

**ANALISIS SENTIMEN TERHADAP PROGRAM MAKAN BERGIZI
GRATIS DENGAN METODE *DEEP LEARNING LONG SHORT TERM
MEMORY* DAN *GATED RECURRENT UNIT* PADA MEDIA SOSIAL X**

(Skripsi)

Oleh

**RHALASYA ELEINA PUTRI
NPM 2217051083**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2026**

ABSTRAK

ANALISIS SENTIMEN TERHADAP PROGRAM MAKAN BERGIZI GRATIS DENGAN METODE *DEEP LEARNING LONG SHORT TERM MEMORY* DAN *GATED RECURRENT UNIT* PADA MEDIA SOSIAL X

Oleh

RHALASYA ELEINA PUTRI

Penelitian ini menganalisis sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG) pada tingkat SMA melalui *tweet* berbahasa Indonesia di media sosial X pada periode Januari-Agustus 2025. Data diperoleh melalui *scrapping data* berdasarkan kata kunci terkait MBG sebanyak 13.486 *tweet* dan setelah penghapusan duplikasi data tersisa 12.476 *tweet*. Pelabelan sentimen dilakukan secara semi-otomatis melalui kombinasi pelabelan manual dan IndoBERT sehingga diperoleh 5.039 kelas negatif, 3.754 kelas netral, dan 3.682 kelas positif. Untuk mengatasi ketidakseimbangan kelas, dilakukan pada data latih menggunakan ADASYN. Selanjutnya, pemodelan klasifikasi sentimen dilakukan dengan model *Long Short-Term Memory* dan *Gated Recurrent Unit*, dan menggunakan *Stratified k-Fold Cross Validation* (k=10). Hasil pengujian dengan model LSTM memberikan performa terbaik dengan akurasi 77,63%, presisi 77,51%, *recall* 77,4%, *F1-score* 77,46%, dan AUC 91,54%, sementara GRU memperoleh akurasi 75,59%, presisi 75,24%, *recall* 75,71%, *F1-score* 75,4%, dan AUC 90,82%. Penelitian ini juga melakukan analisis linguistik dengan n-gram, POS *tagging*, dan *dependency parsing*, serta analisis temporal dan spasial untuk mengidentifikasi pola percakapan dan faktor pendorong sentimen. Secara umum, percakapan didominasi sentimen negatif yang banyak menyoroti isu kualitas makanan dan anggaran, sementara sentimen positif muncul pada narasi manfaat gizi dan pemerataan program, dan sentimen netral cenderung berupa informasi pelaksanaan tanpa penilaian eksplisit.

Kata Kunci: Analisis Sentimen, Program Makan Bergizi Gratis, X, *Long Short Term Memory*, *Gated Recurrent Unit*, *Adaptive Synthetic Sampling Approach*.

ABSTRACT

SENTIMENT ANALYSIS OF THE FREE NUTRITIOUS MEAL PROGRAM USING DEEP LEARNING METHODS LONG SHORT-TERM MEMORY AND GATED RECURRENT UNIT ON X SOCIAL MEDIA

By

RHALASYA ELEINA PUTRI

This study analyzes public sentiment toward the Free Nutritious Meal Program (MBG) at the senior high school level using Indonesian-language tweets on X from January to August 2025. The data were collected through web scraping with MBG-related keywords, resulting in 13.486 tweets, after removing duplicates 12.476 tweets were used for analysis. Sentiment labels were assigned semi-automatically by combining manual labeling and IndoBERT, producing 5.039 negative, 3.754 neutral, and 3.682 positive tweets. To handle class imbalance, ADASYN oversampling was applied to the training data. Sentiment classification was then performed using Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models and evaluated with Stratified k-Fold Cross-Validation (k = 10). The results show that LSTM performed better, with an accuracy of 77.63%, precision of 77.51%, recall of 77.40%, F1-score of 77.46%, and AUC of 91.54%, while GRU achieved an accuracy of 75.59%, precision of 75.24%, recall of 75.71%, F1-score of 75.40%, and AUC of 90.82%. This study also used linguistic analyses n-grams, POS tagging, and dependency parsing, as well as temporal and spatial analyses to examine discussion patterns and the main factors behind public sentiment. Overall, negative sentiment was the most common and often focused on food quality and budget allocation, while positive sentiment highlighted narratives of nutritional benefits and program equity, and neutral sentiment mostly shared implementation information without clear judgment.

Keywords: Sentiment Analysis, Free Nutritious Meal Program, X, Long Short Term Memory, Gated Recurrent Unit, Adaptive Synthetic Sampling Approach.

**ANALISIS SENTIMEN TERHADAP PROGRAM MAKAN BERGIZI
GRATIS DENGAN METODE *DEEP LEARNING LONG SHORT TERM
MEMORY* DAN *GATED RECURRENT UNIT* PADA MEDIA SOSIAL X**

Oleh

RHALASYA ELEINA PUTRI

Proposal Penelitian

**Sebagai Salah Satu Syarat untuk Mencapai Gelar
SARJANA KOMPUTER**

Pada

**Jurusan Ilmu Komputer
Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2026**

Judul Skripsi : **ANALISIS SENTIMEN TERHADAP PROGRAM MAKAN BERGIZI GRATIS DENGAN METODE *DEEP LEARNING LONG SHORT TERM MEMORY* DAN *GATED RECURRENT UNIT* PADA MEDIA SOSIAL X**

Nama Mahasiswa : **Rhalasya Eleina Putri**

Nomor Pokok Mahasiswa : 2217051083

Program Studi : S1-Ilmu Komputer

Fakultas : Matematika dan Ilmu Pengetahuan Alam



1. Komisi Pembimbing

Dewi Asiah Shofiana, S.Komp., M.Kom.

NIP. 199509292020122030

Rahman Taufik, S.Pd., M.Kom.

NIP. 199306272022031007

2. Ketua Jurusan Ilmu Komputer

Dwi Sakethi, S.Si., M.Kom.

NIP. 196806111998021001

3. Ketua Program Studi S1 Ilmu Komputer

Tristiyanto, S.Kom., M.I.S., Ph.D.

NIP. 198104142005011001

MENGESAHKAN

1. Tim Penguji

Ketua : Dewi Asiah Shofiana, S.Komp., M.Kom.



Sekretaris : Rahman Taufik, S.Pd., M.Kom.



Penguji

Bukan Pembimbing : Favorisen Rosyking Lumbanraja, Ph.D.



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam




Dr. Eng. Heri Satria, S.Si., M.Si.

NIP. 197110012005011002

Tanggal Lulus Ujian Skripsi: 10 Maret 2026

PERNYATAAN

Saya yang bertanda tangan di bawah ini.

Nama : Rhalasya Eleina Putri

NPM : 2217051083

Dengan ini menyatakan bahwa skripsi saya yang berjudul “**Analisis Sentimen Terhadap Program Makan Bergizi Gratis Dengan Metode *Deep Learning Long Short Term Memory Dan Gated Recurrent Unit* Pada Media Sosial X**” merupakan hasil karya saya sendiri dan bukan karya orang lain. Seluruh tulisan yang tertuang di skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila di kemudian hari terbukti skripsi saya merupakan hasil penjiplakan atau dibuat orang lain, maka saya bersedia menerima sanksi berupa pencabutan gelar yang telah saya terima.

Bandar Lampung, 7 April 2026



Rhalasya Eleina Putri

NPM. 2217051083

RIWAYAT HIDUP



Penulis dilahirkan pada tanggal 20 April 2004 di Bandar Lampung sebagai anak ketiga dari pasangan Bapak Edwin Hanibal, S.H., M.H (Alm) dan Ibu Astronati Dyah Candrarini, S.H. Penulis menyelesaikan pendidikan Sekolah Dasar (SD) di SD Kartika II-5 pada Tahun 2016. Kemudian dilanjutkan pendidikan menengah pertama di SMP Negeri 23 Bandar Lampung pada Tahun 2019, dan melanjutkan ke pendidikan menengah atas di SMAN 2 Bandar Lampung yang diselesaikan pada Tahun 2022. Pada tahun 2022, penulis terdaftar sebagai mahasiswa Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung melalui jalur SBMPTN. Selama menjadi mahasiswa, penulis melakukan beberapa kegiatan antara lain.

1. Menjadi Anggota Muda Ilmu Komputer (ADAPTER) Himpunan Mahasiswa Jurusan Ilmu Komputer pada periode 2022/2023.
2. Menjadi Anggota Badan Khusus Himpunan Mahasiswa Jurusan Ilmu Komputer pada periode 2022/2023.
3. Menjadi Bendahara Badan Khusus Himpunan Mahasiswa Jurusan Ilmu Komputer pada periode 2023/2024.
4. Menjadi Asisten Dosen Jurusan Ilmu Komputer pada mata kuliah Matematika pada Tahun 2023, Rekayasa Perangkat Lunak (RPL) pada Tahun 2024, dan Basis Data pada Tahun 2025.
5. Menjadi Kepala Departemen IV pada acara *Computer Science Showdown* (CSS) Tahun 2024.
6. Mengikuti Program MSIB Kampus Merdeka di Bangkit *Academy* 2024 Batch 2 pada periode 2024/2025 dengan bidang *Cloud Computing*.

7. Mengikuti Kuliah Kerja Nyata (KKN) periode 2024/2025 di Desa Lingga Pura, Kecamatan Selagai Lingga Lampung Tengah.
8. Melaksanakan kerja praktik di Kantor Pusat Bank Lampung divisi Akuntansi pada Tahun 2025.

MOTTO

“Maka sesungguhnya bersama kesulitan ada kemudahan, sesungguhnya bersama kesulitan ada kemudahan.”

(QS. Al-Insyirah: 5-6)

“You take your plans and you make them true, you’ll see your life in a different view.”

(Reality Club)

“No need to run, just walk and see everything around us.”

(Mark Lee)

PERSEMBAHAN

Alhamdulillahirobbilalamin

Puji syukur kehadiran Allah Subhanahu Wa Ta'ala atas segala rahmat dan karunia-Nya sehingga skripsi ini dapat diselesaikan dengan sebaik-baiknya. Shalawat serta salam selalu tercurahkan kepada junjungan Nabi Agung Muhammad Shallallahu'Alaihi Wasallam.

Kupersembahkan karya ini kepada:

Kedua Orang Tuaku Tercinta

Yang selalu mendukung, memberikan cinta dan kasih sayang yang tak terhingga, serta do'a yang selalu menyertaiku. Kuucapkan terima kasih sebesar-besarnya atas pengorbanan dan perjuangan dalam mendidik dan membesarkanku yang tak akan dapat terbalaskan.

Seluruh Keluarga Besar Ilmu Komputer 2022

Yang senantiasa memberikan semangat dan dukungan.

Almamater Tercinta Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung

Tempat bernaung mengemban semua ilmu untuk menjadi bekal hidup.

SANWACANA

Puji syukur kehadiran Allah Subhanahu Wa Ta'ala atas limpahan nikmat, rahmat dan karunia-Nya. Shalawat serta salam senantiasa tercurahkan kepada junjungan Nabi Muhammad Shallallahu 'Alaihi Wasallam, sehingga penulis dapat menyelesaikan skripsi yang berjudul "**Analisis Sentimen Terhadap Program Makan Bergizi Gratis Dengan Metode *Deep Learning Long Short Term Memory* Dan *Gated Recurrent Unit* Pada Media Sosial X**" dengan baik dan lancar.

Terima kasih penulis ucapkan kepada pihak-pihak yang telah memberi dukungan, bimbingan, dan arahan pada penulis dalam menyelesaikan penyusunan skripsi ini. Ucapan terima kasih penulis tujukan kepada.

1. Ibuku tersayang, Astronati Dyah Candrarini, yang selalu mengajarkan kebaikan, memberikan doa dan motivasi yang luar biasa, serta mencurahkan kasih sayang dengan penuh cinta. Alhamdulillah kini penulis sedikit lagi menyelesaikan studi untuk meraih gelar sarjana. Terima kasih atas dukungan dan kepercayaanmu yang tiada henti bahkan sampai saat ini. Ayahku tercinta, Alm. Edwin Hanibal, yang selama hidupnya senantiasa menyayangi, mengajarkan menjadi tangguh, dan memberikan doa, kepercayaan, serta dukungan baik secara moral maupun material. Terima kasih atas perjuanganmu dalam memberikan kesejahteraan dan kenyamanan hidup bahkan hingga saat ini sehingga penulis bisa berada di titik ini.
2. Bapak Dr. Eng. Heri Satria, S.Si., M.Si. selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam.
3. Bapak Dwi Sakethi S.Si., M.Kom. selaku Ketua Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.

4. Ibu Yunda Heningtyas, M.Kom. selaku Sekretaris Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
5. Bapak Tristiyanto S.Kom., M.I.S., Ph.D. selaku Ketua Program Studi S1 Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung.
6. Bapak Favorisen Rosyking Lumbanraja, Ph.D. selaku Pembahas dan Pembimbing Akademik yang telah memberikan masukan dan saran yang bermanfaat untuk perbaikan penelitian ini.
7. Ibu Dewi Asiah Shofiana, S.Komp., M.Kom. selaku Pembimbing Utama yang telah memberikan arahan, motivasi, semangat, ide, kritik dan saran dalam menyelesaikan penelitian ini.
8. Bapak Rahman Taufik, S.Pd., M.Kom. selaku Pembimbing Kedua yang selalu meluangkan waktu untuk membimbing, memberikan arahan serta kritik dalam menyelesaikan penelitian ini.
9. Bapak dan Ibu Dosen Jurusan Ilmu Komputer FMIPA Universitas Lampung yang telah memberikan ilmu, motivasi dan pengalaman hidup selama penulis menempuh pendidikan di Jurusan Ilmu Komputer Universitas Lampung.
10. Seluruh staf dan karyawan Jurusan Ilmu Komputer, Ibu Ade Nora Mela, Bang Zainuddin, Mas Nofal, Mas Sam dan Pak Dahud yang telah membantu segala urusan administrasi penulis di Jurusan Ilmu Komputer.
11. Kedua Kakakku yang selalu mendukung dan menyemangati selama proses penyusunan skripsi ini. Terima kasih atas doa, dukungan dan tindakan untuk penulis hingga saat ini.
12. Teman-teman seperjuangan grup “Pragib Cantik” yaitu Nawa, Yayas, Patiya, Ancit, Pia, dan Atikah yang selalu mendukung, membantu, menemani, menghibur dan berbagi cerita indah dari awal masa perkuliahan sampai saat ini. Terima kasih telah berteman baik dengan kebersamai dan membantu penulis dalam berbagai hal.
13. Sahabat penulis Aurelia Fabiola Iranda yang selalu mendengarkan keluh kesah penulis dan memberikan motivasi dari sebelum masa perkuliahan.

Terima kasih telah menjadi teman baik penulis dan selalu mendukung serta mengapresiasi penulis hingga saat ini.

14. Teman-teman Pimpinan Himakom 2024 yang sudah mengajarkan banyak hal dalam berorganisasi, memberikan banyak pengalaman dan mendukung penulis.
15. Keluarga Ilmu Komputer 2022 yang telah memberikan pengalaman berharga. Terima kasih telah menjadi rekan kelompok, rekan diskusi dan rekan berjuang selama menjalankan studi di Jurusan Ilmu Komputer Universitas Lampung.
16. Seluruh pihak yang telah membantu dan memberikan dukungan baik secara langsung maupun tidak langsung selama perkuliahan hingga penyelesaian skripsi ini.
17. Diri saya sendiri Rhalasya Eleina yang telah bekerja keras dan selalu semangat sehingga tidak menyerah selama proses penyusunan skripsi ini. Terima kasih karena mampu mengendalikan stress sehingga tidak pernah diopname selama proses penyusunan skripsi dan bertanggung jawab serta menikmati setiap proses studi hingga saat ini.

