PERBANDINGAN ALGORITMA XGBOOST DAN LSTM DALAM ANALISIS SENTIMEN

SKRIPSI

Program Studi Sarjana Informatika
Jurusan Informatika

Oleh:

SYARIFAH RAISTA FITRALIZA
NIM D1041211013

FAKULTAS TEKNIK
UNIVERSITAS TANJUNGPURA
PONTIANAK
2026

HALAMAN PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Syarifah Raista Fitraliza

NIM : D1041211013

menyatakan bahwa dalam SKRIPSI yang berjudul "Perbandingan Algoritma XGBoost dan LSTM dalam Analisis Sentimen" tidak terdapat karya yang pernah diajukan untuk memperoleh gelar sarjana di suatu perguruan tinggi manapun. Sepanjang pengetahuan Saya, tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam Daftar Rujukan.

Demikian pernyataan ini dibuat dengan sebenar-benarnya. Saya sanggup menerima konsekuensi akademis dan hukum di kemudian hari apabila pernyataan yang dibuat ini tidak benar.

Pontianak, 17 Juni 2026



Syarifah Raista Fitraliza
NIM D1041211013

KEMENTERIAN PENDIDIKAN TINGGI, SAINS, DAN TEKNOLOGI
UNIVERSITAS TANJUNGPURA
FAKULTAS TEKNIK
Jalan Prof. Dr. H. Hadari Nawawi Pontianak 78124
Telepon (0561) 740186, WA: +6282152280907
Email : ft@untan.ac.id Website : http://teknik.untan.ac.id

HALAMAN PENGESAHAN

PERBANDINGAN ALGORITMA XGBOOST DAN LSTM DALAM ANALISIS SENTIMEN

Jurusan Informatika
Program Studi Sarjana Informatika

Oleh:

Syarifah Raista Fitraliza
NIM D1041211013

Telah dipertahankan di depan Penguji Skripsi pada tanggal 17 Juni 2026 dan diterima sebagai salah satu persyaratan untuk memperoleh gelar sarjana.

Susunan Penguji Skripsi:

Pembimbing Utama	Penguji Utama
Tursina, S.T., M.Cs.	Dr. Arif Bijaksana Putra Negara, S.T., M.T.
NIP. 197801152002122003	NIP. 197208081998021002
Pembimbing Kedua	Penguji Kedua
Rina Septiriana, S.T., M.Cs.	Niken Candraningrum, S.T., M.Cs.
NIP. 198709232020122001	NIP. 199309272022032010

Pontianak, 17 Juni 2026
Dekan,

Dr.-Ing. Ir. Slamet Widodo, M.T., IPM
NIP. 196712231992031002

HALAMAN PERSEMBAHAN

Dengan penuh rasa syukur dan ketulusan, karya sederhana ini penulis persembahkan kepada:

1. Abah, Mama, dan Kakak yang selalu menjadi tempat pertama penulis untuk pulang. Terima kasih untuk doa yang tidak pernah putus dan kasih sayang yang tidak terbatas. Meski raga Abah sudah tidak lagi ada, semangat dan cintanya senantiasa mengalir dan menjadi kekuatan bagi penulis.
2. Ami dan Ibu yang menjadi orang tua selama di rantauan. Terima kasih atas segala perhatian dan kepedulian tanpa pamrih yang diberikan selama penulis menempuh pendidikan.
3. Dosen Pembimbing Tugas Akhir, Ibu Tursina, S.T., M.Cs. dan Ibu Rina Septiriana, S.T., M.Cs. yang senantiasa mendampingi dan memberikan arahan selama penyusunan Tugas Akhir ini. Terima kasih atas waktu, tenaga, dan pikiran yang tercurahkan dalam setiap diskusi, baik mengenai topik akademik maupun topik lainnya yang membantu memperkaya perspektif penulis. Terima kasih pula karena selalu memberikan motivasi serta keyakinan di saat penulis merasa ragu.
4. Rahma Putri Setyoningsih dan Asmi Januariza, dua orang yang selalu memastikan penulis dalam kondisi baik selama penyusunan Tugas Akhir. Terima kasih atas kesediaan untuk berbagi opini di berbagai kondisi tanpa pernah saling menghakimi.
5. Syifa Nuraini, Lia Apriliana, dan Dhia Hana Muthiah yang membersamai perjalanan studi di Informatika. Kehadiran kalian dalam proses penuh makna ini menjadi salah satu hal yang sangat penulis syukuri. Terima kasih telah menjadi rekan perjuangan yang suportif bagi penulis.

PRAKATA

Puji syukur kehadirat Allah SWT atas kehendak-Nya maka penelitian dan penulisan skripsi dengan judul "Perbandingan Algoritma XGBoost dan LSTM dalam Analisis Sentimen" ini dapat diselesaikan dengan baik. Tujuan dari penyusunan skripsi ini dibuat untuk memenuhi syarat dalam menyelesaikan Program Sarjana (S1) pada Fakultas Teknik, Jurusan Informatika, Universitas Tanjungpura.

