

**Analisis Sentimen Ulasan Aplikasi Shopee Pada Google Play Store  
Menggunakan Metode Bidirectional Long Short-Term Memory  
(BiLSTM)**

**Laporan Tugas Akhir**

Diajukan Untuk Memenuhi Persyaratan  
Guna Meraih Gelar Sarjana  
Informatika Universitas Muhammadiyah Malang



Nurul Hidayah  
(2020103703113)

**PROGRAM STUDI INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS MUHAMMADIYAH MALANG**

**2024**

## LEMBAR PERSETUJUAN

### **Analisis Sentimen Ulasan Aplikasi Shopee Pada Google Play Store Menggunakan Metode Bidirectional Long Short-Term Memory (BiLSTM)**

### **TUGAS AKHIR**

**Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1  
Informatika Universitas Muhammadiyah Malang**

Menyetujui,  
Malang, 20 Januari 2025

Dosen Pembimbing 1



**Christian Sri Kusuma Aditya S.Kom.,**

**M.Kom**

**NIP. 180327021991PNS.**

Dosen Pembimbing 2



**Nur Hayatin S. ST., M.Kom**

**NIP. 10809070476PNS.**

**LEMBAR PENGESAHAN**

**Analisis Sentimen Ulasan Aplikasi Shopee Pada Google Play Store  
Menggunakan Metode Bidirectional Long Short-Term Memory  
(BiLSTM)**

**TUGAS AKHIR**

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1  
Informatika Universitas Muhammadiyah Malang

Disusun Oleh :

**NURUL HIDDAYAH**

**202010370311361**

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis penguji  
pada tanggal 20 Januari 2025

Menyetujui,

Dosen Penguji 1



**Ir. Yufis Azhar S.Kom., M.Kom.**

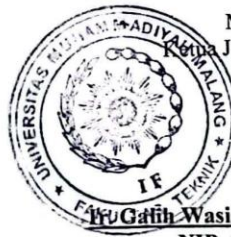
**NIP. 10814100544PNS.**

Dosen Penguji 2



**Vinna Rahmavanti S S.Si., M.Si**

**NIP. 180306071990PNS.**



Mengetahui,  
Ketua Jurusan Informatika



**Yudi Gafit Wasis Wicaksono S.kom. M.Cs.**

**NIP. 10814100541PNS.**

## LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

**NAMA : NURUL HIDDAYAH**

**NIM : 202010370311361**

**FAK./JUR. : Informatika**

Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul “**Analisis Sentimen Ulasan Aplikasi Shopee Pada Google Play Store Menggunakan Metode Bidirectional Long Short-Term Memory (BiLSTM)**” beserta seluruh isinya adalah karya saya sendiri dan bukan merupakan karya tulis orang lain, baik sebagian maupun seluruhnya, kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya ini, atau ada klaim dari pihak lain terhadap keaslian karya saya ini maka saya siap menanggung segala bentuk resiko/sanksi yang berlaku.

Mengetahui,  
Dosen Pembimbing



Christian Sri Kusuma Aditya S.Kom.,  
M.Kom

Malang, 20 Januari 2025  
Yang Membuat Pernyataan



NURUL HIDDAYAH

## ABSTRAK

Analisis sentimen ulasan aplikasi di Google Play Store memiliki tantangan dalam menangkap pola kompleks bahasa alami, terutama pada ulasan berbahasa Indonesia yang sering kali informal dan tidak terstruktur. Model deep learning seperti BiLSTM memerlukan teknik ekstraksi fitur yang efektif untuk meningkatkan akurasi prediksi sentimen. Penelitian ini menggunakan model Bidirectional Long Short-Term Memory (BiLSTM) dengan teknik Word2Vec untuk ekstraksi fitur, memanfaatkan data ulasan aplikasi Shopee dari Google Play Store. Data diproses melalui tahapan preprocessing, meliputi cleaning, tokenizing, normalization, stopword removal, dan stemming, diikuti oleh pelatihan model dengan berbagai kombinasi parameter seperti batch size dan learning rate. Hasil eksperimen menunjukkan bahwa penggunaan batch size 32 dengan learning rate 0.001 memberikan akurasi tertinggi, yaitu 75%, sementara batch size 64 dengan learning rate yang sama juga menghasilkan akurasi 75%. Penurunan learning rate menjadi 0.00001 pada kedua batch size menyebabkan penurunan akurasi yang signifikan, dengan akurasi 68% untuk batch size 32 dan 63% untuk batch size 64. Secara keseluruhan, penggunaan Word2Vec sebagai teknik ekstraksi fitur terbukti efektif dalam meningkatkan kemampuan model dalam menangkap sentimen positif, negatif dan netral dari ulasan aplikasi Shopee. Temuan ini memberikan kontribusi pada pengembangan model analisis sentimen berbasis deep learning untuk aplikasi e-commerce, khususnya dalam konteks aplikasi Shopee.