Penulis menyadari bahwa penyusunan skripsi ini masih jauh dari kata sempurna. Namun, penulis sangat mengharapkan skripsi ini dapat bermanfaat khususnya bagi masyarakat Indonesia dan pemangku jabatan di Indonesia yang membacanya, sehingga dapat menjadi bahan evaluasi program Makan Bergizi Gratis untuk masa depan Indonesia yang lebih cerah.

Bandar Lampung, 7 April 2026

Rhalasya Eleina Putri

NPM. 2217051083

DAFTAR ISI

	Halaman
DAFTAR ISI	xv
DAFTAR TABEL.....	xix
DAFTAR GAMBAR.....	xxi
I. PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah.....	3
1.3. Tujuan Penelitian	4
1.4. Batasan Masalah	4
1.5. Manfaat Penelitian	5
II. TINJAUAN PUSTAKA.....	6
2.1. Analisis Sentimen	6
2.2. <i>Natural Language Processing</i>	6
2.3. <i>Deep Learning</i>	9
2.4. X.....	9
2.5. Makan Bergizi Gratis	10
2.6. <i>Text Mining</i>	10
2.7. IndoBERT	12
2.8. <i>Adaptive Synthetic Sampling Approach</i>	12

2.9.	<i>Stratified k-fold Cross Validation</i>	14
2.10.	<i>Recurrent Neural Network</i>	15
2.11.	<i>Long Short-Term Memory</i>	15
2.12.	<i>Gated Recurrent Unit</i>	18
2.13.	<i>Confusion Matrix</i>	19
2.14.	<i>Receiver Operating Characteristic (ROC)</i>	21
2.15.	<i>Area Under Curve (AUC)</i>	21
2.16.	<i>Functional Programming</i>	22
2.17.	Python	23
2.17.1.	NumPy dan Pandas.....	23
2.17.2.	Scikit-learn dan Imabalanced-learn	23
2.17.3.	Wordcloud dan Matplotlib.....	24
2.17.4.	TensorFlow dan Transformers.....	24
2.18.	Penelitian Terdahulu	25
III.	METODOLOGI PENELITIAN	29
3.1.	Waktu dan Tempat Penelitian	29
3.1.1.	Tempat Penelitian	29
3.1.2.	Waktu Penelitian.....	29
3.2.	Analisis Kebutuhan Sistem.....	30
3.2.1	Perangkat Keras (<i>Hardware</i>).....	30
3.2.2	Perangkat <i>Cloud</i>	30
3.2.3	Perangkat Lunak (<i>Software</i>)	30
3.3.	Tahapan Penelitian	31
3.3.1.	Studi Literatur.....	33
3.3.2.	Pengumpulan Data.....	33
3.3.3.	<i>Exploratory Data Analysis</i>	33
3.3.4.	<i>Data Preprocessing</i>	34
3.3.5.	Pelabelan.....	36
3.3.6.	Pembagian Data.....	37

3.3.7.	<i>Data Balancing</i>	38
3.3.8.	<i>Training Long Short Term Memory (LSTM)</i>	38
3.3.9.	<i>Training Gated Recurrent Unit (GRU)</i>	39
3.3.10.	Evaluasi Model.....	40
3.4.	Analisis Linguistik.....	41
3.4.1.	Analisis Temporal dan Spasial	41
3.4.2.	N-Gram.....	42
3.4.3.	<i>Part Of Speech Tagging</i>	43
3.4.4.	<i>Depedency Parsing</i>	43
IV.	HASIL DAN PEMBAHASAN	44
4.1.	Pengumpulan Data.....	44
4.2.	<i>Exploratory Data Analysis</i>	44
4.2.1.	Distribusi Panjang Kata.....	45
4.2.2.	Distribusi Panjang Karakter	46
4.2.3.	Deteksi Duplikasi <i>Tweet</i>	47
4.3.	<i>Pre-Processing Data</i>	47
4.3.1.	<i>Case Folding</i>	47
4.3.2.	<i>Data Cleaning</i>	48
4.3.3.	Normalisasi Data	49
4.4.	Pelabelan.....	49
4.5.	Pembagian Data	51
4.6.	<i>Data Balancing</i>	51
4.7.	<i>Hyperparameter Tuning</i>	52
4.8.	<i>Training Long Short Term Memory</i>	55
4.8.1.	<i>Stratified k-Fold Cross Validation</i>	56
4.8.2.	<i>Confusion Matrix LSTM</i>	58
4.8.3.	Evaluasi Model LSTM	59
4.9.	<i>Training Gated Recurrent Unit</i>	60
4.9.1.	<i>Stratified k-Fold Cross Validation</i>	61
4.9.2.	<i>Confusion Matrix GRU</i>	63

4.9.3.	Evaluasi Model GRU	64
4.10.	<i>Word Cloud</i>	65
4.11.	Analisis N-gram per Sentimen.....	67
4.11.1.	<i>Tweet</i> Negatif.....	67
4.11.2.	<i>Tweet</i> Positif	69
4.11.3.	<i>Tweet</i> Netral.....	70
4.12.	Analisis <i>Part Of Speech Tagging</i>	71
4.13.	Analisis <i>Dependency Parsing</i>	74
4.13.1.	Level Kata	74
4.13.2.	Level Frasa	78
4.14.	Temuan Faktor Pendorong Sentimen.....	83
4.14.1.	Dinamika Temporal	83
4.14.2.	Perbedaan Spasial.....	84
4.14.3.	Aspek Tiap Sentimen.....	91
V.	SIMPULAN DAN SARAN	95
5.1.	Simpulan	95
5.2.	Saran	96
	DAFTAR PUSTAKA	97

DAFTAR TABEL

Tabel	Halaman
1. <i>Confusion Matrix</i>	19
2. Penelitian Terdahulu.....	25
3. <i>Timeline</i> Penelitian.....	29
4. Statistik Panjang Kata	45
5. Statistik Panjang Karakter.....	46
6. Contoh <i>Case Folding</i>	48
7. Contoh <i>Data Cleaning</i>	48
8. Contoh Normalisasi	49
9. Contoh Pelabelan Manual	49
10. Contoh Pelabelan IndoBERT	50
11. Rentang Nilai <i>Hyperparameter Tuning</i>	53
12. Eksplorasi Skenario <i>Balancing</i>	53
13. Hasil SKCV model LSTM.....	57
14. Hasil SKCV Model GRU.....	61
15. Hasil N-Gram Kelas Negatif.....	67
16. Hasil N-Gram Kelas Positif	69
17. Hasil N-Gram Kelas Netral.....	70
18. Top 5 POS <i>Tagging</i> Kata Sifat	72
19. Top 5 POS <i>Tagging</i> Kata Benda	73
20. Top 5 POS <i>Tagging</i> Kata Kerja.....	73
21. Top 5 Subjek <i>Depedency Parsing</i> Level Kata.....	74
22. Top 5 Objek <i>Depedency Parsing</i> Level Kata.....	75
23. Top 5 <i>Verb Depedency Parsing</i> Level Kata	76
24. Contoh Pola SVO Level Kata	77
25. Top 5 Subjek <i>Depedency Parsing</i> Level Frasa	79

26. Top 5 Objek <i>Depedency Parsing</i> Level Frasa	80
27. Top 5 <i>Verb Depedency Parsing</i> Level Frasa.....	81
28. Contoh Pola SVO Level Frasa.....	81

DAFTAR GAMBAR

Gambar	Halaman
1. Distribusi ADASYN (Alhudhaif, 2021).	13
2. Arsitektur Long Short-Term Memory (Zhang et al., 2023).....	16
3. Arsitektur Gated Recurrent Unit (Zhang et al., 2023).....	18
4. Tahapan Penelitian.....	32
5. Visualisasi Hasil Pelabelan.....	37
6. Histogram Panjang Kata.....	45
7. Histogram Panjang Karakter.....	46
8. Histogram Distribusi ADASYN.....	52
9. Eksplorasi <i>Learning Curves</i> LSTM.....	54
10. Eksplorasi <i>Learning Curves</i> GRU.....	55
11. Grafik <i>Train</i> LSTM.....	57
12. <i>Confusion Matrix</i> LSTM.....	58
13. Kurva ROC Model LSTM.....	60
14. Grafik <i>Train</i> GRU.....	62
15. <i>Confusion Matrix</i> GRU.....	63
16. Kurva ROC Model GRU.....	64
17. <i>Word Cloud</i> Sentimen Negatif.....	65
18. <i>Word Cloud</i> Sentimen Netral.....	66
19. <i>Word Cloud</i> Sentimen Positif.....	67
20. Distribusi POS <i>Tagging</i>	71
21. Grafik Volume <i>Tweet</i> per Bulan.....	83
22. Histogram Sentimen per Pulau.....	84
23. <i>Word Cloud</i> Tiap Pulau.....	85
24. Histogram Sentimen Pulau Papua.....	86
25. Histogram Sentimen Pulau Kalimantan.....	86

26. Histogram Sentimen Pulau Maluku.....	87
27. Histogram Sentimen Pulau Jawa.	88
28. Histogram Sentimen Pulau Sumatera.	89
29. Histogram Sentimen Pulau Sulawesi.....	89
30. Histogram Sentimen Pulau Nusa Tenggara.	90

I. PENDAHULUAN

1.1. Latar Belakang

Perkembangan teknologi digital telah membawa perubahan signifikan dalam berbagai aspek kehidupan, termasuk dalam hal penyebaran informasi dan komunikasi publik. Media sosial menjadi salah satu *platform* utama yang digunakan masyarakat untuk menyampaikan opini, kritik, maupun dukungan terhadap isu-isu terkini. Salah satu isu yang ramai diperbincangkan di media sosial X (sebelumnya Twitter) adalah Program Makan Bergizi Gratis (MBG) yang menjadi salah satu janji utama pasangan Prabowo-Gibran pada Pemilu 2024. Opini masyarakat terbagi menjadi emosi positif, negatif, maupun netral. Berdasarkan perasaan tersebut, dapat digunakan sebagai bahan evaluasi yang membantu pemerintah Indonesia untuk mengkaji ulang program yang didasarkan opini masyarakat.

Program MBG bertujuan untuk memberikan makan siang gratis kepada anak-anak sekolah dan kelompok rentan, sebagai upaya meningkatkan kualitas gizi dan menurunkan angka stunting di Indonesia. Namun, program ini juga menimbulkan kontroversi dan perdebatan publik pada implementasinya di berbagai aspek. Meskipun Program Makan Bergizi Gratis (MBG) bertujuan mulia untuk meningkatkan gizi anak-anak, tetapi pada pelaksanaannya justru memunculkan pro dan kontra di kalangan publik. Opini pro dan kontra banyak disampaikan pada *platform* X. Salah satu target yang paling banyak menuai pro dan kontra adalah MBG dengan target anak-anak di sekolah menengah atas (SMA).

Kementerian Keuangan RI mengalokasikan dana sebesar 71 triliun rupiah pada tahun 2025 dan 255 triliun rupiah pada tahun 2026 untuk Badan Gizi Nasional (BGN) yang tertulis pada Rancangan Undang-Undang Anggaran Pendapatan dan Belanja Negara (RUU APBN) (Indonesia, 2025). Kondisi ini menimbulkan kekhawatiran bagi sebagian masyarakat dan menganggap anggaran tersebut terlalu besar serta dapat mengganggu alokasi dana untuk sektor penting lainnya seperti infrastruktur, pendidikan, dan kesehatan.

Di sisi lain, masih terdapat keraguan mengenai efektivitas serta tata kelola pelaksanaan program, mengingat persoalan transparansi dan efisiensi dalam penggunaan anggaran negara belum sepenuhnya teratasi. Akan tetapi, pengamatan secara langsung melalui media sosial tidak mampu mencerminkan opini publik secara keseluruhan. Hal ini disebabkan oleh algoritma media sosial X yang menyesuaikan konten berdasarkan preferensi dan riwayat interaksi pengguna sehingga menimbulkan fenomena *bubble filter*, yaitu kondisi ketika pengguna hanya terekspos pada informasi yang sejalan dengan pandangannya sendiri (Rehani, 2020). Fenomena ini kerap diperkuat oleh *echo chamber*, yaitu lingkungan komunikasi daring saat opini yang serupa terus bergema melalui interaksi berulang antar pengguna dengan pandangan yang sama (Cinelli *et al.*, 2021). Akibatnya, pengguna yang mendukung MBG cenderung lebih banyak terpapar unggahan positif, sementara yang menolak justru lebih sering menemukan unggahan negatif.

Keadaan tersebut menimbulkan bias informasi dan membentuk persepsi yang parsial sehingga analisis manual tidak dapat menjamin keakuratan maupun validitas hasil. Untuk menjawab keterbatasan ini, pemanfaatan teknologi kecerdasan buatan dengan pendekatan *deep learning* menjadi solusi yang tepat. Model seperti *Long Short-Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU) memiliki kemampuan mengolah ribuan data secara menyeluruh serta memahami kerumitan konteks linguistik, dan mengklasifikasikan sentimen dengan tingkat konsistensi lebih tinggi dibandingkan metode konvensional.

Banyak penelitian yang mengeksplorasi analisis sentimen, salah satunya studi oleh (Khadapi & Pakpahan, 2024) juga membandingkan LSTM dengan BERT untuk analisis sentimen Pemilu 2024 di *Twitter*, metode LSTM mencapai akurasi maksimal 87% sedangkan metode *Bidirectional Encoder Representations from Transformers* (BERT) hanya sebesar 76.48%. Beberapa penelitian mengenai analisis sentimen menggunakan metode *deep learning Gated Recurrent Unit* (GRU). Model ini dikenal karena efisiensi komputasinya dan kemampuannya yang efektif dalam menangkap pola dalam data sekuensial. Penelitian (Edward *et al.*, 2023) menunjukkan efektivitas model GRU dalam menganalisis sentimen *tweet* Donald Trump pada Pemilu 2020 dengan akurasi 93%, presisi 96%, dan *recall* 94% yang menunjukkan kekuatan model dalam mengategorikan sikap dengan tepat. Penelitian (Kawatu, 2024) juga menerapkan metode GRU dalam analisis sentimen media sosial untuk mengklasifikasikan opini publik menjadi positif, negatif, dan netral yang mencapai akurasi keseluruhan 70%.

Berdasarkan uraian tersebut, dapat disimpulkan bahwa analisis sentimen berbasis media sosial X terhadap Program Makan Bergizi Gratis sangat penting dilakukan untuk memahami sentimen dan penyebab tiap sentimen. Selain itu, perbandingan antara metode LSTM dan GRU dalam klasifikasi sentimen dapat memberikan kontribusi ilmiah dalam pengembangan metode analisis sentimen di Indonesia. Penelitian ini diharapkan dapat menjadi referensi bagi pemerintah dalam mengevaluasi dan meningkatkan pelaksanaan Program Makan Bergizi Gratis, serta memberikan kontribusi dalam pengembangan ilmu pengetahuan di bidang analisis sentimen berbasis media sosial.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, didapatkan rumusan masalah yang menjadi fokus dalam penelitian ini sebagai berikut.

- a) Bagaimana sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG) melalui analisis kinerja model *Long Short Term Memory* dan *Gated Recurrent Unit* berdasarkan unggahan di media sosial X?
- b) Bagaimana perbedaan performa model *Long Short Term Memory* dan *Gated Recurrent Unit* dalam mengklasifikasikan sentimen tersebut?
- c) Bagaimana hasil analisis linguistik pada sentimen terhadap Program Makan Bergizi Gratis?

1.3. Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut.

- a) Menganalisis sentimen masyarakat terhadap Program Makan Bergizi Gratis (MBG) berdasarkan data dari media sosial X.
- b) Membandingkan dan mengevaluasi hasil akurasi, presisi, *recall*, *F1-score*, dan *ROC-AUC score* menggunakan metode *Long Short-Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU) terhadap analisis sentimen Program Makan Bergizi Gratis.
- c) Mengetahui faktor penyebab sentimen terhadap Program Makan Bergizi Gratis dengan analisis linguistik.