Penyelesaian penyusunan skripsi ini tidak lepas dari bantuan, arahan dan bimbingan banyak pihak. Oleh sebab itu penulis ingin menyampaikan rasa hormat dan terima kasih yang sebesar-besarkan kepada pihak yang telah turut andil memberikan bantuan dan dukungan kepada penulis.

Secara khusus penulis menyampaikan ucapan terima kasih kepada Ibu Tursina, S.T., M.Cs. selaku dosen Pembimbing I dan Ibu Rina Septiriana, S.T., M.Cs. selaku dosen Pembimbing II serta kepada Bapak Dr. Arif Bijaksana Putra Negara, S.T., M.T. selaku dosen Penguji I dan Ibu Niken Candraningrum, S.T., M.Cs selaku penguji II.

Penulis menyadari masih terdapat kekurangan dalam penyusunan skripsi ini. Dengan segala kerendahan hati penulis menerima kritik dan saran untuk dijadikan sebagai bahan evaluasi dan pengembangan penulisan berikutnya. Semoga hasil penelitian ini dapat memberikan kontribusi yang bermanfaat bagi pembaca dan pihak-pihak yang berkepentingan.

Pontianak, 17 Juni 2026

Penulis,

Syarifah Raista Fitraliza

ABSTRAK

Tingginya volume opini publik berbasis teks di media digital memerlukan analisis sentimen untuk klasifikasi secara otomatis. Pengolahan data teks yang kompleks ini sangat bergantung pada pemilihan algoritma yang tepat. Oleh karena itu, penelitian ini bertujuan membandingkan performa algoritma *Extreme Gradient Boosting* (XGBoost) dan *Long Short-Term Memory* (LSTM) dalam mengklasifikasikan sentimen menjadi positif, netral, atau negatif dengan fokus pada nilai akurasi. Pengujian dilakukan menggunakan dua dataset dengan karakteristik berbeda, yaitu opini Program Makan Bergizi Gratis pada platform X serta ulasan produk makanan dan minuman pada platform Tokopedia. Untuk menangani karakteristik distribusi data pada kedua dataset tersebut, penelitian ini menerapkan penyeimbangan data melalui metode SMOTE dan ADASYN serta skenario tambahan dengan melakukan undersampling terlebih dahulu sebelum dilakukan SMOTE dan ADASYN, diikuti *hyperparameter tuning* berbasis *Random Search* dan *Tree-structured Parzen Estimator* (TPE). Hasil pengujian menunjukkan bahwa LSTM dengan skenario data asli yang dikombinasikan dengan *Random Search* secara konsisten menghasilkan akurasi tertinggi pada kedua dataset, yaitu sebesar 0,8625 pada data Makan Bergizi Gratis dan 0,8062 pada data Tokopedia. Penerapan SMOTE dan ADASYN justru memicu penurunan akurasi yang signifikan pada LSTM akibat gejala *overfitting*, sebab kapasitas arsitektur jaringan saraf tersebut cenderung menghafal titik data sintetis daripada mempelajari pola umum. Sebaliknya, XGBoost menunjukkan stabilitas yang jauh lebih baik terhadap gangguan data sintetis berkat mekanisme pembagian data berbasis aturan yang kaku. Adapun penerapan *hyperparameter tuning* menghasilkan pengaruh yang cenderung stagnan pada hasil akurasi data uji, tetapi proses ini tetap berperan penting dalam mengoptimalkan konfigurasi parameter terbaik untuk setiap arsitektur model. Berdasarkan uji signifikansi statistik menggunakan *paired t-test* terhadap hasil *k-fold cross validation*, perbedaan nilai akurasi antara LSTM dan XGBoost terbukti signifikan secara statistik.

Kata kunci: Analisis Sentimen, XGBoost, LSTM, Penyeimbangan Data, *Hyperparameter tuning*