**Kata Kunci:** Analisis Sentimen, Ulasan Aplikasi, BiLSTM, Word2Vec, Deep Learning.

## ABSTRACT

Sentiment analysis of application reviews on the Google Play Store faces challenges in capturing the complex patterns of natural language, especially in Indonesian reviews that are often informal and unstructured. Deep learning models such as BiLSTM require effective feature extraction techniques to improve sentiment prediction accuracy. This study employs a Bidirectional Long Short-Term Memory (BiLSTM) model with Word2Vec as a feature extraction technique, utilizing review data of the Shopee application collected from the Google Play Store. The data undergoes preprocessing stages, including cleaning, tokenizing, normalization, stopword removal, and stemming, followed by model training with various parameter combinations such as batch size and learning rate. The experimental results show that using a batch size of 32 with a learning rate of 0.001 yields the highest accuracy of 75%, while a batch size of 64 with the same learning rate also results in an accuracy of 75%. Reducing the learning rate to 0.00001 for both batch sizes significantly decreased the accuracy, with the lowest accuracy recorded at 68% to 63% for batch size 32 and 64, respectively. Overall, the use of Word2Vec as a feature extraction technique proved effective in enhancing the model's ability to capture positive, negative, and neutral sentiments from Shopee app reviews. Our finding contributes to the development of sentiment analysis models based on deep learning for e-commerce applications, particularly in the context of Shopee app reviews.

**Keywords:** Sentiment Analysis, App Reviews, Shopee, BiLSTM, Word2Vec, Deep Learning.

## KATA PENGANTAR

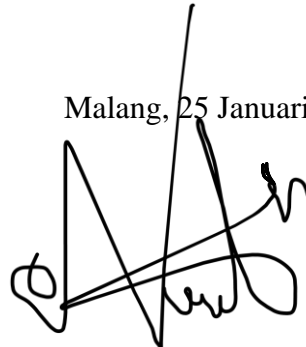
Dengan memanjatkan puji syukur kehadirat Allah SWT. Atas limpahan rahmat dan hidayah-NYA sehingga peneliti dapat menyelesaikan tugas akhir yang berjudul:

**”ANALISIS SENTIMEN ULASAN APLIKASI SHOPEE PADA GOOGLE PLAY STORE MENGGUNAKAN METODE BIDIRECTIONAL LONG SHORT-TERM MEMORY (BILSTM)”**

Di dalam tulisan ini disajikan pokok-pokok bahasa yang meliputi analisis sentimen ulasan pengguna aplikasi Shopee, metode pemrosesan teks, dan implementasi model BiLSTM untuk klasifikasi sentimen. Setiap tahap penelitian dijelaskan secara rinci, mulai dari pengumpulan data ulasan melalui scraping hingga evaluasi model dengan metrik akurasi dan performa lainnya.

Peneliti menyadari sepenuhnya bahwa dalam penulisan tugas akhir ini masih banyak kekurangan dan keterbatasan. Oleh karena itu peneliti mengharapkansaran yang membangun agar tulisan ini bermanfaat bagi perkembangan ilmupengetahuan.