1.4. Batasan Masalah

Adapun batasan masalah dalam penelitian adalah sebagai berikut.

- a) Penelitian ini berfokus pada *tweet* yang berasal dari media sosial X dengan kata kunci terkait Makan Bergizi Gratis dan pelaksanaannya di SMA. Data *tweet* yang digunakan terbatas pada bahasa Indonesia.
- b) Data pada penelitian ini didapatkan dengan rentang waktu dari bulan Januari 2025 – Agustus 2025.

- c) Pengolahan data dilakukan menggunakan bahasa pemrograman Python.
- d) Metode yang digunakan hanya *Long Short-Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU) tanpa membandingkan dengan algoritma lain.
- e) Penelitian tidak membahas implementasi teknis program MBG, tetapi fokus pada analisis sentimen publik.

1.5. Manfaat Penelitian

Adapun manfaat dari penelitian ini adalah sebagai berikut.

- a) Dapat membantu pemerintah untuk mengetahui opini masyarakat tentang Program Makan Bergizi Gratis (MBG) sehingga hasilnya dapat dijadikan bahan pertimbangan dalam memperbaiki atau mengembangkan program tersebut kedepannya.
- b) Memberikan kontribusi ilmiah dalam perbandingan *Long Short Term Memory* dan *Gated Recurrent Unit* untuk analisis sentimen terhadap Program Makan Bergizi Gratis.

II. TINJAUAN PUSTAKA

2.1. Analisis Sentimen

Analisis sentimen adalah cabang dari *Natural Language Processing* (NLP) yang berfokus pada identifikasi, ekstraksi, dan klasifikasi opini dari sebuah teks. Tujuannya adalah menentukan pandangan penulis terhadap suatu topik bersifat positif, negatif, atau netral. Dalam konteks yang lebih luas, analisis sentimen berfungsi sebagai alat untuk memahami pandangan publik secara masif terhadap suatu produk, layanan, atau isu sosial, seperti program pemerintah (Naufal & Kusuma, 2022). Analisis sentimen juga sebagai sarana untuk mengungkap sikap dan emosi yang tersirat dalam data berukuran besar, seperti yang umumnya ditemukan pada media sosial (Fahreza & Setiawan, 2024).

2.2. *Natural Language Processing*

Natural Language Processing (NLP) adalah salah satu bidang dalam kecerdasan buatan yang berfokus pada interaksi komputer dan bahasa manusia. NLP bertujuan mengembangkan sistem yang mampu memahami, menafsirkan, dan menghasilkan bahasa alami sebagaimana manusia berkomunikasi. Pada awal perkembangannya, NLP lebih banyak diarahkan pada pengembangan teori mengenai pemahaman bahasa, namun seiring waktu pendekatan ini berubah menjadi disiplin yang lebih bersifat teknis dan rekayasa. Hal ini disebabkan oleh semakin berkembangnya teknologi komputasi serta melimpahnya data bahasa digital (Lenci & Padó, 2022). NLP berperan penting karena analisis sentimen merupakan salah satu aplikasi utama dari bidang tersebut. Tantangan dalam pengolahan teks di

media sosial, seperti penggunaan bahasa tidak baku, singkatan, hingga keberadaan opini ganda dalam satu kalimat, membutuhkan pendekatan yang lebih kuat seperti *Deep Learning* dibandingkan metode tradisional.

Dalam linguistik, bahasa dapat dianalisis pada beberapa tingkatan (*levels of language*). Tingkatan ini kemudian diadopsi ke dalam NLP sebagai kerangka untuk merancang algoritma dan sistem pemrosesan bahasa. Tingkatan tersebut antara lain fonologi, morfologi, leksikal, sintaksis, semantik, serta wacana dan pragmatik (Jurafsky & Martin, 2018).

A. Tingkat Fonologi dan Fonetik

Tingkat fonologi membahas satuan bunyi (fonem) dan pola pengucapan dalam suatu bahasa. Sedangkan fonetik berhubungan dengan sifat fisik bunyi ujaran. Pada NLP berbasis teks, aspek fonologi tidak selalu menjadi fokus utama, namun tetap relevan pada tugas yang berkaitan dengan pengenalan ujaran (*speech recognition*) dan sintesis ucapan (*text-to-speech*) yang harus memetakan sinyal suara menjadi teks atau sebaliknya.

B. Tingkat Morfologi

Morfologi mempelajari struktur internal kata seperti morfem, afiks, afiksasi, reduplikasi, dan komposisi. Penelitian (Nugraha, 2024) menekankan bahwa morfologi merupakan aspek fundamental linguistik yang berkaitan dengan struktur dan pembentukan kata dan berfungsi sebagai dasar untuk memahami bagaimana morfem membentuk kata dan variasi maknanya. Dalam NLP, analisis morfologi digunakan untuk *stemming* dan lemmatisasi, pemisahan kata terikat, dan penandaan bentuk kata.

C. Tingkat Leksikal

Tingkat leksikal berfokus pada kata sebagai unit dasar (leksikon) dan informasi terkaitnya, seperti kelas kata dan makna dasar. Pada NLP, tingkat ini mencakup tokenisasi, pembentukan kamus kata/n-gram, dan yang terpenting, representasi vektor kata (*embedding*). Penelitian (Torregrossa *et al.*, 2021) mencatat bahwa *embedding* kata telah

terbukti efektif untuk banyak tugas pemrosesan bahasa alami dengan menyediakan representasi kata yang mengintegrasikan pengetahuan sebelumnya sehingga kata-kata yang serupa secara semantik dan sintaksis memiliki vektor yang berjarak dekat dalam ruang *embedding*. Representasi leksikal ini kemudian menjadi masukan standar untuk berbagai tugas hilir seperti klasifikasi teks, analisis sentimen, dan pemodelan sekuens.

D. Tingkat Sintaksis

Sintaksis mempelajari struktur kalimat dan hubungan antar kata di dalam kalimat. Dalam NLP, informasi sintaktis biasanya diperoleh melalui:

- 1) *Part-of-speech (POS) tagging*, yang memberi label kelas kata untuk setiap token.
- 2) *Constituency parsing*, yang memecah kalimat menjadi frasa-frasa seperti NP (*noun phrase*) dan VP (*verb phrase*).
- 3) *Dependency parsing*, yang memodelkan hubungan ketergantungan seperti *nsubj*, *obj*, *obl* untuk memetakan siapa melakukan apa kepada siapa/apa.

Model modern memanfaatkan POS dan *dependency graph* untuk tugas ekstraksi informasi maupun analisis opini berbasis aspek. Dikarenakan struktur subjek–predikat–objek membantu merumuskan pola “siapa melakukan apa terhadap siapa”.

E. Tingkat Semantik

Semantik berhubungan dengan makna kata, frasa, dan kalimat. Pada level ini, sistem tidak hanya mengenali bentuk, tetapi juga mencoba memutuskan makna yang tepat dalam konteks. Beberapa tugas semantik penting dalam NLP antara lain:

- 1) *Word Sense Disambiguation (WSD)* untuk memilih makna kata yang paling sesuai dengan konteks.
- 2) *Named Entity Recognition (NER)* untuk mengenali entitas seperti nama orang, organisasi, lokasi, dan lain-lain.

3) *Semantic Role Labeling* (SRL) untuk memetakan peran semantik seperti pelaku, pasien, dan penerima dalam sebuah peristiwa.

F. Tingkat Wacana dan Pragmatik

Tingkat wacana (*discourse*) melihat hubungan antar kalimat dalam teks yang lebih panjang, koherensi paragraf, struktur argumen, dan koreferensi (rujukan kata ganti ke entitas sebelumnya). Tingkat pragmatik menyoroti makna yang bergantung pada konteks situasi, tujuan penutur, dan pengetahuan bersama.

2.3. *Deep Learning*

Deep Learning merupakan salah satu cabang dari *machine learning* yang menggunakan arsitektur *neural networks* dengan banyak lapisan untuk memodelkan data kompleks. *Deep learning* cocok untuk *Natural Language Processing* (NLP) karena dapat mengekstrak representasi kata berdasarkan konteks kalimat secara otomatis. *Deep learning* dikembangkan untuk menduplikasi cara kerja otak manusia dalam mengenali pola sehingga dapat menangani data dalam skala besar dan tingkat kesulitan tinggi. Hal ini menjadikan *deep learning* sangat relevan dalam pemrosesan bahasa alami (NLP) karena data berupa teks memiliki variasi bentuk, panjang, serta konteks yang dinamis (Ain *et al.*, 2017). Model-model berbasis *neural networks* dinilai lebih unggul dalam memahami hubungan antar kata dalam sebuah kalimat.

2.4. X

X adalah sebuah *platform* media sosial yang sebelumnya bernama Twitter. X setelah diakuisisi oleh X Corp memiliki tujuan memastikan kebebasan berbicara dan akselerator. X berbasis *micro-blogging* yang memungkinkan pengguna menyampaikan opini dalam format singkat (Sanusi *et al.*, 2021). *Tweet* tidak hanya berisi teks biasa, tetapi juga disertai metadata penting seperti *mentions* (@), *hashtags* (#), serta fitur *retweet* yang memungkinkan penyebaran opini secara luas. Bahasa yang digunakan di Twitter cenderung

informal dan variatif, mencakup slang, singkatan, serta ekspresi budaya seperti sarkasme atau humor. Hal ini sering menyulitkan proses analisis karena makna yang tersembunyi dan bergantung konteks. Karakter khas ini menambah kompleksitas *preprocessing* teks dan menuntut strategi representasi data yang lebih cermat.

2.5. Makan Bergizi Gratis

Program makan bergizi gratis merupakan kebijakan pemerintah untuk menyediakan asupan makanan sehat dan seimbang bagi anak-anak sekolah. Kebijakan ini dirancang untuk menjamin terpenuhinya kebutuhan gizi yang seimbang, sehingga dapat mendukung proses tumbuh kembang anak sekaligus meningkatkan kualitas sumber daya manusia di masa depan. Seperti yang dijelaskan dalam penelitian kebijakan terbaru, program ini dimaksudkan untuk memberikan makanan dengan kandungan gizi lengkap agar anak-anak dapat memperoleh energi yang cukup dalam menjalani aktivitas sehari-hari, khususnya dalam proses belajar di sekolah. Meskipun memiliki manfaat yang besar, implementasi program ini di lapangan menghadapi berbagai tantangan (Merlinda & Yusuf, 2025). Salah satu persoalan yang sering muncul adalah keberlanjutan pendanaan, mengingat program ini membutuhkan alokasi anggaran yang cukup besar dan konsisten. Selain itu, mekanisme distribusi makanan yang merata dan tepat sasaran juga menjadi isu penting.

2.6. *Text Mining*

Text mining adalah proses menganalisis data berbasis teks untuk menemukan pola, makna, dan pengetahuan baru. Informasi diperoleh dengan memprediksi pola dan tren menggunakan model statistik. Teks yang diproses meliputi teks terstruktur dan tidak terstruktur (Pratiwi *et al.*, 2021). Tujuan utama dari *text mining* adalah mengungkap wawasan dan pengetahuan baru yang tersembunyi dalam teks. Fungsi lain *text mining*, dapat mengidentifikasi pola-pola yang tidak terlihat secara kasat mata,

menemukan hubungan antara entitas teks dan memahami sentimen yang terkandung dalam teks. Agar teks dapat diolah secara lebih optimal, dilakukan tahap *preprocessing* untuk mereduksi *noise*, menyetarakan format, dan menyederhanakan struktur data. Menurut (Ma'aly *et al.*, 2024), *preprocessing* berfungsi untuk mengubah teks tidak terstruktur menjadi lebih terorganisasi sehingga memudahkan algoritme klasifikasi maupun analisis sentimen. Beberapa tahapan utama *preprocessing* dalam *text mining* antara lain sebagai berikut:

A. *Cleaning*

Cleaning adalah proses menghapus komponen teks yang tidak diperlukan atau bersifat *noise*. Karakter unik seperti tanda baca titik (.), koma (,), tanda tanya (?), tandaseru (!), tanda petik (‘), dan sebagainya, serta angka numerik (0-9) dan simbol lainnya (*,#,%) dihilangkan (Kurniawan & Mustikasari, 2021).

B. *Case Folding*

Case folding adalah cara untuk menyetarakan huruf secara merata pada sebuah dokumen dengan mengubah semua huruf kapital menjadi huruf kecil. Penyetarakan huruf dengan *case folding* membantu algoritma pengolahan bahasa alami mengenali kata dengan konsisten tanpa dipengaruhi variasi penulisan huruf besar maupun kecil (Saeed *et al.*, 2021).

C. *Normalization*

Normalization adalah proses menyetarakan kata yang memiliki variasi penulisan agar menjadi bentuk standar. Contoh dalam bahasa Indonesia adalah mengubah kata tidak baku seperti “gak”, “ga”, atau “nggak” menjadi “tidak”, serta “yg” menjadi “yang”. Penelitian (Bustamin *et al.*, 2025) menunjukkan bahwa normalisasi *slang* pada teks *Twitter* berbahasa Indonesia dapat meningkatkan akurasi analisis sentimen sebesar 3,47% dibandingkan data yang tidak dinormalisasi. Dengan demikian, normalisasi berperan penting dalam mengurangi keragaman kata yang sebenarnya memiliki makna serupa.

D. *Tokenizing*

Tokenizing adalah proses memecah teks menjadi unit-unit kecil yang disebut token. Pada tahap ini, kalimat atau paragraf diurai menjadi potongan-potongan kata, frasa, atau simbol yang lebih sederhana (Kurniawan & Mustikasari, 2021).

2.7. **IndoBERT**

IndoBERT merupakan model yang didasarkan pada *Bidirectional Encoder Representations from Transformers* (BERT) yang dirancang khusus untuk bahasa Indonesia. IndoBERT diperkenalkan oleh penelitian (Wilie *et al.*, 2024) yang merupakan model yang menjalani pelatihan dengan menggunakan korpus besar bahasa Indonesia sehingga memudahkan dalam memahami struktur dan konteks bahasa lebih baik dibandingkan dengan metode yang lebih konvensional. Melalui bantuan arsitektur *Transformer*, IndoBERT telah menunjukkan kinerja yang sangat baik dalam berbagai tugas pemrosesan bahasa alami (NLP) untuk bahasa Indonesia, termasuk analisis sentimen dan pengenalan emosi (Christian *et al.*, 2025).

Dalam penelitian ini, IndoBERT tidak digunakan sebagai model klasifikasi akhir, melainkan berfungsi sebagai *teacher model* untuk memberi label otomatis (*pseudo-labeling*) pada data yang belum dilabeli secara manual. Pendekatan ini memungkinkan beberapa data tetap dilabeli dengan cara manual sebagai *gold label*, sementara sisa data diprediksi oleh IndoBERT sebagai *silver label*. Oleh karena itu, dalam konteks penelitian ini IndoBERT berfungsi sebagai penyedia pengetahuan (*knowledge provider*) pada tahap pelabelan, bukan sebagai model prediksi akhir.

2.8. *Adaptive Synthetic Sampling Approach*

Ketidakeimbangan kelas adalah masalah yang sering muncul dalam klasifikasi data, termasuk dalam analisis teks dan sentimen. Situasi ini terjadi ketika satu kelas memiliki jumlah data yang jauh lebih besar

daripada kelas lainnya, sehingga model cenderung berfokus pada pola kelas mayoritas dan tidak mengenali karakteristik kelas minoritas. Untuk mengatasi masalah ini, berbagai metode oversampling telah dikembangkan, salah satunya ADASYN (*Adaptive Synthetic Sampling Approach*) yang pertama kali diperkenalkan oleh (He *et al.*, 2008).



Gambar 1. Distribusi ADASYN (Alhudhaif, 2021).