ABSTRACT

The high volume of text-based public opinion in digital media necessitates sentiment analysis for automated classification. Processing this complex textual data heavily relies on selecting the appropriate algorithm. Therefore, this study aims to compare the performance of Extreme Gradient Boosting (XGBoost) and Long Short-Term Memory (LSTM) algorithms in classifying sentiment into positive, neutral, or negative, with a primary focus on accuracy. Testing was conducted using two datasets with distinct characteristics: opinions on the Free Nutritious Meal Program (Makan Bergizi Gratis) on the X platform, and food and beverage product reviews on Tokopedia. To address the data distribution characteristics of both datasets, this study applied data balancing techniques using SMOTE and ADASYN, alongside an additional scenario involving undersampling prior to SMOTE and ADASYN. This was followed by hyperparameter tuning based on Random Search and Tree-structured Parzen Estimator (TPE). The experimental results demonstrate that LSTM using the original data scenario combined with Random Search consistently yields the highest accuracy across both datasets, achieving 0.8625 on the Free Nutritious Meal data and 0.8062 on the Tokopedia data. Conversely, the application of SMOTE and ADASYN triggered a significant decline in LSTM accuracy due to overfitting, as the neural network architecture's capacity tends to memorize synthetic data points rather than learning generalized patterns. In contrast, XGBoost exhibited substantially better stability against synthetic data interference, owing to its rigid, rule-based data splitting mechanism. Furthermore, while the implementation of hyperparameter tuning exerted a stagnant effect on the test data accuracy, the process remained crucial in optimizing the best parameter configurations for each model architecture. Based on statistical significance testing using a paired t-test on the k-fold cross-validation results, the difference in accuracy values between LSTM and XGBoost was proven to be statistically significant.

Keywords: Sentiment Analysis, XGBoost, LSTM, Data Balancing, Hyperparameter tuning.

DAFTAR ISI

HALAMAN PERNYATAAN ii
HALAMAN PERSEMBAHAN iv
PRAKATA v
ABSTRAK vi
ABSTRACT vii
DAFTAR ISI viii
DAFTAR GAMBAR xii
DAFTAR TABEL xiv
DAFTAR KODE PROGRAM xvi
DAFTAR LAMPIRAN xvii
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Perumusan Masalah 4
1.3 Tujuan Penelitian 4
1.4 Pembatasan Masalah 5
1.5 Sistematika Penulisan 5
BAB II TINJAUAN PUSTAKA 7
2.1 Penelitian Terkait 7
2.2 Landasan Teori 11
2.1.1 *Natural Language Processing* (NLP) 11
2.1.2 *Text Mining* 12
2.1.3 Analisis Sentimen 12
2.1.4 *Text Preprocessing* 13
2.1.5 TF-IDF 14
2.1.6 Ketidakseimbangan Data 15

2.1.7 *Undersampling* 15
2.1.8 SMOTE 15
2.1.9 ADASYN 16
2.1.10 *Hyperparameter Tuning* 16
2.1.11 *Random Search* 16
2.1.12 *Tree-structured Parzen Estimator* (TPE) 18
2.1.13 *Extreme Gradient Boosting* (XGBoost) 18
2.1.14 *Long Short-Term Memory* (LSTM) 20
2.1.15 *Cross-Validation* 22
2.1.16 Uji Normalitas Shapiro-Wilk 23
2.1.17 *Paired Sample T-test* 24
2.1.18 *Confusion Matrix* 25
2.1.19 Python 26
BAB III METODOLOGI PENELITIAN 27
3.1 Data Penelitian 27
3.2 Perangkat Penelitian 27
3.3 Metodologi Penelitian 28
3.1.1 Pengumpulan Data 29
3.1.2 Pelabelan Data 29
3.1.3 Analisis Data 29
3.1.4 *Text Preprocessing* 29
3.1.5 Pembagian Dataset 30
3.1.6 TF-IDF 30
3.1.7 Penyeimbangan data 33
3.1.8 *Hyperparameter Tuning* 35
3.1.9 Pemodelan 38

3.1.10 Skenario Pengujian 40
3.1.11 Validasi 42
3.1.12 Uji Signifikansi Statistik 43
3.1.13 Evaluasi 44
3.1.14 Analisis Hasil 45
3.1.15 Penarikan Kesimpulan 45
BAB IV HASIL DAN ANALISIS 46
4.1 Hasil Implementasi 46
4.1.1 Pengumpulan Data 46
4.1.2 Pelabelan Data 48
4.1.3 Analisis Data 48
4.1.4 *Text Preprocessing* 54
4.1.5 Pembagian Dataset 63
4.1.6 TF-IDF 64
4.1.7 Penyeimbangan data 65
4.1.8 Implementasi Model 68
4.1.9 Hasil Validasi 73
4.1.10 Hasil Uji Signifikansi Statistik 80
4.1.11 Hasil Evaluasi 93
4.2 Analisis Hasil 102
4.2.1 Analisis Hasil Klasifikasi 102
4.2.2 Analisis Perbandingan Akurasi Data Latih dan Data Uji 109
4.2.3 Analisis Perbandingan Akurasi Skenario Utama pada Data MBG dan Tokopedia 116
BAB V PENUTUP 119
5.1 Kesimpulan 119