Malang, 25 Januari 2025



Nurul Hidayah

## DAFTAR ISI

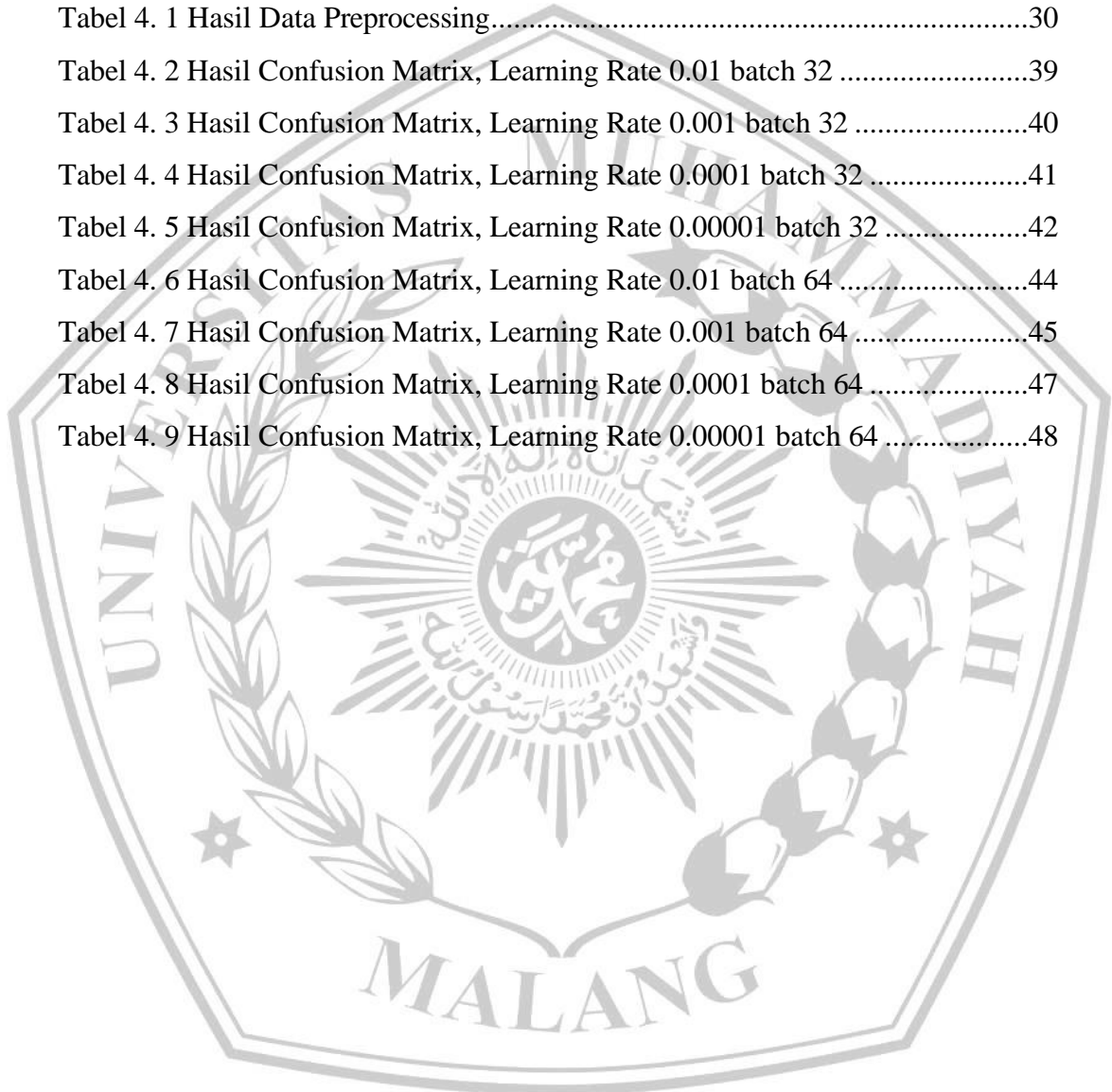
<b>LEMBAR PERSETUJUAN</b> .....	<b>ii</b>
<b>LEMBAR PENGESAHAN</b> .....	<b>iii</b>
<b>LEMBAR PERNYATAAN</b> .....	<b>iv</b>
<b>ABSTRAK</b> .....	<b>v</b>
<b>ABSTRACT</b> .....	<b>vi</b>
<b>LEMBAR PERSEMBAHAN</b> .....	<b>vii</b>
<b>KATA PENGANTAR</b> .....	<b>1</b>
<b>DAFTAR ISI</b> .....	<b>2</b>
<b>DAFTAR TABEL</b> .....	<b>4</b>
<b>DAFTAR GAMBAR</b> .....	<b>5</b>
<b>BAB 1 PENDAHULUAN</b> .....	<b>6</b>
1.1 Latar Belakang .....	6
1.2 Rumusan Masalah .....	8
1.3 Tujuan Penelitian .....	8
1.4 Batasan Masalah .....	9
<b>BAB 2 TINJAUAN PUSTAKA</b> .....	<b>10</b>
2.1 Google PlayStore .....	10
2.2 Analisis Sentimen .....	10
2.3 Oversampling .....	11
2.4 Word2Vec .....	12
2.5 BiLSTM .....	13
2.6 Kajian Penelitian Terdahulu .....	15