Secara konseptual, ADASYN merupakan pengembangan lebih lanjut dari algoritma SMOTE (*Synthetic Minority Oversampling Technique*) dengan perbedaan pada mekanisme pembuatan data sintetis. SMOTE mendistribusikan data baru secara merata ke seluruh kelas minoritas, sedangkan ADASYN melakukannya secara adaptif, yaitu berdasarkan tingkat kesulitan setiap sampel untuk dibedakan dari kelas mayoritas.

Proses ADASYN dimulai dengan menghitung rasio ketidakseimbangan antar tiap kelas sentimen yang diikuti dengan algoritma *K-Nearest Neighbors* (KNN) untuk mengevaluasi tingkat kesulitan setiap data minoritas. Sampel yang berada di dekat batas keputusan (*decision boundary*) atau memiliki lebih banyak tetangga dari kelas mayoritas akan mendapatkan bobot kesulitan yang lebih tinggi. Secara matematis, tingkat kesulitan tiap sampel minoritas x_i , dinyatakan pada Persamaan 1 dan 2:

$$r_i = \frac{\Delta_i}{k} \quad (1)$$

dengan Δ_i menunjukkan jumlah tetangga dari kelas mayoritas, dan k adalah jumlah total tetangga yang dipertimbangkan (Talukder *et al.*, 2025). Nilai r_i digunakan untuk menentukan jumlah data sintetis yang akan dibuat untuk tiap sampel:

$$G_i = r_i \cdot G \quad (2)$$

dengan G merupakan total jumlah data sintetis yang dihasilkan untuk seluruh kelas minoritas (Talukder *et al.*, 2025).

Dengan pendekatan ini, ADASYN menekankan pembangkitan data sintetis di area yang lebih kompleks, sehingga membantu model klasifikasi untuk belajar lebih baik pada wilayah *decision boundary* dan mengurangi bias terhadap kelas mayoritas.

2.9. *Stratified k-fold Cross Validation*

Stratified k-Fold Cross Validation (SKCV) adalah *varian k-Fold Cross Validation* yang memastikan proporsi setiap kelas di setiap *fold* sesuai dengan distribusi kelas dalam *dataset* asli. Daripada membagi data secara acak, SKCV mengelompokkan contoh berdasarkan label. Kemudian membagi setiap kelompok menjadi k bagian dan menggabungkannya sehingga setiap *fold* mempertahankan komposisi kelas yang serupa. Pendekatan ini cocok untuk data yang tidak seimbang karena mencegah pembentukan *fold* yang hampir tidak memiliki sampel kelas minoritas sehingga menghasilkan estimasi kinerja model yang lebih stabil dan tidak terlalu optimis.

Penelitian (Szeghalmy & Fazekas, 2023) melakukan studi komparatif antara *stratified cross-validation* dan variasi *distribution-balanced stratified cross-validation* pada ratusan *dataset* yang tidak seimbang. Mereka menemukan bahwa memilih metode validasi berstrata sangat penting untuk mencegah hasil yang menyesatkan akibat pergeseran data.

Dengan melatih dan menguji model pada k kombinasi data yang terstratifikasi, peneliti memperoleh rata-rata metrik seperti *F1-score* dan AUC yang lebih representatif karena tiap *fold* memuat proporsi kelas minoritas yang wajar.

2.10. *Recurrent Neural Network*

Recurrent Neural Network (RNN) adalah keluarga jaringan saraf (*neural network*) yang dirancang untuk memodelkan data berurutan dengan memanfaatkan keadaan tersembunyi (*hidden state*) yang diperbarui secara rekursif pada tiap langkah waktu. Mekanisme rekursif ini memungkinkan RNN menyimpan informasi konteks jangka pendek dan mengekspresikan ketergantungan temporal sehingga banyak dipakai pada pemrosesan bahasa alami, pengenalan wicara, dan deret waktu. Tinjauan komprehensif menegaskan bahwa RNN tetap menjadi fondasi penting untuk pemodelan sekuens meskipun arsitektur modern semakin beragam (Mienye *et al.*, 2024).

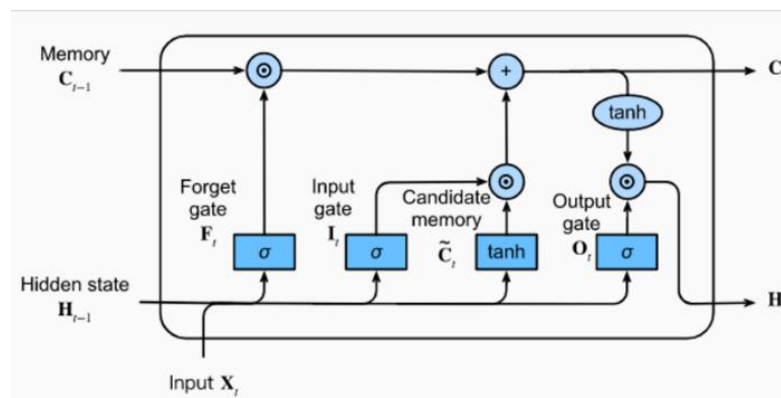
2.11. *Optuna Bayesian Optimization*

Optuna adalah pustaka untuk *hyperparameter tuning* untuk mencari konfigurasi model terbaik berdasarkan nilai metrik evaluasi yang telah ditentukan. Pada penelitian ini, optuna dilakukan dengan pendekatan *Bayesian Optimization* yang populer melalui TPE (*Tree-structured Parzen Estimation*) (Yadav, 2025). Pada optuna, proses pencarian disusun dalam sebuah *study*, yaitu keseluruhan sesi eksperimen yang terdiri dari banyak *trial*. Pada tiap *trial* mengusulkan satu set hyperparameter, kemudian model dilatih menggunakan set tersebut dan menghasilkan skor evaluasi. Skor ini disimpan sebagai riwayat sehingga percobaan berikutnya dapat dipilih secara adaptif, bukan sekedar mencoba kombinasi acak. Selain itu, optuna menyediakan mekanisme *pruning* untuk menghentikan *trial* yang sejak awal menunjukkan performa buruk, sehingga waktu dan sumber daya komputasi dapat digunakan lebih efisien.

Pemilihan nilai *hyperparameter* pada tiap trial dipilih oleh *sampler*, secara default optuna menggunakan sampler berbasis TPE yang belajar dari hasil trial sebelumnya untuk memfokuskan pencarian pada wilayah *hyperparameter* yang berpeluang menghasilkan performa lebih baik (Yadav, 2025). Secara konsep, TPE sampler memiliki keterkaitan dengan teorema Bayes, namun tetap berbeda dengan pendekatan *Gaussian Process* yang langsung memodelkan $p(y|x)$, TPE memodelkan $p(x|y)$ dan membangun dua model probabilistik terhadap nilai *hyperparameter*, dimana $l(x)$ untuk konfigurasi yang hasilnya bagus, dan $g(x)$ untuk sisanya. Kandidat *hyperparameter* berikutnya dipilih dengan prinsip memaksimalkan rasio $\frac{l(x)}{g(x)}$, sehingga nilai yang lebih sering muncul pada percobaan yang sukses dan jarang muncul pada yang gagal akan lebih diprioritaskan.

2.12. Long Short-Term Memory

Long Short-Term Memory (LSTM) adalah salah satu arsitektur dari *Recurrent Neural Network* (RNN) yang dirancang khusus untuk mengatasi berbagai permasalahan dalam pemrosesan data berurutan, seperti teks atau deret waktu. (Hochreiter & Schmidhuber, 1997) memperkenalkan arsitektur *Long Short-Term Memory* (LSTM) yang dirancang khusus untuk mempertahankan informasi dalam jangka panjang melalui mekanisme *cell state* dan *gating mechanism*.



Gambar 2. Arsitektur *Long Short-Term Memory* (Zhang et al., 2023).

Berdasarkan arsitekturnya, LSTM memiliki struktur yang lebih kompleks dibandingkan RNN sederhana. Gambar 1 menunjukkan arsitektur LSTM dibangun atas tiga komponen utama *gate*, yaitu *forget gate*, *input gate*, dan *output gate*. *Forget gate* bertugas memilih informasi dari memori lama yang perlu dihapus, *input gate* berfungsi menentukan informasi baru yang relevan untuk disimpan dalam *cell state*, sementara *output gate* mengontrol informasi yang akan dihasilkan dan diteruskan ke lapisan berikutnya. Ketiga gerbang ini bekerja dengan menggunakan fungsi aktivasi *sigmoid* (σ) dan hiperbolik *tangent* (\tanh) yang mengatur aliran informasi antara *input* (x_t), *hidden state* sebelumnya (h_{t-1}), dan *cell state* (c_t).

Berikut persamaan matematis yang digunakan untuk menggambarkan cara kerja gerbang-gerbang tersebut, ditunjukkan pada Persamaan 3 sampai 8.

$$\text{Forget Gate: } f_t = \sigma(W_{fh}[h_{t-1}] + W_{fx}[x_t] + b_f) \quad (3)$$

$$\text{Input Gate: } i_t = \sigma(W_{ih}[h_{t-1}] + W_{ix}[x_t] + b_i) \quad (4)$$

$$\text{Candidate Vector: } \hat{C}_t = \tanh(W_{ch}[h_{t-1}] + W_{cx}[x_t] + b_c) \quad (5)$$

$$\text{Cell State Update: } C_t = f_t * C_{t-1} + i_t * \hat{C}_t \quad (6)$$

$$\text{Output Gate: } O_t = \sigma(W_{oh}[h_{t-1}] + W_{ox}[x_t] + b_o) \quad (7)$$

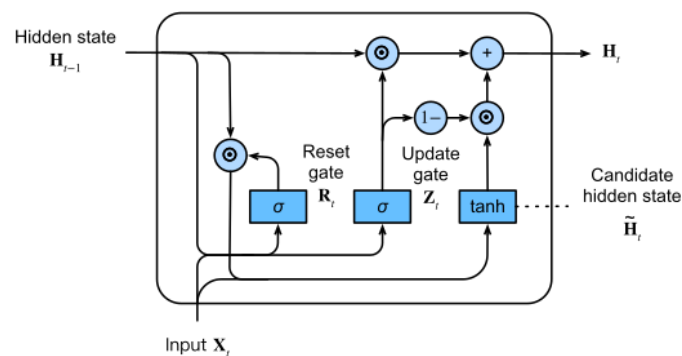
$$\text{Final Output: } h_t = O_t * \tanh(C_t) \quad (8)$$

Keterangan x_t adalah nilai input, W merepresentasikan bobot, b Adalah bias, h_{t-1} adalah *output* dari waktu $t - 1$, C_{t-1} merupakan *memory cell state* pada *cell* sebelumnya, dan h_t adalah *output* akhir (Puteri, 2023).

Dengan mekanisme ini, LSTM dapat menjaga informasi penting dalam rentang waktu panjang sekaligus membuang data yang tidak relevan. Hal tersebut menjadikan LSTM unggul dalam berbagai permasalahan pemrosesan bahasa alami, khususnya analisis sentimen di media sosial (Puteri, 2023).

2.13. Gated Recurrent Unit

Gated Recurrent Unit (GRU) merupakan salah satu cabang *Recurrent Neural Network* (RNN) yang dirancang untuk mengatasi masalah *vanishing gradient* yang menghambat kemampuan model dalam memproses data sekuensial yang panjang, sama seperti halnya *Long Short-Term Memory*. Namun, GRU menawarkan keunggulan berupa arsitektur yang lebih sederhana, yang membuatnya lebih cepat dalam komputasi dan efisien dalam penggunaan memori.



Gambar 3. Arsitektur *Gated Recurrent Unit* (Zhang et al., 2023).

Gambar 3 menunjukkan perbedaan utama antara GRU dan LSTM yang terletak pada jumlah gerbangnya. Jika LSTM memiliki tiga gerbang, GRU hanya memiliki dua gerbang utama yaitu *update gate* dan *reset gate*. Gerbang-gerbang ini bekerja bersama untuk mengontrol seberapa banyak informasi dari masa lalu yang harus dipertahankan dan seberapa banyak informasi baru yang harus dimasukkan.

Berikut persamaan matematis yang menggambarkan cara kerja gerbang tersebut, ditunjukkan Persamaan 9 sampai 12:

$$\text{Reset Gate: } \mathbf{r}_t = \sigma(\mathbf{W}_r \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_r) \quad (9)$$

$$\text{Update Gate: } \mathbf{z}_t = \sigma(\mathbf{W}_z \cdot [\mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_z) \quad (10)$$

$$\text{Candidate Hidden State: } \tilde{\mathbf{h}}_t = \tanh(\mathbf{W} \cdot [\mathbf{r}_t \times \mathbf{h}_{t-1}, \mathbf{x}_t] + \mathbf{b}_h) \quad (11)$$

$$\text{Final Hidden State: } \mathbf{h}_t = (\mathbf{1} - \mathbf{z}_t) \times \mathbf{h}_{t-1} + \mathbf{z}_t \times \hat{\mathbf{h}}_t \quad (12)$$

dengan h_{t-1} merupakan representasi *state*, σ adalah lapisan gerbang *sigmoid*, dan *tan h* merupakan lapisan jaringan fungsi aktivasi *tan h*, serta W_h, W_z, W_r merepresentasi matriks bobot (Kawatu, 2019).

Dalam konteks analisis sentimen pada data media sosial, efektivitas GRU telah terbukti secara empiris. Kesederhanaan arsitekturnya, yang tidak mengorbankan performa, menjadikan GRU pilihan yang kuat dan efisien untuk membandingkan kinerja dengan model LSTM dalam penelitian ini.

2.14. *Confusion Matrix*

Confusion matrix merupakan tabel yang menunjukkan perbandingan antara informasi kelas aktual dan kelas prediksi dalam sebuah sistem klasifikasi. Tabel ini digunakan untuk menghitung kinerja proses klasifikasi atau model dengan nilai benar-salah yang telah diketahui (Pratiwi *et al.*, 2021).

Tabel 1. *Confusion Matrix*

<i>Confusion Matrix</i>	Kelas Prediksi	
	Negatif	Positif
Negatif	<i>True Negative (TN)</i>	<i>False Positive (FP)</i>
Positif	<i>False Negative (FN)</i>	<i>True Positive (TP)</i>

Tabel 1 menunjukkan tentang *confusion matrix* yang akan dijelaskan definisinya sebagai berikut.

- 1) *True Negative (TN)* adalah data sentimen negatif yang diprediksi benar sebagai negatif.
- 2) *True Positive (TP)* adalah data sentimen positif yang diprediksi benar sebagai positif.
- 3) *False Negative (FN)* adalah data sentimen positif yang diprediksi salah sebagai negatif.
- 4) *False Positive (FP)* adalah data sentimen negatif yang diprediksi salah sebagai positif.

Untuk klasifikasi teks, akurasi biasanya diuji menggunakan berbagai metrik dari *confusion matrix* sehingga dapat memperoleh nilai akurasi, presisi, *recall*, dan *f1-score* dengan Persamaan 13 sampai 16 berikut.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (13)$$

Persamaan 13 menunjukkan rumus perhitungan dari akurasi. Akurasi merupakan perhitungan semua data yang diklasifikasikan seperti positif maupun negatif dalam pengujiannya. Nilai akurasi didapat dari kedekatan antara nilai prediksi sistem dan prediksi manusia (Sartika & Saluza, 2022).

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (14)$$

Persamaan 14 menunjukkan rumus perhitungan *precision*. *Precision* adalah perbandingan data aktual dengan jumlah data yang diprediksi. Nilai *precision* dihasilkan dari nilai prediksi positif berbanding dengan jumlah nilai yang positif (Sartika & Saluza, 2022).