5.2 Saran 120

DAFTAR RUJUKAN 122

DAFTAR GAMBAR

Gambar 2.1 Ilustrasi Random search 17
Gambar 3.1 Metodologi Penelitian 28
Gambar 3.2 Proses TF-IDF 30
Gambar 3.3 Proses SMOTE 34
Gambar 3.4 Proses ADASYN 34
Gambar 3.5 Proses *Undersampling* + SMOTE 35
Gambar 3.6 Proses *Undersampling* + ADASYN 35
Gambar 3.7 Proses *Random Search* 36
Gambar 3.8 Proses TPE 37
Gambar 3.9 Alur Pemodelan XGBoost 39
Gambar 3.10 Alur Pemodelan LSTM 39
Gambar 3.11 Arsitektur LSTM 40
Gambar 4.1 Distribusi Jumlah Data per Label Data MBG 49
Gambar 4.2 Jumlah Kata per Label Data MBG 50
Gambar 4.3 Rata-rata dan Variasi per Label Data MBG 50
Gambar 4.4 Distribusi Jumlah Data per Label Data Tokopedia 51
Gambar 4.5 Jumlah Kata per Label Data Tokopedia 52
Gambar 4.6 Rata-rata dan Variasi per Label Data Tokopedia 53
Gambar 4.7 Hasil TF-IDF Data MBG 65
Gambar 4.8 Hasil TF-IDF Data Tokopedia 65
Gambar 4.9 Distribusi Jumlah Data per Label Data MBG 66
Gambar 4.10 Distribusi Jumlah Data MBG (Kombinasi Penyeimbangan Data) per Label 67
Gambar 4.11 Distribusi Jumlah Data per Label Data Tokopedia 68
Gambar 4.12 Akurasi Validasi Skenario Utama XGBoost pada Data MBG 74
Gambar 4.13 Akurasi Validasi Skenario Utama LSTM pada Data MBG 74
Gambar 4.14 Akurasi Validasi Skenario Utama XGBoost pada Data Tokopedia 76
Gambar 4.15 Akurasi Validasi Skenario Utama LSTM pada Data Tokopedia ... 76
Gambar 4.16 Akurasi Validasi XGBoost Skenario Tambahan Data MBG 78
Gambar 4.17 Akurasi Validasi LSTM Skenario Tambahan Data MBG 79
Gambar 4.18 *Confusion Matrix* Performa Terbaik Data MBG 94

Gambar 4.19 *Confusion Matrix* Performa Terburuk Data MBG 95
Gambar 4.20 *Confusion Matrix* Performa Terbaik Data Tokopedia..... 96
Gambar 4.21 *Confusion Matrix* Performa Terbaik Data Tokopedia..... 97
Gambar 4.22 *Confusion Matrix* Performa Terburuk Data Tokopedia 98
Gambar 4.23 *Confusion Matrix* Performa Terbaik Skenario Tambahan Data MBG 99
Gambar 4.24 *Confusion Matrix* Performa Terburuk Skenario Tambahan Data MBG 100
Gambar 4.25 *Confusion Matrix* Performa Terburuk Skenario Tambahan Data MBG 101
Gambar 4.26 Perbandingan Akurasi Skenario Utama Data MBG..... 102
Gambar 4.27 Perbandingan Akurasi Skenario Utama Data Tokopedia..... 105
Gambar 4.28 Perbandingan Akurasi Skenario Tambahan Data MBG 107
Gambar 4.29 Perbandingan Akurasi Latih dan Uji XGBoost pada Skenario Utama Data MBG 110
Gambar 4.30 Perbandingan Akurasi Latih dan Uji LSTM pada Skenario Utama Data MBG 110
Gambar 4.31 Perbandingan Akurasi Latih dan Uji XGBoost pada Skenario Utama Data Tokopedia 112
Gambar 4.32 Perbandingan Akurasi Latih dan Uji LSTM pada Skenario Utama Data Tokopedia 112
Gambar 4.33 Perbandingan Akurasi Latih dan Uji XGBoost pada Skenario Tambahan Data MBG 114
Gambar 4.34 Perbandingan Akurasi Latih dan Uji LSTM pada Skenario Tambahan Data MBG 114
Gambar 4.35 Perbandingan Akurasi XGBoost pada Data MBG dan Tokopedia 116
Gambar 4.36 Perbandingan Akurasi LSTM pada Data MBG dan Tokopedia ... 117