<b>BAB 3 METODOLOGI PENELITIAN.....</b>	<b>20</b>
3.1    Jenis Penelitian .....	20
3.2    Tahapan Penelitian .....	20
3.2.1    Data Scraping .....	20
3.2.2    Data Preprocessing .....	21
3.2.3    Data Labelling .....	22
3.2.4    Visualisasi Data .....	24
3.2.5    Oversampling .....	24
3.2.6    Data Splitting .....	25
3.2.7    Word2Vec .....	25
3.2.8    Model BiLSTM .....	27
3.2.9    Evaluasi .....	27
<b>BAB 4 HASIL DAN PEMBAHASAN .....</b>	<b>30</b>
4.1    Data Preprocessing .....	30
4.2    Visualisasi Data .....	33
4.3    Oversampling .....	35
4.4    Word2vec .....	35
4.5    Model BiLSTM .....	36
4.6    Evaluasi .....	37
4.7    Analisis Hasil .....	48
<b>BAB 5 KESIMPULAN .....</b>	<b>50</b>
<b>DAFTAR PUSTAKA .....</b>	<b>51</b>

## DAFTAR TABEL

Tabel 2. 1 Penelitian terdahulu tentang analisis sentimen menggunakan BiLSTM .....	15
Tabel 3. 1 Tabel berikut menyajikan contoh ulasan dan skor yang mencerminkan masing-masing kategori sentimen.....	23
Tabel 4. 1 Hasil Data Preprocessing.....	30
Tabel 4. 2 Hasil Confusion Matrix, Learning Rate 0.01 batch 32 .....	39
Tabel 4. 3 Hasil Confusion Matrix, Learning Rate 0.001 batch 32 .....	40
Tabel 4. 4 Hasil Confusion Matrix, Learning Rate 0.0001 batch 32 .....	41
Tabel 4. 5 Hasil Confusion Matrix, Learning Rate 0.00001 batch 32 .....	42
Tabel 4. 6 Hasil Confusion Matrix, Learning Rate 0.01 batch 64 .....	44
Tabel 4. 7 Hasil Confusion Matrix, Learning Rate 0.001 batch 64 .....	45
Tabel 4. 8 Hasil Confusion Matrix, Learning Rate 0.0001 batch 64 .....	47
Tabel 4. 9 Hasil Confusion Matrix, Learning Rate 0.00001 batch 64 .....	48



## DAFTAR GAMBAR

Gambar 2. 1 Logo Google PlayStore .....	10
Gambar 2. 2 Arsitekture Word2Vec [19] .....	12
Gambar 2. 3 Arsitektur BiLSTM [16].....	14
Gambar 3. 1 Tahapan Penelitian .....	20
Gambar 3. 2 Data ulasan Shopee .....	21
Gambar 3. 3 Tahapan data preprocessing.....	22
Gambar 3. 4 Visualisasi data.....	24
Gambar 3. 5 Arsitektur BiLSTM [16].....	27
Gambar 4. 1 10 Kata Teratas.....	33
Gambar 4. 2 Word Cloud .....	34
Gambar 4. 3 Distribusi sentimen ulasan.....	34
Gambar 4. 4 Distribusi label setelah oversampling .....	35
Gambar 4. 5 Hasil Word2vec .....	36
Gambar 4. 7 Syntax Implementasi BiLSTM.....	36
Gambar 4. 8 Confusion Matrix Batch Size 32, Learning Rate 0.01.....	38
Gambar 4. 9 Confusion Matrix Batch Size 32, Learning Rate 0.001.....	40
Gambar 4. 10 Confusion Matrix Batch Size 32, Learning Rate 0.0001.....	41
Gambar 4. 11 Confusion Matrix Batch Size 32, Learning Rate 0.00001.....	42
Gambar 4. 12 Confusion Matrix Batch Size 64, Learning Rate 0.01.....	44
Gambar 4. 13 Confusion Matrix Batch Size 64, Learning Rate 0.001.....	45
Gambar 4. 14 Confusion Matrix Batch Size 64, Learning Rate 0.0001.....	46
Gambar 4. 15 Confusion Matrix Batch Size 64, Learning Rate 0.00001.....	47