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (15)$$

Persamaan 15 menunjukkan rumus perhitungan *recall*. *Recall* adalah nilai yang menunjukkan tingkat keberhasilan untuk mengetahui kembali informasi data positif dan negatif yang benar. *Recall* didapat dari total nilai *true* positif dibanding nilai aktual positif (Sartika & Saluza, 2022).

$$F1 - score = 2 \times \frac{precision \times recall}{precision + recall} \quad (16)$$

Persamaan 16 menunjukkan rumus perhitungan *F1-score*. *F1-score* digunakan untuk mengukur keseimbangan antara *precision* dan *recall*, sehingga metrik ini tetap relevan meskipun data bersifat tidak seimbang. Pada kondisi tersebut, *F1-score* memberikan gambaran yang lebih adil tentang performa model dibandingkan akurasi karena mempertimbangkan kesalahan kelas minoritas maupun mayoritas (Sartika & Saluza, 2022).

2.15. Receiver Operating Characteristic (ROC)

Receiver Operating Characteristic (ROC) merupakan metode evaluasi kinerja model klasifikasi yang digunakan untuk menggambarkan kemampuan model dalam membedakan kelas pada berbagai nilai ambang keputusan (*threshold*). ROC memetakan *true positive rate* pada sumbu y dan *false positive rate* pada sumbu x, sehingga terlihat pertukaran antara kemampuan menangkap kelas positif dan risiko salah mengklasifikasikan kelas negatif. *True Positive Rate* (TPR) dikenal sebagai *recall* atau *sensitivity*, menunjukkan proporsi data positif yang berhasil diklasifikasikan dengan benar. Sedangkan *False Positive Rate* (FPR) menunjukkan proporsi data negatif yang keliru diklasifikasikan sebagai positif. ROC umumnya dipakai ketika model menghasilkan probabilitas kontinu, lalu *threshold* diubah-ubah untuk menghitung TPR dan FPR pada setiap titik. Dengan cara ini, ROC membantu membandingkan model tanpa terkunci pada satu *threshold* (Chicco & Jurman, 2023).

Pada penelitian analisis sentimen untuk kelas negatif dan positif, ROC dapat digunakan untuk menilai seberapa baik skor probabilitas dari model yang memisahkan dua kelas tersebut dan membantu memilih *threshold* yang sesuai tujuan. Pada kasus sentimen multi-kelas (positif–netral–negatif), praktik yang sering dipakai adalah memperluas ROC dengan skema *one-vs-rest* (OvR), yaitu menghitung ROC untuk setiap kelas secara bergantian dengan memperlakukan satu kelas sebagai “positif” dan kelas lainnya sebagai “negatif”. Hasil evaluasi dapat diringkas melalui rata-rata seperti *macro average* atau *micro average* sesuai strategi yang digunakan.

2.16. Area Under Curve (AUC)

Area Under Curve (AUC) adalah luas area di bawah kurva ROC yang merangkum kinerja model di seluruh *threshold* menjadi satu nilai. Secara interpretatif, AUC merepresentasikan peluang bahwa model memberi skor lebih tinggi pada contoh positif acak dibanding contoh negatif acak. Nilai

AUC berada pada rentang 0.5 sampai 1.0, pada 0.5 setara tebakan acak dan 1.0 menunjukkan pemisahan sempurna. Dalam penelitian analisis sentimen, AUC dapat digunakan sebagai ringkasan tambahan untuk menggambarkan kualitas pemisahan kelas berdasarkan skor prediksi model. Untuk skenario multi-kelas, AUC umumnya dihitung melalui pendekatan seperti *one-vs-rest* dan dapat diringkas menggunakan rata-rata sesuai kebutuhan pelaporan. Pada data yang distribusi kelasnya tidak seimbang, AUC tetap bermanfaat sebagai indikator kemampuan diskriminatif, namun sebaiknya tetap disertai metrik seperti *precision*, *recall*, dan *F1-score* agar gambaran kinerja model lebih lengkap dan mudah diinterpretasikan dalam konteks sentimen (Chicco & Jurman, 2023).

2.17. *Functional Programming*

Paradigma pemrograman fungsional (*functional programming*) adalah cara berpikir dan menulis program yang menempatkan fungsi sebagai unit utama penyusun program. Pada pendekatan ini, proses komputasi dipandang sebagai penerapan fungsi-fungsi yang mengolah data, sehingga program terasa seperti rangkaian transformasi yang rapi dan terstruktur (Dyer, 2022). Dalam praktik modern, pemrograman fungsional sering hadir di bahasa yang bersifat *multi-paradigm* (tidak murni fungsional), dan banyak digunakan karena cocok dengan kebutuhan pengembangan perangkat lunak saat ini seperti modularitas, kemudahan pengujian, dan pengelolaan sistem komputasi yang kompleks.

Bahasa pemrograman multiparadigma adalah bahasa yang memungkinkan gaya pemrograman berbeda digunakan dalam bahasa yang sama. Beberapa contoh bahasa pemrograman multiparadigma adalah Python, JavaScript, Java, dan C++ (Vidal, 2023). Bahasa-bahasa ini menawarkan fitur dan sintaks yang menerapkan paradigma pemrograman yang berbeda. Hal ini memungkinkan mereka untuk memanfaatkan kekuatan masing-masing paradigma dan menggunakan alat dan *library* untuk setiap gaya pemrograman. Kemampuan untuk menggabungkan paradigma yang

berbeda dalam satu bahasa memberikan fleksibilitas dan efisiensi dalam pengembangan

2.18. Python

Python dikenal sebagai bahasa pemrograman multi-paradigma, artinya Python memungkinkan penulisan program dengan gaya imperatif, berorientasi objek, dan juga fungsional dalam satu bahasa (Thaker & Shukla, 2020). Dalam praktik *data science* dan *machine learning*, Python banyak digunakan karena sintaksnya ringkas serta ekosistem pustaka ilmiah yang matang. Pada penelitian ini, Python berperan sebagai *environment* untuk membangun *pipeline* analisis sentimen. *Library* digunakan sebagai implementasi teknis untuk merealisasikan *pipeline* secara konsisten.

2.18.1. NumPy dan Pandas

Library NumPy digunakan sebagai dasar komputasi numerik berbasis array. Dalam penelitian ini, NumPy membantu pengolahan data numerik seperti manipulasi label, operasi array pada data hasil tokenisasi/padding, serta perhitungan agregasi sederhana yang dibutuhkan pada pelatihan dan evaluasi (Harris et al., 2020). *Library* pandas digunakan untuk pengolahan data tabular, khususnya membaca *dataset* CSV, pembersihan kolom, seleksi data, serta penyimpanan hasil ke *file* CSV (W3Schools, 2026).

2.18.2. Scikit-learn dan Imbalanced-learn

Library scikit-learn dipakai pada tahapan persiapan data, validasi, ekstraksi fitur seperti n-gram, dan evaluasi model. Pada tahapan persiapan data digunakan untuk encoding label (LabelEncoder), pada validasi digunakan StratifiedKFold, lalu pada tahap evaluasi untuk menghitung metrik performa, dan untuk n-gram digunakan

CountVectorizer yang mengubah kumpulan dokumen teks menjadi matriks jumlah kemunculan token (Scikit-learn developers, 2025). *Library* imbalanced-learn digunakan untuk menangani masalah ketidakseimbangan kelas yang umum terjadi pada data sentimen media sosial. Pada penelitian ini, teknik over-sampling seperti ADASYN digunakan untuk menambah sampel kelas minoritas secara adaptif. Imbalanced-learn memang dikembangkan sebagai *toolbox* khusus untuk mengatasi *imbalanced datasets* dan kompatibel dengan alur kerja scikit-learn (Lemaitre et al., 2017).

2.18.3. Wordcloud dan Matplotlib

Library wordcloud digunakan untuk memvisualisasikan kata yang sering muncul pada data teks dalam bentuk “awan kata”. Pada penelitian ini, WordCloud menggambarkan kecenderungan kata dominan pada sentimen tertentu (Python Software Foundation, 2026). *Library matplotlib* digunakan untuk membuat visualisasi hasil penelitian, misalnya grafik *learning curve (loss/accuracy)*, *confusion matrix*, *ROC curve*, grafik tren temporal/spasial, dan plot pendukung lain yang disimpan ke folder hasil. Visualisasi ini membantu pembaca memahami performa model dan pola data secara lebih intuitif (Hunter, 2007).

2.18.4. TensorFlow dan Transformers

Library TensorFlow digunakan (melalui API Keras) untuk membangun dan melatih model *deep learning* berbasis LSTM dan GRU. Pada penelitian ini, tensorflow digunakan pada tahap representasi teks dan pembangunan model serta *callback* (TensorFlow, 2026). *Library transformers* digunakan untuk tahapan pelabelan menggunakan IndoBERT (Wolf et al., 2020).

2.19. Penelitian Terdahulu

Berdasarkan kajian literatur sebelumnya, disusun Tabel 2 yang menunjukkan ringkasan penelitian terdahulu yang relevan.

Tabel 2. Penelitian Terdahulu

No	Komponen Penelitian Terdahulu				
	Judul Penelitian	Penulis	Metode	Data	Hasil
1	Analisis Program Makan Bergizi Gratis dengan <i>Support Vector Machine</i> (SVM) pada Aplikasi X.	(Amaria & Hidayati, 2025)	<i>Support Vector Machine</i> (SVM) dengan ekstraksi fitur <i>Term Presence, Bag of Words</i> , dan TF-IDF.	Berjumlah 1.378 <i>tweet</i> terkait Program Makan Bergizi Gratis di <i>platform</i> X, diproses menjadi 1.151 <i>data</i> setelah <i>preprocessing</i> .	Penelitian ini menggunakan model SVM dengan tiga jenis ekstraksi fitur, dan hasil terbaik diperoleh pada kombinasi SVM dengan TF-IDF dengan akurasi 77,5%, presisi 77,51%, <i>recall</i> 90,34%, dan <i>F1-score</i> 83,44%. Uji lanjut menggunakan validasi silang <i>5-Fold</i> yang dipadukan dengan SMOTE menghasilkan akurasi rata-rata 75,54% dengan performa yang tetap konsisten. Temuan ini menunjukkan bahwa TF-IDF lebih unggul dibandingkan metode lain karena mampu memberikan representasi kata yang lebih baik serta stabil dalam menghadapi variasi data.

Tabel 2. Lanjutan

No	Komponen Penelitian Terdahulu				
	Judul Penelitian	Penulis	Metode	Data	Hasil
2	<i>Sentiment Perspective of Government's Free Nutritious Meal Policy on Social Media X using Indo-BERT and Bi-LTSM.</i>	(Subarkah <i>et al.</i> , 2025)	Algoritma Indo-BERT dan Bi-LSTM.	Pengambilan data dari media sosial X yang berjumlah 3.319 data dengan sentimen positif berjumlah 2.513, sentimen negatif berjumlah 716, dan netral berjumlah 71.	Penelitian ini menggunakan algoritma Indo-BERT yang menghasilkan akurasi sebesar 80% dan dengan algoritma Bi-LSTM menghasilkan akurasi sebesar 78%.
3	Analisis Sentimen Program Makan Bergizi Gratis Menggunakan <i>Bidirectional Gated Recurrent Unit.</i>	(Krisnawan <i>et al.</i> , 2025)	Algoritma <i>Bidirectional Gated Recurrent Unit</i> (Bi-GRU).	Pengambilan data dari media sosial X yang berjumlah 1.405 data dengan 798 data negatif dan 607 data positif.	Hasil klasifikasi sentimen menggunakan algoritma Bi-GRU dengan bantuan FastText menghasilkan akurasi sebesar 80%, presisi 80%, <i>recall</i> 80%, dan <i>f1-score</i> 80%.

Tabel 2. Lanjutan

No	Komponen Penelitian Terdahulu				
	Judul Penelitian	Penulis	Metode	Data	Hasil
4	Analisis Sentimen Program Makan Gratis Pada Platform X Menggunakan Algoritma <i>Naïve Bayes</i>	(Laila <i>et al.</i> , 2025)	Algoritma <i>Naïve Bayes</i> dengan pelabelan semi-otomatis berbasis IndoBERT.	Pengambilan data menggunakan 3.600 <i>tweet</i> berbahasa Indonesia dari Platform X pada periode Februari 2024 – Juni 2025.	Penelitian ini menggunakan metode <i>Naïve Bayes</i> menghasilkan akurasi 86,46%, presisi 86,55%, <i>recall</i> 95,25%, dan <i>F1-score</i> 90,77%. Hasil <i>10-Fold Cross Validation</i> menunjukkan rata-rata akurasi 86,74%. Penelitian membuktikan <i>Naïve Bayes</i> stabil dan efisien untuk klasifikasi teks pendek seperti <i>tweet</i> .
5	<i>An Integrated Random Forest for Analyzing Public Sentiment on the “Makan Bergizi Gratis” Program</i>	(Ramadhan i & Khoirunnisa, 2025)	Algoritma <i>Random Forest</i> , dibandingkan dengan <i>AdaBoost</i> , <i>XGBoost</i> , <i>Bagging</i> , dan <i>Stacking</i> , menggunakan TF-IDF.	Dataset berjumlah 3.417 <i>tweet</i> dengan sentimen negatif dan positif dikumpulkan dari platform X pada periode Januari – April 2025.	Penelitian ini menerapkan metode <i>Random Forest</i> tanpa <i>stemming</i> memberikan hasil terbaik: akurasi 91,15% dan ROC-AUC 95,46%. Visualisasi menunjukkan kata dominan pada sentimen positif (gizi, makan, gratis, sehat) dan negatif (tidak, kerja, pemerintah, rakyat). Studi ini menegaskan efektivitas <i>ensemble learning</i> untuk analisis opini publik kebijakan sosial.

Tabel 2. Lanjutan

No	Komponen Penelitian Terdahulu				
	Judul Penelitian	Penulis	Metode	Data	Hasil
6	Analisis Sentimen <i>Tweet</i> COVID-19 menggunakan <i>Word Embedding</i> dan Metode <i>Long Short-Term Memory</i>	(Rahman <i>et al.</i> , 2021)	Model <i>deep learning Long Short-Term Memory</i> dengan integrasi <i>word embedding</i>	Pengambilan data menggunakan data Twitter 1364 <i>tweet</i> berbahasa Indonesia yang diberikan sentimen positif, negatif, dan netral yang di- <i>crawling</i> secara mandiri.	Penelitian ini menggunakan metode LSTM dengan <i>word embedding</i> untuk analisis sentimen <i>tweet</i> terkait COVID-19 memperoleh akurasi 81%, presisi 80%, <i>recall</i> 80%, dan F1 81%. Jika dibandingkan dengan <i>Naïve Bayes</i> (akurasi 74%) dan RNN (71%), maka LSTM terbukti lebih unggul dengan selisih 7–10%, sehingga dinilai paling efektif dalam mengklasifikasikan sentimen publik terhadap kebijakan karantina.
7	<i>The Application of Gated Recurrent Unit in Social Media Sentiment Analysis.</i>	(Kawatu, 2019)	Algoritma <i>Gated Recurrent Unit</i> (GRU).	Pengambilan <i>tweet</i> berjumlah 2677 opini pelanggan tentang layanan ekspedisi barang di Twitter.	Penelitian ini menerapkan metode GRU pada 2.677 <i>tweet</i> terkait layanan ekspedisi. Model ini menghasilkan akurasi sebesar 70%, dengan nilai presisi, <i>recall</i> , dan F1 bervariasi: sentimen netral (0,63; 0,79; 0,70), positif (0,74; 0,60; 0,67), dan negatif (0,75; 0,69; 0,72). Temuan ini menunjukkan bahwa GRU cukup efektif terutama dalam mendeteksi sentimen negatif.

3.2. Analisis Kebutuhan Sistem

Perangkat penelitian merupakan alat pendukung yang digunakan dalam penelitian ini adalah sebagai berikut.