DAFTAR TABEL

Tabel 2.1 Penelitian Terkait 7
Tabel 2.2 *Confusion Matrix* 25
Tabel 3.1 Sampel Data Perhitungan Manual TF-IDF 30
Tabel 3.2 Hasil Perhitungan Manual *Term Frequency* 31
Tabel 3.3 Hasil Perhitungan Manual *Inverse Document Frequency* 32
Tabel 3.4 Hasil Perhitungan Manual *Term Frequency-Inverse Document* 33
Tabel 3.5 *Hyperparameter* yang dioptimalkan 37
Tabel 3.6 Skenario Utama 41
Tabel 3.7 Skenario Tambahan 42
Tabel 4.1 Hasil Pengumpulan Data MBG 47
Tabel 4.2 Hasil Pengumpulan Data Tokopedia 47
Tabel 4.3 Hasil Pelabelan Data MBG 48
Tabel 4.4 Hasil *Cleaning* Data MBG 56
Tabel 4.5 Hasil *Cleaning* Data Tokopedia 56
Tabel 4.6 Hasil *Case Folding* Data MBG 56
Tabel 4.7 Hasil *Case Folding* Data Tokopedia 57
Tabel 4.8 Hasil Normalisasi Data MBG 59
Tabel 4.9 Hasil Normalisasi Data Tokopedia 59
Tabel 4.10 Hasil Tokenisasi Data MBG 59
Tabel 4.11 Hasil Tokenisasi Data Tokopedia 60
Tabel 4.12 Hasil *Filtering* Data MBG 60
Tabel 4.13 Hasil *Filtering* Data Tokopedia 61
Tabel 4.14 Hasil *Stemming* Data MBG 61
Tabel 4.15 Hasil *Stemming* Data Tokopedia 62
Tabel 4.16 Hasil *Text Preprocessing* Data MBG 62
Tabel 4.17 Hasil *Text Preprocessing* Data Tokopedia 63
Tabel 4.18 Pembagian Dataset MBG 64
Tabel 4.19 Pembagian Dataset Tokopedia 64
Tabel 4.20 *Hyperparameter* XGBoost Skenario Utama Data MBG 70
Tabel 4.21 *Hyperparameter* XGBoost Skenario Utama Data Tokopedia 70
Tabel 4.22 *Hyperparameter* XGBoost Skenario Tambahan Data MBG 70

Tabel 4.23 *Hyperparameter* LSTM Skenario Utama Data MBG 72
Tabel 4.24 *Hyperparameter* LSTM Skenario Utama Data Tokopedia 72
Tabel 4.25 *Hyperparameter* LSTM Skenario Tambahan Data MBG 72
Tabel 4.26 Nilai Akurasi Validasi Skenario Utama Data MBG 81
Tabel 4.27 Hasil Uji Normalitas Skenario Utama Data MBG 81
Tabel 4.28 Hasil Uji *Paired T-test* Skenario Utama Data MBG 82
Tabel 4.29 Peringkat Signifikansi Perubahan Antar Skenario Utama Data MBG 83
Tabel 4.30 Hasil Uji *Paired T-test* Antar Algoritma Skenario Utama Data MBG 83
Tabel 4.31 Nilai Akurasi Validasi Skenario Utama Data Tokopedia 85
Tabel 4.32 Hasil Uji Normalitas Skenario Utama Data Tokopedia 85
Tabel 4.33 Hasil Uji *Paired T-test* Skenario Utama Data Tokopedia 86
Tabel 4.34 Peringkat Signifikansi Perubahan Antar Skenario Utama Data Tokopedia 87
Tabel 4.35 Hasil Uji *Paired T-test* Antar Algoritma Skenario Utama Data Tokopedia 88
Tabel 4.36 Nilai Akurasi Validasi Skenario Tambahan Data MBG 89
Tabel 4.37 Hasil Uji Normalitas Skenario Tambahan Data MBG 90
Tabel 4.38 Hasil Uji *Paired T-test* Skenario Tambahan Data MBG 90
Tabel 4.39 Peringkat Signifikansi Perubahan Antar Skenario Tambahan Data MBG 91
Tabel 4.40 Hasil Uji *Paired T-test* Antar Algoritma Skenario Tambahan Data MBG 92
Tabel 4.41 Hasil Evaluasi Skenario Utama Data MBG 96
Tabel 4.42 Hasil Evaluasi Skenario Utama Data Tokopedia 99
Tabel 4.43 Hasil Evaluasi Skenario Tambahan Data MBG 102
Tabel 4.44 Perbandingan Akurasi Setiap Skenario terhadap Skenario Default pada Skenario Utama Data MBG 103
Tabel 4.45 Perbandingan Akurasi Setiap Skenario terhadap Skenario Default pada Skenario Utama Data Tokopedia 105
Tabel 4.46 Perbandingan Akurasi Setiap Skenario terhadap Skenario Default pada Skenario Tambahan Data MBG 107

