1 Penelitian Terdahulu.....	8
2.2 Teori Penunjang.....	13

2.2.1 Analisi Sentimen	14
2.2.2 <i>Naïve Bayes</i>	15
2.2.3 <i>K - Nearest Neighbors</i>	22
2.2.4 Google Colaboratory	24
BAB III METODOLOGI PENELITIAN	26
3.1 Alur Penelitian	26
3.2 Studi Pustaka	27
3.3 Pengumpulan Data	27
3.4 Pelabelan Data	27
3.5 <i>Cleaning Awal</i>	28
3.6 <i>Preprocessing</i>	28
3.6.1 <i>Case Folding</i>	28
3.6.2 <i>Remove Punctuation</i>	28
3.6.3 <i>Stopword Removal</i>	28
3.6.4 <i>Tokenizing</i>	29
3.6.5 <i>Stemming</i>	29
3.7 Visualisasi	29
3.7.1 Diagram Lingkaran	29
3.7.2 <i>Word Cloud</i>	30
3.8 Klasifikasi	30
3.8.1 <i>Splitting Data</i>	30
3.8.2 Pembobotan Kata	33
3.8.3 <i>Naïve Bayes</i>	35
3.8.4 <i>K - Nearest Neighbors</i>	36
3.9 Hasil	37
3.9.1 Akurasi	38
3.9.2 <i>Precision</i>	38
3.9.3 <i>Recall</i>	39
3.9.4 <i>F1 – Score</i>	39
3.9.5 <i>Confusion Matrix</i>	40
3.9.6 Perbandingan Akurasi	42
BAB IV HASIL DAN PEMBAHASAN	43

4.1 Crawling Data.....	43
4.2 Pelabelan Data.....	44
4.3 <i>Cleaning</i> Awal	44
4.4 <i>Preprocessing</i>	46
4.4.1 <i>Case Folding</i>	46
4.4.2 <i>Remove Punctuation</i>	47
4.4.3 <i>Stopword Removal</i>	48
4.4.4 <i>Tokenizing</i>	50
4.4.5 <i>Stemming</i>	51
4.5 Inisialisasi Sentimen.....	52
4.6 Visualisasi.....	54
4.6.1 Diagram Lingkaran	54
4.6.2 <i>Word Clouds</i>	56
4.7 Klasifikasi	59
4.7.1 <i>Splitting</i> Data.....	59
4.7.2 Pembobotan Kata	61
4.7.3 Metode <i>Naïve Bayes</i>	63
4.7.4 Prediksi Metode <i>Naïve Bayes</i>	64
4.7.5 <i>Confusion Matrix</i> Metode <i>Naïve Bayes</i>	65
4.7.6 Metode <i>K-Nearest Neighbor</i>	67
4.7.7 Prediksi Metode <i>K-Nearest Neighbor</i>	68
4.7.8 <i>Confusion Matrix</i> Metode <i>K-Nearest Neighbor</i>	69
4.8 Hasil	70
4.8.1 Akurasi Metode <i>Naïve Bayes</i>	71
4.8.2 Akurasi Metode <i>K-Nearest Neighbor</i>	73
4.8.3 Perbandingan Metode <i>Naïve Bayes</i> dan <i>K-Nearest Neighbor</i>	76
4.9 Hasil Analisis	78
BAB V PENUTUP	81
5.1 Kesimpulan	81
5.2 Saran	82
DAFTAR PUSTAKA	83
LAMPIRAN.....	90

DAFTAR TABEL

Tabel 2.1 Perbandingan Penelitian Terdahulu	10
Tabel 4. 1 Perbandingan metode <i>Naïve Bayes</i> dan <i>K-Nearest Neighbor</i>	77

DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian.....	26
Gambar 3. 2 <i>Confusion Matrix</i> Tiga Kelas (Positif, Negatif, dan Netral)	42
Gambar 4. 1 <i>Crawling</i> Data Menggunakan <i>Google Colab</i>	43
Gambar 4. 2 Hasil <i>Crawling</i> Data Dalam Bentuk File <i>Excel</i>	44
Gambar 4. 3 Pelabelan Data Dalam Bentuk File <i>Excel</i>	44
Gambar 4. 4 <i>Cleaning</i> Awal.....	45
Gambar 4. 5 Hasil <i>Cleaning</i> Awal	46
Gambar 4. 6 <i>Case Folding</i>	47
Gambar 4. 7 Hasil <i>Case Folding</i>	47
Gambar 4. 8 <i>Remove Punctuation</i>	48
Gambar 4. 9 Hasil <i>Remove Punctuation</i>	48
Gambar 4. 10 <i>Normalisasi</i>	49
Gambar 4. 11 <i>Remove Punctuation</i>	49
Gambar 4. 12 Hasil <i>Remove Punctuation</i>	50
Gambar 4. 13 <i>Tokenzing</i>	50
Gambar 4. 14 Hasil <i>Tokenzing</i>	51
Gambar 4. 15 <i>Stemming</i>	52
Gambar 4. 16 Hasil <i>Stemming</i>	52
Gambar 4. 17 Inisialisasi Sentimen.....	53
Gambar 4. 18 Hasil Inisialisasi Sentimen	54
Gambar 4. 19 Digram Lingkaran	55
Gambar 4. 20 Hasil Diagram Lingkaran	56
Gambar 4. 21 <i>Word Clouds</i>	57
Gambar 4. 22 <i>Word Clouds</i> Positif	58
Gambar 4. 23 <i>Word Clouds</i> Netral.....	58
Gambar 4. 24 <i>Word Clouds</i> Negatif.....	59
Gambar 4. 25 <i>Splitting</i> Data.....	60
Gambar 4. 26 Hasil <i>Splitting</i> Data	61
Gambar 4. 27 Pembobotan Kata	62
Gambar 4. 28 Hasil Pembobotan Kata	63

Gambar 4. 29 Metode <i>Naïve Bayes</i>	64
Gambar 4. 30 Prediksi Metode <i>Naïve Bayes</i>	65
Gambar 4. 31 <i>Confusion Matrix</i> Metode <i>Naïve Bayes</i>	65
Gambar 4. 32 Hasil <i>Confusion Matrix</i> Metode <i>Naïve Bayes</i>	66
Gambar 4. 33 Metode <i>K-Nearest Neighbor</i>	68
Gambar 4. 34 Prediksi Metode <i>K-Nearest Neighbor</i>	68
Gambar 4. 35 <i>Confusion Matrix</i> Metode <i>K-Nearest Neighbor</i>	69
Gambar 4. 36 Hasil <i>Confusion Matrix</i> Metode <i>K-Nearest Neighbor</i>	70
Gambar 4. 37 Akurasi Metode <i>Naive Bayes</i>	72
Gambar 4. 38 Hasil Akurasi Metode <i>Naive Bayes</i>	73
Gambar 4. 39 Akurasi Metode <i>K-Nearest Neighbor</i>	75
Gambar 4. 40 Hasil Akurasi Metode <i>K-Nearest Neighbor</i>	76