3.2.1 Perangkat Keras (*Hardware*)

- a. Perangkat : ASUS VivoBook X513EAN_K513EA
- b. Prosesor : Intel Core i5
- c. Memori (RAM) : 8 GB
- d. Penyimpanan : SSD 512 GB

3.2.2 Perangkat *Cloud*

- a. Platform Google Colab

Google Colab digunakan untuk melakukan analisis sentimen dengan menjalankan *script* Python untuk pengumpulan data, pra-pemrosesan, visualisasi data, pelabelan data, dan pemodelan data dengan spesifikasi dibawah ini.

- i. Akselerator : GPU T4
- ii. Virtual RAM :15 GB
- iii. RAM Sistem : 12.7 GB
- iv. Virtual Disk : 112.6 GB

- b. Platform Kaggle

Kaggle digunakan untuk menjalankan *script* bagian *hypertuning* dengan akselerator GPU P100 untuk mengefesiensikan waktu komputasi.

3.2.3 Perangkat Lunak (*Software*)

- a. Windows 11 Pro 64-bit

Sistem operasi yang mendasari komputer, memastikan bahwa semua perangkat lunak dapat berjalan dengan baik di dalamnya.

b. Microsoft Excel

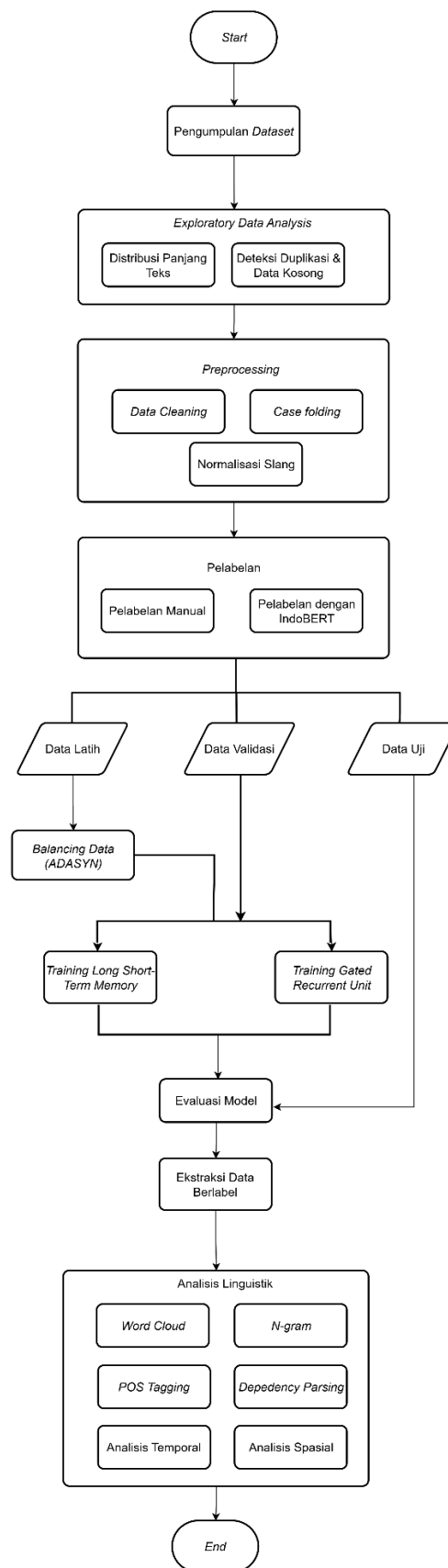
Microsoft Excel digunakan untuk melakukan pelabelan data secara manual dan *review* data.

c. *Web browser* Google Chrome

Web browser seperti Google Chrome digunakan untuk mengakses sumber data *online*, melakukan penelusuran, dan mengakses *platform-platform* yang relevan dengan data yang akan dianalisis.

3.3. Tahapan Penelitian

Pada penelitian ini, terdapat beberapa tahapan utama yang dilakukan untuk menganalisis sentimen *tweets* terhadap program makan bergizi gratis. Setiap tahapan berperan penting dalam memastikan keakuratan hasil analisis sentimen dan analisis linguistik. Rangkaian studi ini dilaksanakan sesuai dengan tahapan yang tertera pada Gambar 4. Tahap pertama mencakup pengumpulan informasi berupa *tweet* dari *platform* media sosial X (*Twitter*) dengan menggunakan kata kunci yang relevan dengan program "Makan Bergizi Gratis" lalu difilter hanya dalam cakupan SMA. Selanjutnya data melalui tahap *Exploratory Data Analysis*, lalu dilanjutkan tahap *preprocessing*. *Dataset* yang telah dibersihkan kemudian dilewati fase pelabelan yang dimulai dengan data yang telah diberikan label secara manual yang kemudian diperbanyak menggunakan *teacher model* IndoBERT untuk menambah jumlah data yang berlabel. *Dataset* tersebut kemudian dibagi menjadi dua bagian yaitu data latih, data uji, dan data validasi. Pada pembagian data ini digunakan metode *Stratified K-Fold Cross Validation*. Data latih disesuaikan menggunakan metode *ADASYN* untuk mencapai distribusi kelas yang lebih seimbang. Selanjutnya, pelatihan model menggunakan dua algoritma LSTM dan GRU serta dilanjutkan dengan evaluasi model pada data uji.



Gambar 4. Tahapan Penelitian.

3.3.1. Studi Literatur

Langkah awal dalam penelitian ini adalah melakukan kajian pustaka seperti mengumpulkan informasi dari jurnal, buku, dan artikel yang berkaitan dengan topik penelitian. Selanjutnya, teori-teori yang berhubungan dengan penelitian ini dikembangkan, termasuk definisi, penjelasan, dan kelebihan serta kelemahan dari setiap metode yang diterapkan, khususnya *Long Short-Term Memory* (LSTM) dan *Gated Recurrent Unit* (GRU).

3.3.2. Pengumpulan Data

Pengumpulan data dilakukan dengan memanfaatkan *tools tweet-harvest* yang dikembangkan oleh (Satria, 2024). Alasan pemilihan *tools* ini karena *tweet harvest* termasuk salah satu *tools* yang mudah digunakan untuk mengumpulkan data. Pengguna hanya perlu memasukkan *auth token* dari akun *Twitter* pribadi untuk melakukan pengumpulan, dan pemilihan data dengan kurun waktu yang peneliti tentukan. Data yang diambil berupa informasi *tweet* seperti tanggal *tweet* dibuat, isi *tweet*, jumlah favorit, *image url*, bahasa, lokasi, jumlah *quote*, jumlah *reply*, *tweet url*, *user id*. Pada penelitian ini, saya melakukan *crawling* data dengan kata kunci “makan bergizi gratis”, “MBG”, “program gizi siswa”, dan “program makan Prabowo”. Data yang terkumpul kemudian difilter dengan kata kunci yang memfokuskan case study di SMA. Pengumpulan data dilakukan dari periode Januari 2025 – Agustus 2025 yang berjumlah 13.486 *tweet* dengan variabel *created_at*, *full_text*, dan label dengan format CSV.

3.3.3. Exploratory Data Analysis

Tujuan utama tahap ini adalah menilai kualitas data mentah, memahami karakteristik dasar teks, serta mengidentifikasi masalah potensial seperti duplikasi dan nilai kosong yang dapat

memengaruhi tahapan berikutnya. Langkah pertama EDA adalah pemeriksaan struktur data. Berkas kumpulan *tweet* dibaca ke dalam *dataframe* kemudian diperiksa jumlah baris, nama kolom, serta tipe data pada setiap kolom, misalnya teks *tweet*, tanggal unggah, dan informasi lain yang tersedia. Pada tahap ini juga dihitung jumlah nilai kosong pada tiap kolom untuk menentukan apakah ada baris yang perlu dibuang atau dilengkapi sebelum diproses lebih lanjut.

Selanjutnya dilakukan analisis panjang teks. Untuk setiap *tweet* dihitung panjang dalam satuan karakter dan jumlah kata. Nilai ini kemudian diringkas dalam bentuk statistik deskriptif (nilai minimum, maksimum, rata-rata, dan kuartil) dan divisualisasikan menggunakan histogram. Analisis ini digunakan untuk melihat sebaran panjang *tweet*, mendeteksi keberadaan teks yang sangat pendek atau sangat panjang, serta menjadi dasar penentuan batas panjang urutan ketika data akan dipersiapkan sebagai masukan model.

Tahap berikutnya adalah deteksi duplikasi. Teks *tweet* terlebih dahulu dinormalkan secara sederhana, misalnya dengan mengubah huruf menjadi bentuk kecil (*lowercase*) dan menghapus spasi ganda di awal maupun akhir. Setelah itu dilakukan pencarian *tweet* dengan isi yang persis sama melalui pengelompokan berdasarkan teks. Dari hasil ini dihitung jumlah dan proporsi *tweet* duplikat terhadap keseluruhan data. *Tweet* yang teridentifikasi sebagai duplikat kemudian dihapus sehingga model nantinya tidak belajar dari contoh yang sama berulang kali.

3.3.4. Data Preprocessing

Preprocessing data adalah tahapan penting sebelum proses pelabelan dan pemodelan. Tujuannya untuk memastikan bahwa teks yang akan dijadikan masukan bagi model bersih, konsisten, dan tidak mengandung duplikat yang dapat mempengaruhi efektivitas

algoritma. Berdasarkan penelitian (Ma'aly *et al.*, 2024), kualitas *preprocessing* berpengaruh besar terhadap kinerja model khususnya dalam analisis sentimen pada teks media sosial. Pada penelitian ini, *preprocessing* dilakukan secara berurutan sebagai berikut:

a) *Case Folding*

Case folding adalah proses standarisasi seluruh teks diubah menjadi huruf kecil untuk menyamakan bentuk kata. Contohnya, kata "Gratis" dan "gratis" dianggap sebagai satu sehingga model tidak membedakan keduanya.

b) *Cleaning*

Cleaning adalah proses pembersihan dilakukan untuk menghilangkan unsur yang tidak relevan seperti URL, penyebutan (@namapengguna), tagar, dan simbol non-alfabet. Langkah ini bertujuan untuk mengurangi gangguan dalam data teks. Menurut (Ma'aly *et al.*, 2024), penghapusan elemen non-linguistik dapat memperbaiki kualitas representasi teks dalam tugas NLP Bahasa Indonesia.

c) *Normalization*

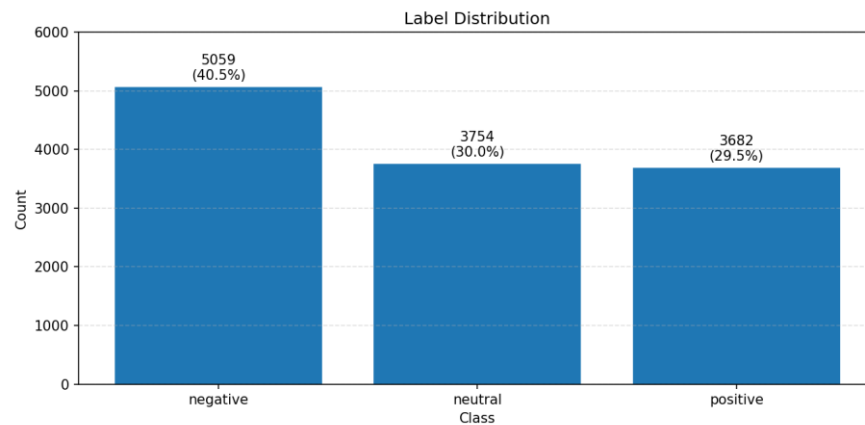
Normalization adalah proses mengubah *tweet* yang menggunakan bahasa gaul atau singkatan seperti "gak", "yg", "tp", dan "gw". Oleh sebab itu, dilakukan normalisasi dengan menggunakan kamus yang menerjemahkan bahasa gaul ke dalam kata standar seperti "gak" menjadi "tidak," "yg" menjadi "yang," dan "skrg" menjadi "sekarang". Normalisasi ini berfungsi untuk mengurangi variasi bentuk kata yang memiliki makna yang sama. Kamus bahasa untuk proses normalisasi diambil dari penelitian (Ibrohim & Budi, 2018) dan penelitian (Salsabila *et al.*, 2018) yang mencakup banyak *slang* dalam bahasa Indonesia.

Hasil akhir dari *preprocessing* adalah *dataset* yang bersih, teratur, dan representatif yang dapat digunakan untuk pelabelan dengan IndoBERT serta untuk pemodelan LSTM dan GRU.

3.3.5. Pelabelan

Data bersih yang sudah melalui tahapan *preprocessing* sebanyak 13.486 data teks. Tahap pelabelan memiliki fungsi penting dalam penelitian ini karena kualitas label yang dihasilkan berdampak pada kinerja model analisis sentimen. Proses pelabelan dilakukan melalui dua metode, yaitu anotasi manual dalam jumlah terbatas serta pelabelan otomatis dengan bantuan *teacher model* berbasis IndoBERT. Pada pelabelan manual dilakukan pada 1000 data dan dilabeli dalam tiga kelas sentimen sebagai data acuan (*gold label*). Pelabelan manual dilakukan menggunakan pedoman anotasi yang mengacu pada kamus sentimen Bahasa Indonesia dari penelitian terdahulu milik (Abidin, 2024; Wahid & SN, 2016). Penentuan label manual didasarkan pada penilaian dominan penulis dengan indikator positif berupa dukungan ataupun apresiasi, indikator negatif berupa kritik tidakpuasan dan keluhan, serta indikator netral untuk pernyataan informatif ataupun berita.

Penelitian (Chen *et al.*, 2020) menegaskan bahwa dalam kerangka *weakly supervised learning*, sedikit label manual dapat membantu model dalam menganotasi data lainnya. Proses pelabelan tidak perlu sepenuhnya dilakukan oleh manusia yang memakan banyak waktu dan biaya. Setelah pelabelan manual selesai, data yang belum terlabel diproses dengan model *indobenchmark/indobert-large-pl* dari *Hugging Face*. Ini adalah versi IndoBERT yang dikembangkan sedari proyek IndoNLU dan dilatih di korpus Indo4B yang merupakan kumpulan data berisi 4 miliar kata dalam bahasa Indonesia dari berbagai sumber. Jumlah keseluruhan data yang telah diberikan label sentimen dapat dilihat pada Gambar 4.



Gambar 5. Visualisasi Hasil Pelabelan.

Hasil dari pemberian label gabungan antara manual dan IndoBERT menghasilkan 5039 opini negatif, 3754 opini netral, dan 3682 opini positif. *Dataset final* inilah yang selanjutnya digunakan dalam proses pembagian data dan pelatihan model klasifikasi berbasis LSTM serta GRU.

3.3.6. Pembagian Data

Setelah seluruh data teks selesai diberi label sentimen dan melalui tahapan *pre-processing*, dilakukan pembagian data untuk keperluan pelatihan dan evaluasi model. Pada penelitian ini digunakan pendekatan *Stratified K-Fold Cross Validation* (SKCV) dengan nilai $K=10$. Pendekatan ini dipilih agar pemanfaatan data lebih efisien dan evaluasi kinerja model menjadi lebih stabil, mengingat distribusi kelas sentimen yang tidak seimbang.

Dalam *stratified 10-fold cross validation*, keseluruhan korpus yang telah berlabel dibagi menjadi sepuluh *fold* dengan memperhatikan proporsi kelas negatif, netral, dan positif pada setiap *fold*. Pada setiap iterasi, sembilan *fold* digabungkan sebagai data latih dan satu *fold* sisanya digunakan sebagai data uji. Proses ini diulang sebanyak sepuluh kali sehingga setiap *fold* sekali menjadi data uji dan sembilan kali menjadi bagian data latih. Dengan cara ini, semua

data akan berkontribusi baik pada proses pelatihan maupun pengujian, namun tidak pada iterasi yang sama.

Untuk setiap model yang dibangun, pelatihan dilakukan secara terpisah pada masing-masing *fold* menggunakan parameter dan arsitektur yang sama. Metrik evaluasi seperti akurasi, presisi, *recall*, dan *F1-score* dihitung pada data uji di setiap *fold*.