DAFTAR KODE PROGRAM

Kode Program 4.1 *Crawling* Data 46
Kode Program 4.2 Menghapus URL 54
Kode Program 4.3 Menghapus HTML 55
Kode Program 4.4 Menghapus Simbol 55
Kode Program 4.5 Menghapus Angka 55
Kode Program 4.6 Menghapus *Username* 55
Kode Program 4.7 Menghapus *Hashtag* 56
Kode Program 4.8 *Case Folding* 56
Kode Program 4.9 Normalisasi 58
Kode Program 4.10 Tokenisasi 59
Kode Program 4.11 *Filtering* 60
Kode Program 4.12 *Stemming* 61
Kode Program 4.13 Pembagian Dataset 63
Kode Program 4.14 TF-IDF 64
Kode Program 4.15 ADASYN 66
Kode Program 4.16 SMOTE 66
Kode Program 4.17 *Undersampling* 66
Kode Program 4.18 *Random Search* pada XGBoost 69
Kode Program 4.19 TPE pada XGBoost 69
Kode Program 4.20 Implementasi Model XGBoost 70
Kode Program 4.21 *Random Search* pada LSTM 71
Kode Program 4.22 TPE pada LSTM 71
Kode Program 4.23 *Import Library* Model LSTM 73
Kode Program 4.24 Implementasi LSTM 73
Kode Program 4.25 Implementasi *Stratified Kfold Cross Validation* 73
Kode Program 4.26 *Confusion Matrix* 93

DAFTAR LAMPIRAN

LAMPIRAN A *Classification Report* A-127

LAMPIRAN B Hasil *Text Preprocessing* B-1

BAB I
PENDAHULUAN

1.1 Latar Belakang

Ruang publik digital saat ini telah menjadi media utama bagi masyarakat untuk mengekspresikan opini, baik terhadap suatu kebijakan publik maupun evaluasi terhadap suatu produk. Interaksi publik yang intens ini secara akumulatif menghasilkan volume data teks berskala besar yang bersifat tidak terstruktur. Dalam upaya memproses dan merepresentasikan informasi yang relevan dari kumpulan data teks yang kompleks tersebut, diperlukan implementasi pendekatan *text mining*. *Text mining* merupakan salah satu bidang khusus dalam *data mining* yang memiliki definisi menambang data berupa teks dengan sumber data yang biasanya didapatkan dari dokumen. Tujuannya adalah mencari kata-kata yang dapat mewakili isi dari dokumen sehingga dapat dilakukan analisis keterhubungan antardokumen (Nur Rahmah & Rini Sulistiyowati, 2022.). Salah satu cabang dari *text mining* yang secara spesifik berfokus pada ekstraksi opini, emosi, dan evaluasi masyarakat terhadap suatu entitas atau isu adalah analisis sentimen. Sebagai bagian dari *Natural Language Processing* (NLP), analisis sentimen bertujuan untuk mengklasifikasikan opini dalam teks menjadi kategori positif, negatif, atau netral (Anggriyani & Fakhriza, 2024).

Dalam penerapannya, kinerja dan generalisasi model analisis sentimen sangat dipengaruhi oleh karakteristik data teks yang dianalisis. Untuk menguji keandalan algoritma secara komprehensif, penggunaan dataset dengan domain dan gaya bahasa yang berbeda menjadi sangat penting. Oleh karena itu, penelitian ini menggunakan dua sumber data yang berbeda sebagai objek analisis sentimen. Dataset pertama merupakan data opini masyarakat di media sosial X (sebelumnya Twitter) mengenai Program Makan Bergizi Gratis (MBG), sebuah inisiatif strategis pemerintah Indonesia untuk meningkatkan kualitas gizi anak-anak (RI, 2024). Pemilihan platform X didasarkan pada tingginya tingkat interaksi pengguna serta karakteristik data teksnya yang bersifat informal dan kaya akan representasi opini publik. Sebagai pembanding sekaligus penguat argumen performa algoritma, penelitian ini menggunakan dataset kedua yang terdiri dari ulasan produk makanan dan minuman di *e-commerce* Tokopedia, yang bersumber dari Kaggle. Dataset

ulasan produk ini memiliki struktur teks yang lebih spesifik pada pengalaman konsumen dan evaluasi produk, sehingga perbandingan kedua dataset ini dapat membuktikan konsistensi performa algoritma yang diusulkan dalam menangani keragaman teks bahasa Indonesia.

Pemilihan algoritma menjadi salah satu aspek penting dalam analisis sentimen karena menentukan sejauh mana data teks yang tidak terstruktur dapat diolah menjadi informasi yang akurat. Berbagai algoritma *machine learning* telah digunakan di beberapa penelitian dengan hasil performa yang beragam. Penelitian ini memfokuskan pada dua di antaranya, yaitu *Extreme gradient boosting* (XGBoost) dan *Long Short Term Memory* (LSTM), untuk melihat bagaimana keduanya bekerja dalam mengolah opini pengguna media sosial X mengenai Program MBG dan ulasan produk Tokopedia.