## DAFTAR PUSTAKA

- [1] E. F. Santika, "10 Negara Dengan Proyeksi Pertumbuhan E-Commerce Tertinggi Global 2024 (2024)." Accessed: Aug. 13, 2024. [Online]. Available: <https://databoks.katadata.co.id/datapublish/2024/04/29/ecdb-proyeksi-pertumbuhan-e-commerce-indonesia-tertinggi-sedunia-pada-2024>
- [2] W. Huang, M. Lin, And Y. Wang, "Sentiment Analysis Of Chinese E-Commerce Product Reviews Using Ernie Word Embedding And Attention Mechanism," *Applied Sciences*, Vol. 12, No. 14, 2022, Doi: 10.3390/App12147182.
- [3] R. Jihan Syanita And A. Ghina, "Enrichment: Journal Of Management Influence Of Factor Application Between Utilizing Modified United Theory Of Acceptance And Use Of The Technology 2 Model (A Study On Shopee Users In Indonesia)," 2023. Doi: <https://doi.org/10.35335/Enrichment.V13i1.1245>.
- [4] M. Hadwan, M. Al-Sarem, F. Saeed, And M. A. Al-Hagery, "An Improved Sentiment Classification Approach For Measuring User Satisfaction Toward Governmental Services' Mobile Apps Using Machine Learning Methods With Feature Engineering And Smote Technique," *Applied Sciences*, Vol. 12, No. 11, 2022, Doi: 10.3390/App12115547.
- [5] L. Budi Ilmawan And E. Winarko, "Aplikasi Mobile Untuk Analisis Sentimen Pada Google Play," *Ijccs*, Vol. 9, No. 1, Pp. 53–64, 2015.
- [6] N. Z. B. Jannah And K. Kusnawi, "Comparison Of Naïve Bayes And Svm In Sentiment Analysis Of Product Reviews On Marketplaces," *Sinkron*, Vol. 8, No. 2, Pp. 727–733, Mar. 2024, Doi: 10.33395/Sinkron.V8i2.13559.
- [7] A. K. Ni Komang, I. M. A. D. Suarjaya, And I. M. S. Raharja, "Classification Of Public Figures Sentiment On Twitter Using Big Data Technology," *Journal Of Informatics And Telecommunication Engineering*, Vol. 6, No. 1, Pp. 157–169, Jul. 2022, Doi: 10.31289/Jite.V6i1.7329.

- [8] N. Afrianto, D. H. Fudholi, And S. Rani, “Prediksi Harga Saham Menggunakan Bilstm Dengan Faktor Sentimen Publik,” *Jurnal Resti (Rekayasa Sistem Dan Teknologi Informasi)*, Vol. 6, No. 1, Pp. 41–46, Feb. 2022, Doi: 10.29207/Resti.V6i1.3676.
- [9] F. Masri, D. Saepudin, And D. Adytia, “Forecasting Of Sea Level Time Series Using Deep Learning Rnn, Lstm, And Bilstm, Case Study In Jakarta Bay, Indonesia,” Bandung, Aug. 2020. Accessed: Aug. 13, 2024. [Online]. Available: [https://Openlibrary.Telkomuniversity.Ac.Id/Pustaka/Files/161205/Jurnal\\_Eproc/Peramalan-Seri-Waktu-Permukaan-Laut-Menggunakan-Deep-Learning-Rnn-Lstm-Dan-Bilstm-Studi-Kasus-Di-Teluk-Jakarta-Indonesia.Pdf](https://Openlibrary.Telkomuniversity.Ac.Id/Pustaka/Files/161205/Jurnal_Eproc/Peramalan-Seri-Waktu-Permukaan-Laut-Menggunakan-Deep-Learning-Rnn-Lstm-Dan-Bilstm-Studi-Kasus-Di-Teluk-Jakarta-Indonesia.Pdf)
- [10] X. Zhang, X. Wang, H. Li, S. Sun, And F. Liu, “Monthly Runoff Prediction Based On A Coupled Vmd-Ssa-Bilstm Model,” *Sci Rep*, Vol. 13, No. 1, P. 13149, 2023, Doi: 10.1038/S41598-023-39606-4.
- [11] E. A. Pradana, S. T. Shaufiah, N. Selviandro, S. Kom, And M. Kom, “Feature-Based Opinion Menggunakan Algoritma High Adjective Count Dan Max Opinion Score (Studi Kasus Review Di Google Play).”
- [12] B. Setiawan, “Analisis Sentimen Multilabel Terhadap Ujaran Kebencian Menggunakan Metode Indo Bert Lite Dan Bidirectional Lstm-Cnn Dengan Grid Search Hyperparameter Optimization,” Uin Syarif Hidayatullah, Jakarta, 2022.