3.3.7. Data Balancing

Hasil analisis distribusi label menunjukkan bahwa kelas negatif jauh lebih banyak daripada kelas netral dan positif. Ketidakseimbangan ini berpotensi membuat model hanya menghafal pola kelas mayoritas dan mengabaikan kelas lainnya. Untuk mengurangi bias tersebut dilakukan *oversampling* dengan *Adaptive Synthetic Sampling Approach* (ADASYN). ADASYN diterapkan pada representasi fitur data latih di setiap *fold*, bukan pada keseluruhan data sehingga tidak terjadi kebocoran informasi ke data uji. Pada setiap iterasi, algoritma ini menghitung derajat kesulitan tiap sampel pada kelas minoritas berdasarkan tetangga terdekatnya, kemudian menghasilkan sampel sintetis baru secara adaptif terutama di area yang sulit dipisahkan dari kelas mayoritas. Proses ini dilakukan hingga distribusi kelas pada data latih di tiap *fold* menjadi lebih seimbang, sementara distribusi data uji tetap dibiarkan apa adanya.

3.3.8. Training Long Short Term Memory (LSTM)

Pada tahap ini, *dataset* akan digunakan untuk melatih model *Long Short-Term Memory* (LSTM) yang berfungsi sebagai model klasifikasi utama. Proses klasifikasi sentimen menggunakan LSTM dimulai dari teks yang telah ditokenisasi, kemudian diubah menjadi representasi numerik menggunakan *embedding layer* bawaan dari model LSTM. Lapisan ini secara otomatis membangun representasi

vektor dari setiap kata berdasarkan proses pembelajaran selama pelatihan. Representasi tersebut selanjutnya diproses oleh lapisan LSTM untuk mempelajari pola hubungan antar kata serta konteks kalimat yang menggambarkan arah sentimen. Representasi akhir dari memori LSTM kemudian diteruskan ke lapisan *output* yang biasanya menggunakan fungsi aktivasi seperti *softmax* untuk menghasilkan distribusi probabilitas bagi setiap kategori sentimen. Dengan cara ini, LSTM dapat membangun pemahaman kontekstual yang kuat terhadap data berlabel dan menggunakannya untuk memprediksi sentimen dalam teks baru yang tidak terdapat dalam data pelatihan.

3.3.9. *Training Gated Recurrent Unit (GRU)*

Proses klasifikasi sentimen menggunakan GRU dimulai dari teks yang telah ditokenisasi. Data teks yang telah melalui proses tokenisasi dan *padding* diubah menjadi urutan angka berukuran seragam, lalu diolah oleh lapisan *embedding* bawaan dari GRU. Lapisan *embedding* ini bersifat *trainable*, artinya representasi vektor dari setiap kata dipelajari langsung selama pelatihan model tanpa menggunakan *pre-trained embedding* eksternal. Pendekatan ini membuat model dapat beradaptasi secara optimal terhadap karakteristik bahasa dan konteks dalam dataset yang telah dilabel.

Selama proses pelatihan, GRU mengidentifikasi hubungan antar kata dalam urutan teks untuk menangkap pola sentimen di setiap kalimat. Struktur jaringan GRU yang lebih sederhana memungkinkan model bekerja dengan waktu pelatihan yang lebih cepat dibandingkan LSTM, namun tetap mampu mengenali memori jangka panjang dalam data sekuensial. *Output* dari lapisan GRU kemudian diteruskan ke *dense layer* dengan fungsi aktivasi *softmax* untuk menghasilkan probabilitas pada masing-masing kategori sentimen, yaitu positif, netral, dan negatif.

3.3.10. Evaluasi Model

Tahap evaluasi dilakukan untuk menilai kinerja model LSTM dan GRU setelah proses pelatihan. Penilaian dilaksanakan pada 20% data uji yang tidak melalui proses *balancing* sehingga hasilnya mencerminkan kemampuan model dalam menghadapi data dari dunia nyata. Evaluasi menggunakan kombinasi metrik berbasis *confusion matrix* dan metrik berbasis kurva agar gambaran kinerja model lebih menyeluruh, terutama saat distribusi kelas tidak seimbang. Hal ini penting karena pada skenario multikelas dengan perbedaan proporsi kelas yang besar, sebagian metrik dapat lebih sensitif terhadap distribusi kelas. Berikut beberapa metrik yang dipakai untuk evaluasi model:

a) Akurasi

Persentase dari prediksi yang benar di seluruh data uji.

b) Presisi

Persentase prediksi positif yang benar dibandingkan dengan semua prediksi positif yang dibuat.

c) *Recall*

Kemampuan model untuk mengidentifikasi data yang benar-benar termasuk dalam suatu kategori.

d) *F1-score*

Rata-rata harmonis antara presisi dan *recall* yang digunakan untuk menilai keseimbangan di antara keduanya, khususnya untuk data yang tidak seimbang.

e) ROC dan AUC (ROC-AUC)

ROC mengevaluasi perubahan kinerja ketika *threshold* bervariasi, sedangkan AUC merangkum kurva ROC menjadi satu nilai yang merepresentasikan kemampuan diskriminasi model. Pada penelitian multikelas, ROC-AUC dihitung menggunakan pendekatan *one-vs-rest* (OvR) untuk setiap kelas,

kemudian dapat diringkas menggunakan *macro-average* agar tiap kelas memperoleh bobot yang setara dalam pelaporan.

Selain metrik kuantitatif, *confusion matrix* juga digunakan untuk menggambarkan distribusi hasil prediksi model pada label sebenarnya sehingga memungkinkan analisis yang lebih mendalam tentang pola kesalahan klasifikasi. Menurut penelitian (Tamami *et al.*, 2025), menekankan pentingnya pemanfaatan metrik seperti akurasi, presisi, *recall*, dan skor F1 untuk memastikan model tidak hanya berkinerja baik dalam kelas mayoritas tetapi juga konsisten dalam kelas minoritas.

3.4. Analisis Linguistik

3.4.1. Analisis Temporal dan Spasial

Analisis temporal digunakan untuk melihat perubahan intensitas dan distribusi sentimen dari waktu ke waktu. Dalam fase ini, setiap *tweet* yang ditandai dengan sentimen ditetapkan ke periode bulanan menggunakan stempel waktu. Untuk setiap bulan, jumlah total *tweet* dan proporsi *tweet* negatif, netral, dan positif dihitung. Hasil perhitungan ini kemudian disajikan sebagai histogram yang menunjukkan jumlah *tweet* bulanan dan distribusi sentimen.

Tujuan analisis ini adalah untuk mengidentifikasi bulan dengan peningkatan aktivitas percakapan dan untuk menyelidiki hal yang berkaitan dengan peningkatan tersebut misalnya, peningkatan sentimen negatif selama periode tertentu. Informasi temporal ini kemudian dihubungkan dengan topik-topik program MBG pada periode tersebut menggunakan analisis n-gram.

Analisis spasial dilakukan untuk meninjau perbedaan sentimen pada pulau-pulau besar di Indonesia. Dikarenakan tidak semua *tweet* berisi informasi lokasi yang eksplisit, hanya *tweet* dengan detail lokasi seperti nama kota, kabupaten, atau provinsi yang

diekstrak dari teks *tweet* untuk digunakan dalam analisis ini. Nama-nama tempat kemudian dipetakan ke beberapa pulau besar di Indonesia seperti Sumatra, Jawa, Kalimantan, Sulawesi, Papua, Kepulauan Maluku, dan Kepulauan Nusa Tenggara.

Setelah pemetaan, jumlah *tweet* negatif, netral, dan positif untuk setiap pulau dihitung, termasuk persentase setiap sentimen. Hasilnya divisualisasikan sebagai histogram sentimen untuk setiap pulau. Analisis ini membantu mengungkap pola, seperti pulau-pulau yang dicirikan oleh sentimen yang lebih positif, lebih negatif, atau relatif seimbang yang memungkinkan korelasi dengan kondisi sosial dan ekonomi atau implementasi program-program tertentu di wilayah tersebut.

3.4.2. N-Gram

Analisis N-gram diterapkan pada teks yang telah diproses sebelumnya meliputi penghapusan URL, penyebutan, tanda baca tertentu, normalisasi huruf besar/kecil, dan penghapusan *stop word*. Stop word yang digunakan diambil dari penelitian milik (Lestari *et al.*, 2017) dan hanya digunakan untuk analisis N-Gram. Setelah langkah ini, unigram dan bigram dihasilkan untuk setiap kategori sentimen, frekuensinya dihitung, dan diurutkan berdasarkan frekuensi. Proses ini dilakukan secara terpisah untuk sentimen negatif, netral, dan positif guna memudahkan identifikasi pola kata umum dalam setiap kategori.

Hasil analisis N-gram tidak digunakan untuk menyempurnakan model, melainkan untuk mengungkap isu utama yang memengaruhi tiap sentimennya. Pada Bab 4, daftar unigram dan bigram dengan frekuensi tertinggi berfungsi sebagai dasar untuk mengidentifikasi tema-tema utama yang memengaruhi sentimen negatif, netral, dan positif terhadap program MBG.

3.4.3. *Part Of Speech Tagging*

Part of Speech (POS) Tagging diimplementasikan menggunakan *library* Stanza Bahasa Indonesia melalui proses yang mencakup tokenisasi, ekspansi multi-*word* token, penandaan kelas kata, dan lemmatisasi. Setiap token dalam sebuah *tweet* ditandai dengan label POS seperti, nomina, verba, adjektiva, atau adverbial. Kemudian digabungkan dengan label sentimen IndoBERT untuk menghasilkan distribusi POS untuk setiap kategori sentimen.

Persentase kemunculan setiap label POS (nomina, verba, adjektiva, dll) dalam setiap kategori sentimen akan dihitung dan kumpulan token dengan frekuensi tertinggi untuk setiap kategori POS diidentifikasi. Daftar kata ini dianalisis lebih lanjut untuk menunjukkan bagaimana nomina, verba, atau adjektiva tertentu mendominasi setiap sentimen dan bagaimana pola-pola ini mencerminkan persepsi pengguna internet terhadap MBG.

3.4.4. *Dependency Parsing*

Dependency Parsing diterapkan pada korpus yang sama menggunakan Stanza untuk menentukan struktur dependensi antar elemen kalimat. Berdasarkan hasil analisis, didapatkan sejumlah relasi, seperti *nsubj* (subjek nomina), *obj* (objek), dan *root* (kata kerja utama). Informasi ini digabungkan dengan label sentimen untuk menghitung subjek, predikat, dan objek yang paling sering muncul dalam setiap kategori sentimen. Selain tingkat token, hasil *dependency parsing* juga digunakan untuk menyusun frasa subjek maupun objek dengan mengelompokkan token yang berkaitan, seperti nomina dengan kata sifat ke dalam unit-unit seperti "program makan bergizi gratis" atau "MBG" (untuk siswa). Hal ini akan memungkinkan penjelasan yang lebih tepat tentang hubungan "siapa melakukan apa, dengan siapa/apa" yang paling sering muncul dalam *tweet* negatif, netral, dan positif.

V. SIMPULAN DAN SARAN

5.1. Simpulan

Berdasarkan uraian, penjelasan, dan pengujian yang telah dilakukan terdapat beberapa kesimpulan diantaranya sebagai berikut.

1. Hasil penelitian menunjukkan bahwa percakapan masyarakat pada *platform X* tentang Program Makan Bergizi Gratis didominasi oleh sentimen negatif, diikuti sentimen netral, sedangkan sentimen positif jumlahnya paling sedikit. Sentimen negatif terutama muncul ketika pengguna menyoroiti kasus dugaan keracunan, kualitas menu, dan isu anggaran, sedangkan sentimen positif menguat ketika MBG dibingkai sebagai investasi gizi, generasi emas, dan pemerataan ke daerah seperti Papua. *Tweet* netral umumnya berupa laporan kegiatan dan informasi pelaksanaan program tanpa penilaian eksplisit.
2. Perbandingan kinerja model menunjukkan bahwa LSTM sedikit lebih baik dibanding GRU pada data uji. Model LSTM menghasilkan akurasi 77.63%, *precision* 77.51%, *recall* 77.4%, *F1-score* 77.66%, dan ROC AUC 91.54%. Model GRU memperoleh akurasi 75.59%, *precision* 75.24%, *recall* 75.71%, *F1-score* 75.40%, dan ROC AUC 90.82%. Selisihnya tidak besar, tetapi LSTM konsisten unggul pada *F1-score* dan sedikit lebih tinggi pada ROC AUC sehingga lebih layak dijadikan model utama untuk analisis sentimen lanjutan.

3. Analisis *n-gram*, *POS Tagging*, dan *dependency parsing* menunjukkan bahwa perbedaan sentimen lebih banyak ditentukan oleh kombinasi verba–objek dibanding sekadar siapa subjeknya. Sentimen negatif muncul ketika tindakan seperti makan, jadi, pakai, bikin, bayar diarahkan pada objek bernuansa risiko (makanan basi, keracunan, anggaran, uang, rakyat). Sentimen positif muncul ketika verba bernuansa perbaikan (lanjut, dukung, tingkat, tumbuh, bantu, bangun) dikaitkan dengan objek manfaat (anak, gizi, stunting, ekonomi, masa depan bangsa). Sentimen netral didominasi struktur subjek-verba-objek yang melaporkan kegiatan (kunjungan, peninjauan, pemberian makanan) tanpa ekspresi emosional. Dapat disimpulkan, citra MBG di mata publik sangat bergantung pada bagaimana hubungan tindakan dan objek tersebut dikomunikasikan dalam wacana daring.

5.2. Saran

Berdasarkan hasil penelitian, didapatkan saran untuk melanjutkan penelitian ini sebagai berikut.

1. Data penelitian dapat diperluas tidak hanya dari media sosial X, tetapi juga dari *platform* lain seperti TikTok, Instagram, dan Youtube agar representasi opini publik lebih beragam.
2. Sumber data disesuaikan dengan target audiens seperti untuk target SMA lebih banyak menggunakan media sosial Tiktok dibandingkan X.
3. Penelitian berikutnya dapat membandingkan lebih banyak arsitektur model, misalnya variasi *transformer* dan model *hybrid*.
4. Penelitian berikutnya dapat membahas kemungkinan *buzzer* di tiap kelas sentimen.

DAFTAR PUSTAKA

- Abidin, A. (2024). *Kamus Custom Preprocessing Analisis Sentimen Indonesia*.
<https://github.com/ArhamZaldin/Kamus-Custom-Preprocessing-Analisis-Sentimen-Indonesia/blob/main/sentiment-words.json>
- Ain, Q. T., Ali, M., Riaz, A., Noureen, A., Kamran, M., Hayat, B., & Rehman, A. (2017). Sentiment Analysis Using Deep Learning Techniques: A Comprehensive Review. *International Journal of Multimedia Information Retrieval (IJACSA)*, 8(6), 424–433. <https://doi.org/10.1007/s13735-023-00308-2>
- Alhudhaif, A. (2021). A Novel Multi-Class Imbalanced EEG Signals Classification Based On The Adaptive Synthetic Sampling (ADASYN) Approach. *PeerJ Computer Science*, 7, e523. <https://doi.org/10.7717/peerj-cs.523>
- Amaria, S. C., & Hidayati, N. (2025). Analisis Program Makan Bergizi Gratis Dengan Support Vector Machine (Svm) Pada Aplikasi X. *Jurnal Sistem Informasi*, 12(2), 16–24.
- Belaroussi, R., Noufe, S. C., & Dupin, F. (2025). Polarity of Yelp Reviews : A BERT – LSTM Comparative Study. *Big Data and Cognitive Computing*, 9(5), 1–25. <https://doi.org/https://doi.org/10.3390/bdcc9050140>
- Bustamin, A., Prayogi, A. A., Siswanto, D., Rafrin, M., & Nurdin, A. (2025). Text Normalization for Indonesian Slang Words in Sentiment Analysis Development. *ICIC Express Letters*, 16(2), 121–129. <https://doi.org/10.24507/icicelb.16.02.121>
- Chen, T., Kornblith, S., Swersky, K., Norouzi, M., & Hinton, G. (2020). Big self-supervised models are strong semi-supervised learners. *34th Conference on Neural Information Processing Systems*, 1–18.