XGBoost merupakan algoritma berbasis pohon keputusan yang menggunakan teknik *boosting* untuk meningkatkan akurasi prediksi. Algoritma ini dirancang untuk efisiensi komputasi, mampu menangani data berskala besar, serta memiliki mekanisme regularisasi yang menjaga stabilitas performa (Wang et al., 2023). Keunggulan tersebut menjadikan XGBoost banyak digunakan dalam klasifikasi teks, khususnya pada data teks yang direpresentasikan dalam bentuk numerik seperti TF-IDF. Penelitian (Safitri et al., 2024) menunjukkan bahwa XGBoost lebih unggul dibandingkan SVM dalam klasifikasi sentimen publik terkait pemilihan presiden 2024, dengan akurasi tertinggi mencapai 78%. Hal ini sejalan dengan penelitian (Abdalrdha et al., 2023), yang mendapatkan hasil bahwa XGBoost mampu mencapai akurasi hingga 100% dalam mendeteksi konten kejahatan pada tweet.

Di sisi lain, *Long short-term memory* (LSTM) merupakan pengembangan dari *Recurrent Neural Network* (RNN) yang dirancang untuk mengatasi masalah *long-term dependency* pada data sekuensial. LSTM mampu mengingat informasi jangka panjang dan tetap memperbarui informasi baru, sehingga sangat cocok untuk memahami teks yang bersifat urutan kata. Dalam penelitian analisis sentimen opini publik terhadap kampanye boikot produk Israel, data teks digunakan diproses melalui teknik TF-IDF untuk memberikan bobot pada kata-kata penting. Penggunaan TF-IDF ini terbukti efektif dalam membantu LSTM mengidentifikasi

pola urutan yang merepresentasikan sentimen secara lebih akurat (Panggabean et al., 2025). Penelitian (Tampubolon, 2025) juga menunjukkan bahwa LSTM unggul dalam memprediksi sentimen positif dan negatif pada dataset yang lebih besar, seperti tweet terkait layanan transportasi online di Indonesia, dengan akurasi mencapai 82,15%. Hasil tersebut memperkuat bahwa LSTM lebih efektif dalam menangani data yang kompleks dan memiliki konteks sekuensial yang kuat.

Dalam perbandingan performa, penelitian (Pandian, 2025) mengevaluasi LSTM, BiLSTM, dan XGBoost menggunakan fitur TF-IDF pada teks berbahasa Tamil. Hasilnya menunjukkan bahwa masing-masing algoritma memiliki keunggulan karakteristik yang berbeda, XGBoost mencapai akurasi tertinggi sebesar 0,86 karena efisiensi pelatihannya, sementara LSTM dengan akurasi 0,75 tetap menonjol karena kemampuannya dalam memetakan pola hubungan antar fitur kata melalui arsitektur jaringan saraf.

Keunggulan dari masing-masing algoritma tidak menjamin model yang dibangun untuk analisis sentimen dapat memberikan hasil yang baik karena dalam praktiknya sering kali menghadapi beberapa tantangan, salah satunya terkait distribusi data yang tidak seimbang. Opini masyarakat bisa didominasi sentimen tertentu yang dapat mempengaruhi hasil analisis. Ketidakseimbangan ini dapat menurunkan performa model *machine learning* karena algoritma cenderung lebih bias terhadap label mayoritas (Imani et al., 2025). Untuk mengatasi permasalahan ini, berbagai teknik penyeimbangan data dikembangkan agar distribusi label menjadi lebih proporsional dan model dapat lebih sensitif terhadap pola dari label minoritas. Salah satu pendekatan yang banyak digunakan adalah *Synthetic Minority Oversampling Technique* (SMOTE), yang membangkitkan sampel sintetis melalui interpolasi antar data minoritas sehingga memperluas ruang keputusan model. Penelitian (Singgih, 2021) menunjukkan bahwa penerapan SMOTE dapat meningkatkan *f1-score*, akurasi dan AUC dibandingkan sebelum dilakukan penyeimbangan data. Di sisi lain, *Adaptive Synthetic Sampling* (ADASYN) menawarkan strategi lebih adaptif dengan memfokuskan penambahan sampel pada area yang sulit dipelajari yang juga dapat meningkatkan akurasi pada model LSTM dalam melakukan deteksi *hate speech* (Wenando et al., 2025).