- [13] S. Mutmainah, D. H. Fudholi, And S. Hidayat, “Analisis Sentimen Dan Pemodelan Topik Aplikasi Telemedicine Pada Google Play Menggunakan Bilstm Dan Lda,” *Jurnal Media Informatika Budidarma*, Vol. 7, No. 1, P. 312, Jan. 2023, Doi: 10.30865/Mib.V7i1.5486.
- [14] J. Liu, “Sentiment Classification Of Social Network Text Based On At-Bilstm Model In A Big Data Environment,” *International Journal Of Information Technologies And Systems Approach*, Vol. 16, No. 2, 2023, Doi: 10.4018/Ijitsa.324808.
- [15] P. Ayuningtyas *Et Al.*, “Perbandingan Metode Word2vec Model Skipgram Pada Ulasan Aplikasi Linkaja Menggunakan Algoritma Bidirectional Lstm Dan Support Vector Machine,” Vol. 12, No. 1, 2024, Doi: 10.26418/Justin.V12i1.72530.
- [16] D. I. Puteri, “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*, Vol. 11, No. 1, Pp. 35–43, May 2023, Doi: 10.34312/Euler.V11i1.19791.
- [17] S. Sandiwarno, “Penerapan Machine Learning Untuk Prediksi Bencana Banjir,” *Jurnal Sistem Informasi Bisnis*, Vol. 14, No. 1, Pp. 62–76, Jan. 2024, Doi: 10.21456/Vol14iss1pp62-76.
- [18] Alvi Rahmy Royyan And Erwin Budi Setiawan, “Feature Expansion Word2vec For Sentiment Analysis Of Public Policy In Twitter,” *Jurnal Resti*

(*Rekayasa Sistem Dan Teknologi Informasi*), Vol. 6, No. 1, Pp. 78–84, Feb. 2022, Doi: 10.29207/Resti.V6i1.3525.

- [19] D. Jatnika, M. A. Bijaksana, And A. A. Suryani, “Word2vec Model Analysis For Semantic Similarities In English Words,” In *Procedia Computer Science*, Elsevier B.V., 2019, Pp. 160–167. Doi: 10.1016/J.Procs.2019.08.153.
- [20] M. Nina Adlini, A. Hanifa Dinda, S. Yulinda, O. Chotimah, And S. Julia Merliyana, “Metode Penelitian Kualitatif Studi Pustaka,” 2022.
- [21] D. Chicco, N. Tötsch, And G. Jurman, “The Matthews Correlation Coefficient (Mcc) Is More Reliable Than Balanced Accuracy, Bookmaker Informedness, And Markedness In Two-Class Confusion Matrix Evaluation,” *Biodata Min*, Vol. 14, No. 1, P. 13, 2021, Doi: 10.1186/S13040-021-00244-Z.
- [22] Wen, H., & Zhao, J. (2023). *Sentiment Analysis Model of Imbalanced Comment Texts Based on BiLSTM*. <https://doi.org/10.21203/rs.3.rs-2434519/v1>



UNIVERSITAS  
MUHAMMADIYAH  
MALANG

# FAKULTAS TEKNIK

## INFORMATIKA

informatika.umm.ac.id | informatika@umm.ac.id



### FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : Nurul Hidayah  
 NIM : 202010370311361  
 Judul TA : Analisis Sentimen Ulasan Aplikasi Shopee Pada Google Play Store Menggunakan Metode Bidirectional Long Short-Term Memory (BILSTM)

#### Hasil Cek Plagiarisme dengan Turnitin

No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%) *
1.	Bab 1 – Pendahuluan	10 %	6%
2.	Bab 2 – Daftar Pustaka	25 %	9%
3.	Bab 3 – Analisis dan Perancangan	25 %	6%
4.	Bab 4 – Implementasi dan Pengujian	15 %	7%
5.	Bab 5 – Kesimpulan dan Saran	5 %	4%
6.	Makalah Tugas Akhir	20%	10%

*\*) Hasil cek plagiarisme diisi oleh pemeriksa (staf TU)*

*\*) Maksimal 5 kali (4 Kali sebelum ujian, 1 kali sesudah ujian)*

Mengetahui,

Pemeriksa (Staff TU)

*[Signature]*  
 (.....)



Kampus I  
 Jl. Bestung 1 Malang, Jawa Timur  
 P. +62 341 531 253 (Malang)  
 F. +62 341 466 435

Kampus II  
 Jl. Bendungan Sutam No 168 Malang, Jawa Timur  
 P. +62 341 551 149 (Malang)  
 F. +62 341 582 060

Kampus III  
 Jl. Raya Tlogomas No 248 Malang, Jawa Timur  
 P. +62 341 464 318 (Malang)  
 F. +62 341 466 435  
 E. webmaster@umm.ac.id