- Chicco, D., & Jurman, G. (2023). The Matthews Correlation Coefficient (MCC) Should Replace The ROC AUC As The Standard Metric For Assessing Binary Classification. *Bio Data Mining*, 16(4), 1–23. <https://doi.org/10.1186/s13040-023-00322-4>
- Christian, W., Adamlu, D., Yu, A., & Suhartono, D. (2025). Leveraging IndoBERT and DistilBERT for Indonesian Emotion Classification in E-Commerce Reviews. *He 10th International Conference on Computer Science and Computational Intelligence 2025 Leveraging*, 00. <http://arxiv.org/abs/2509.14611>
- Cinelli, M., de Francisci Morales, G., Galeazzi, A., Quattrociocchi, W., & Starnini, M. (2021). The Echo Chamber Effect On Social Media. *Proceedings of the National Academy of Sciences of the United States of America*, 118(9), 1–8. <https://doi.org/10.1073/pnas.2023301118>
- Dyer, R. (2022). An Exploratory Study on the Predominant Programming Paradigms in Python Code. In *Proceedings of the 30th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE '22), November 14–18, 2022, Singapore, Singapore* (Vol. 1, Issue 1). Association for Computing Machinery. <https://doi.org/10.1145/3540250.3549158>
- Edward, A., Zevini, B. J., & Abiola, H. M. (2023). Analyzing Election Sentiments in Tweets with Gated Recurrent Units. *Asian Journal of Research in Computer Science*, 16(4), 125–132. <https://doi.org/10.9734/ajrcos/2023/v16i4376>
- Fahreza, S., & Setiawan, E. B. (2024). Sentiment Analysis on Presidential Election 2024 Using Word2Vec and Gated Recurrent Unit (GRU) with Genetic Algorithm Optimization. *International Journal on Information and Communication Technology*, 10(1), 62–77. <https://doi.org/10.21108/ijoict.v10i1.903>
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F.,

- Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- He, H., Bai, Y., Garcia, E. A., & Li, S. (2008). ADASYN: Adaptive Synthetic Sampling Approach For Imbalanced Learning. *Proceedings of the International Joint Conference on Neural Networks*, 1322–1328. <https://doi.org/10.1109/IJCNN.2008.4633969>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineer*, 9(3), 90–95. <https://doi.org/10.1109/MCSE.2007.55>
- Ibrohim, M. O., & Budi, I. (2018). A Dataset and Preliminaries Study for Abusive Language Detection in Indonesian Social Media. *Procedia Computer Science*, 135, 222–229. <https://doi.org/10.1016/j.procs.2018.08.169>
- Indonesia, P. R. (2025). *Rancangan Undang-Undang Anggaran Pendapatan dan Belanja Negara Tahun Anggaran 2026*. Kementerian Keuangan. <https://anggaran.kemenkeu.go.id/api/Medias/2a68f41e-e6a4-476f-ae46-d8d4c0d8c62d>
- Jurafsky, D., & Martin, J. H. (2018). Speech And Language Processing. In *Stanford University* (Vol. 10, Issue 3).
- Kastrati, M., Kastrati, Z., Shariq, A., & Marenglen, I. (2024). Leveraging Distant Supervision And Deep Learning For Twitter Sentiment And Emotion Classification. *Journal of Intelligent Information Systems*, 62(10), 1045–1070. <https://doi.org/10.1007/s10844-024-00845-0>
- Kawatu, C. I. (2019). *The Application Of Gated Recurrent Unit In Social Media Sentiment Analysis*.
- Khadapi, M., & Pakpahan, V. M. (2024). Analisis Sentimen Berbasis Jaringan LSTM dan BERT terhadap Diskusi Twitter tentang Pemilu 2024. *JUKI: Jurnal Komputer Dan Informatika*, 6(2), 130–137.

- Krisnawan, Rabbani, Z. A., Trimono, & Idhom, M. (2025). Analisis sentimen program makan bergizi gratis menggunakan bidirectional gated recurrent unit. *Jurnal Penerapan Teknologi Informasi Dan Komunikasi*, 4(3), 282–294. <https://doi.org/10.24246/itexplore.v4i3.2025.pp282-294>
- Kurniawan, A. A., & Mustikasari, M. (2021). Implementasi Deep Learning Menggunakan Metode CNN dan LSTM untuk Menentukan Berita Palsu dalam Bahasa Indonesia. *Jurnal Informatika Universitas Pamulang*, 5(4), 544–552. <https://doi.org/10.32493/informatika.v5i4.6760>
- Laila, M. M., Hasan, F. N., & Kuntoro, A. Y. (2025). Analisis Sentimen Program Makan Gratis Pada Platform X Menggunakan Algoritma Naïve Bayes. *Jurnal Ilmiah Informatika*, 13(2), 259–264.
- Lemaitre, G., Nogueira, F., & Aridas, C. K. (2017). Imbalanced-learn: A Python Toolbox to Tackle the Curse of Imbalanced Datasets in Machine Learning. *Journal of Machine Learning Research*, 18(1), 559-563 TensorFlow. (n.d.). Module:tf. Tensorflow.O. <http://arxiv.org/abs/1512.03502>
- Lenci, A., & Padó, S. (2022). Editorial: Perspectives For Natural Language Processing Between AI, Linguistics And Cognitive Science. *Frontiers in Artificial Intelligence*, 5, 1–3. <https://doi.org/10.3389/frai.2022.1059998>
- Lestari, A. R. T., Perdana, R. S., & Fauzi, M. A. (2017). Analisis Sentimen Tentang Opini Pilkada DKI Jakarta 2017 Pada Dokumen Twitter Berbahasa Indonesia Menggunakan Naïve Bayes. *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 1(December), 1718–1724. <http://j-ptiik.ub.ac.id>
- Ma'aly, A. N., Pramesti, D., Fathurahman, A. D., & Fakhurroja, H. (2024). Exploring Sentiment Analysis for the Indonesian Presidential Election Through Online Reviews Using Multi-Label Classification with a Deep Learning Algorithm. *Information (Switzerland)*, 15(11), 1–33. <https://doi.org/10.3390/info15110705>
- Merlinda, A. A., & Yusuf, Y. (2025). Analisis Program Makan Gratis Prabowo Subianto Terhadap Strategi Peningkatan Motivasi Belajar Siswa di Sekolah Tinjauan dari Perspektif Sosiologi Pendidikan. *Ranah Research : Journal of*

Multidisciplinary Research and Development, 7(2), 1364–1373.
<https://doi.org/10.38035/rrj.v7i2.1360>

Mienye, I. D., Swart, T. G., & Obaido, G. (2024). Recurrent Neural Networks: A Comprehensive Review of Architectures, Variants, and Applications. *Information (Switzerland)*, 15(9), 1–34. <https://doi.org/10.3390/info15090517>

Mustapha, B., Zhou, Y., Nawel, B., Chunyan, S., & Zhitao, X. (2025). Journal of Radiation Research and Applied Sciences Optimized attention U-Net for enhanced lung and area of infection segmentation in chest X-Rays and CT scans. *Journal of Radiation Research and Applied Sciences*, 18(3), 1–24. <https://doi.org/10.1016/j.jrras.2025.101650>

Naufal, M. F., & Kusuma, S. F. (2022). Analisis Sentimen pada Media Sosial Twitter Terhadap Kebijakan Pemberlakuan Pembatasan Kegiatan Masyarakat Berbasis Deep Learning. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 8(1), 44–49. <https://doi.org/10.26418/jp.v8i1.49951>

Nugraha, D. S. (2024). A Morphological Analysis of the Indonesian Suffixation: A Look at the Different Types of Affixes and Their Semantic Changes. *GEMA Online Journal of Language Studies*, 24(4), 109–132. <https://doi.org/10.17576/gema-2024-2404-07>

Pratiwi, R. W., H, S. F., Dairoh, D., Af'idah, D. I., A, Q. R., & F, A. G. (2021). Analisis Sentimen Pada Review Skincare Female Daily Menggunakan Metode Support Vector Machine (SVM). *Journal of Informatics, Information System, Software Engineering and Applications (INISTA)*, 4(1), 40–46. <https://doi.org/10.20895/inista.v4i1.387>

Puteri, D. I. (2023). Implementasi Long Short Term Memory (LSTM) dan Bidirectional Long Short Term Memory (BiLSTM) Dalam Prediksi Harga Saham Syariah. *Euler: Jurnal Ilmiah Matematika, Sains Dan Teknologi*, 11(1), 35–43. <https://doi.org/10.34312/euler.v11i1.19791>

Python Software Foundation. (2026). *Wordcloud 1.9.6*. Pypi.Org. https://pypi.org/project/wordcloud/?utm_source=chatgpt.com

- Rahman, M. Z., Sari, Y. A., & Yudistira, N. (2021). Analisis Sentimen Tweet COVID-19 Menggunakan Word Embedding dan Metode Long Short-Term Memory (LSTM). *Jurnal Pengembangan Teknologi Informasi Dan Ilmu Komputer*, 5(11), 5120–5127. <http://j-ptiik.ub.ac.id>
- Rahmawati, A., Sulandari, W., Subanti, S., & Yudhanto, Y. (2023). Penerapan Metode Reccurent Neural Network dengan Pendekatan Long Short-Term Memory (LSTM) untuk Meramalkan Harga Saham Hybe Corporation The Application of Recurrent Neural Network Method with the Long Short-Term Memory (LSTM) Approach to Forecast Hybe. *Jurnal Bumigora Information Technology*, 5(1), 65–76. <https://doi.org/10.30812/bite/v5i1.2973>
- Ramadhani, N. G., & Khoirunnisa, A. (2025). An Integrated Random Forest for Analyzing Public Sentiment on the “Makan Bergizi Gratis” Program Nur. *Journal of Information Systems and Informatics*, 7(3), 2314–2328. <https://doi.org/10.51519/journalisi.v6i2.759>
- Rehani, S. (2020). Social Media Personalization Algorithms and the Emergence of Filter Bubbles. *ISChannel*, 15(1), 20–24.
- Saeed, M. Y., Awais, M., Younas, M., Shah, M. A., Khan, A., Uddin, M. I., & Mahmoud, M. (2021). An abstractive summarization technique with variable length keywords as per document diversity. *Computers, Materials and Continua*, 66(3), 2409–2423. <https://doi.org/10.32604/cmc.2021.014330>
- Salsabila, N. A., Ardhito Winatmoko, Y., Akbar Septiandri, A., & Jamal, A. (2018). Colloquial Indonesian Lexicon. *Proceedings of the 2018 International Conference on Asian Language Processing, IALP 2018*, 226–229. <https://doi.org/10.1109/IALP.2018.8629151>
- Sanusi, R., Astuti, F. D., & Buryadi, I. Y. (2021). Analisis Sentimen Pada Twitter Terhadap Program Kartu Pra Kerja Dengan Recurrent Neural Network. *Jurnal Informatika Dan Komputer*, 5(2), 89–99. <https://doi.org/10.26798/jiko.v5i2.645>
- Sartika, D., & Saluza, I. (2022). Penerapan Metode Principal Component Analysis (PCA) Pada Klasifikasi Status Kredit Nasabah Bank Sumsel Babel Cabang

- KM 12 Palembang Menggunakan Metode Decision Tree. *The Generic Journal*, 14(2), 45–49. <https://doi.org/10.18495/generic.v14i2.130>
- Satria, H. (2024). *Cara Mendapatkan Data (Crawl) Twitter X - 30 Maret 2024*. www.youtube.com. <https://youtu.be/BxryBUtUqEk?si=Lya8TrjSUKqwb3-f>
- Scikit-learn developers. (2025). *API Reference*. Scikit-Learn.Org. <https://scikit-learn.org/stable/api/index.html>
- Subarkah, P., Ikhsan, A. N., Anggraeni, E., & Sabaniyah, A. P. (2025). Sentiment Perspective of Government's Free Nutritious Meal Policy on Social Media X using Indo-BERT and Bi-LTSM. *Journal of Technology and Informatics (JoTI)*, 7(2), 201–210. <https://doi.org/10.37802/joti.v7i2.1065>
- Szeghalmy, S., & Fazekas, A. (2023). A Comparative Study of the Use of Stratified Cross-Validation and Distribution-Balanced Stratified Cross-Validation in Imbalanced Learning. *Sensors*, 23(4), 1–27. <https://doi.org/10.3390/s23042333>
- Talukder, M. A., Talaat, A. S., Kazi, M., & Khraisat, A. (2025). XAI-HD: An Explainable Artificial Intelligence Framework For Heart Disease Detection. *Artificial Intelligence Review*, 58(12), 1–78. <https://doi.org/10.1007/s10462-025-11385-6>
- Tamami, G., Triyanto, W. A., & Muzid, S. (2025). Sentiment Analysis Mobile JKN Reviews Using SMOTE Based LSTM. *Indonesian Journal of Computing and Cybernetics Systems*, 19(1), 13–24. <https://doi.org/10.22146/ijccs.101910>
- TensorFlow. (2026). *Module:tf*. Tensorflow.Org. https://www.tensorflow.org/api_docs/python/tf
- Thaker, N., & Shukla, A. (2020). Python as Multi Paradigm Programming Language. *International Journal of Computer Applications*, 177(31), 38–42. <https://doi.org/10.5120/ijca2020919775>
- Torregrossa, F., Allesiaro, R., Claveau, V., Kooli, N., & Gravier, G. (2021). A Survey On Training And Evaluation Of Word Embeddings. *International Journal of Data Science and Analytics*, 11(2), 85–103.

<https://doi.org/10.1007/s41060-021-00242-8>

- Vidal, S. (2023). *Apa itu Bahasa Pemrograman Multiparadigma?* Tecnobits. <https://tecnobits.com/id/Apa-itu-bahasa-pemrograman-multiparadigma%3F/>
- Wahid, D. H., & SN, A. (2016). Peringkasan Sentimen Esktraktif di Twitter Menggunakan Hybrid TF-IDF dan Cosine Similarity. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 10(2), 207. <https://doi.org/10.22146/ijccs.16625>
- Wilie, B., Vincentio, K., Winata, G. I., Cahyawijaya, S., Li, X., Lim, Z. Y., Soleman, S., Mahendra, R., Fung, P., Bahar, S., & Purwarianti, A. (2024). IndoNLU: Benchmark and Resources for Evaluating Indonesian Natural Language Understanding. *Proceedings of the 1st Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics and the 10th International Joint Conference on Natural Language Processing*, 843–857. <https://doi.org/10.18653/v1/2020.aacl-main.85>
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Funtowicz, M., Davison, J., Shleifer, S., Platen, P. Von, Ma, C., Jernite, Y., Plu, J., Xu, C., Scao, T. Le, Gugger, S., Drame, M., ... Rush, A. M. (2020). Transformers : State-of-the-Art Natural Language Processing. *Proceedings of the 2020 EMNLP (Systems Demonstrations)*, 38–45.
- Yadav, A. (2025). *Bayesian Hyperparameter Tuning with Optuna: A Beginner-Friendly Guide*. Medium. <https://medium.com/@avinashyadav16/bayesian-hyperparameter-tuning-with-optuna-a-beginner-friendly-guide-6a426f43c76e>
- Zhang, A., Lipton, Z., Li, M., & Smola, A. J. (2023). Dive Into Deep Learning. In *Journal of the American College of Radiology* (Vol. 17, Issue 5). <https://doi.org/10.1016/j.jacr.2020.02.005>