Selain distribusi data, kualitas hasil analisis sentimen juga dipengaruhi oleh pengaturan *hyperparameter* pada model yang digunakan. *Hyperparameter* adalah parameter yang ditentukan sebelum proses pelatihan dimulai dan berperan penting dalam mengendalikan cara model belajar dari data. Tanpa penyesuaian yang tepat, model berisiko terlalu sederhana sehingga gagal menangkap pola atau terlalu kompleks sehingga mengalami *overfitting*. Mengatasi permasalahan ini, penelitian ini menggunakan metode *hyperparameter tuning* dengan dua pendekatan, yaitu *random search* dan *tree structured parzen estimator* (TPE). Kedua pendekatan ini telah terbukti dapat meningkatkan performa algoritma dalam melakukan klasifikasi. Penerapan *random search* dalam melakukan klasifikasi penyakit diabetes dengan algoritma XGBoost mampu meningkatkan akurasi dari 75% menjadi 95% (Abdurrahman et al., 2022). Di sisi lain, TPE bekerja sebagai proses iteratif yang memanfaatkan data historis dari *hyperparameter* yang telah dievaluasi untuk membangun model probabilistik menggunakan metode Bayesian yang telah terbukti dapat meningkatkan akurasi pada model LSTM dalam tugas analisis sentimen ulasan Aplikasi Recharge di Google Play Store jika dibandingkan dengan model tanpa TPE (Kautsar, 2025).

Berdasarkan uraian yang telah dipaparkan di atas, perbedaan karakteristik antara XGBoost dan LSTM ini menjadi menarik untuk dilakukan perbandingan dalam melakukan analisis sentimen. Penambahan skenario penyeimbangan data dan *hyperparameter tuning* diterapkan untuk memastikan perbandingan yang adil. Penelitian ini diharapkan dapat memberikan pemahaman mengenai kelebihan dan kekurangan masing-masing algoritma ketika dikombinasikan dengan berbagai skenario dalam menganalisis opini publik.

1.2 Perumusan Masalah

Berdasarkan latar belakang tersebut, maka dapat dirumuskan permasalahan yang akan dibahas dalam penelitian ini yaitu melakukan perbandingan performa antara algoritma *Extreme Gradient Boosting* (XGBoost) dan *Long Short-Term Memory* (LSTM) dalam mengenali sentimen positif, negatif, dan netral.

1.3 Tujuan Penelitian

Berdasarkan rumusan masalah, tujuan dari penelitian ini adalah untuk mengetahui algoritma dengan performa terbaik antara algoritma *Extreme Gradient*

Boosting (XGBoost) dan *Long Short-Term Memory* (LSTM) dalam mengenali sentimen positif, negatif, dan netral.

1.4 Pembatasan Masalah

Dari identifikasi masalah yang ada dalam penelitian ini, maka diperlukan pembatasan masalah agar dalam pengkajian yang dilakukan lebih terfokus dan tidak menyimpang dari masalah-masalah yang ingin dipecahkan. Berikut merupakan uraian batasan masalah dari penelitian ini:

1. Penelitian ini akan berfokus pada pengenalan tiga kategori sentimen utama, yaitu positif, negatif, dan netral.
2. Data yang digunakan tweet terkait makan bergizi gratis dari tanggal 6 Januari sampai 6 Februari 2025 dan data ulasan produk makanan dan minuman di Tokopedia.
3. Penyeimbangan data dilakukan dengan menggunakan *Synthetic Minority Oversampling Technique* (SMOTE) dan *Adaptive Synthetic Sampling (*ADASYN).
4. *Hyperparameter tuning* dilakukan dengan metode *Random search* dan *Tree-structured Parzen Estimator*(TPE).

1.5 Sistematika Penulisan

Sistematika penulisan penelitian ini disusun untuk mempermudah dalam mengetahui pembahasan yang ada pada skripsi ini. Adapun sistematika penulisan penelitian disusun dalam 5 (lima) bab yang terdiri dari Bab I Pendahuluan, Bab II Tinjauan Pustaka, Bab III Metodologi Penelitian, Bab IV Hasil dan Pembahasan, dan Bab V Penutup.

Bab I Pendahuluan adalah bab yang berisi latar belakang penelitian, rumusan masalah, tujuan penelitian, batasan masalah, dan sistematika penulisan.

Bab II Tinjauan Pustaka adalah bab yang landasan teori yang berkaitan dengan topik penelitian. Dalam hal ini adalah penelitian terkait yang telah dilakukan oleh peneliti-peneliti sebelumnya, teori terkait dengan *Natural language processing*, Analisis Sentimen, *Text preprocessing*, Pembobotan Kata, Penanganan Ketidakseimbangan Data, Algoritma XGBoost dan LSTM, *Stratified K-Fold* Cross *Validation*, Uji Signifikansi Statistik, dan *Confusion Matrix*.

Bab III Metodologi Penelitian adalah bab yang merincikan tahapan penelitian, mulai dari data yang digunakan, dan proses yang dilakukan sebelum pemodelan menggunakan algoritma XGBoost dan LSTM, hingga evaluasi model yang telah dibuat.

Bab IV Hasil dan Pembahasan adalah bab yang berisi hasil penjabaran terkait hasil dari implementasi tahapan metode penelitian pada bab III.

Bab V Penutup adalah bab yang berisi kesimpulan dari penelitian yang dilakukan dan saran untuk penelitian yang telah dilakukan sebagai bahan untuk pengembangan penelitian.