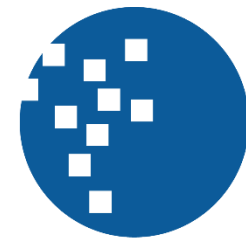


**ANALISA SENTIMEN APLIKASI *MOBILE BANKING* BCA DI
PLATFORM GOOGLE PLAY MENGGUNAKAN *FINE-
TUNING* DISTILBERT**



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

SKRIPSI

Christopher Darren

00000054804

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2025**

**ANALISA SENTIMEN APLIKASI MOBILE BANKING BCA DI
PLATFORM GOOGLE PLAY MENGGUNAKAN *FINE-
TUNING* DISTILBERT**



Diajukan sebagai Salah Satu Syarat untuk Memperoleh

Gelar Sarjana Komputer

Christopher Darren

00000054804

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG**

2025

HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Christopher Darren

Nomor Induk Mahasiswa : 00000054804

Program Studi : Sistem Informasi

Skripsi dengan judul:

ANALISA SENTIMEN APLIKASI MOBILE BANKING BCA DI PLATFORM GOOGLE PLAY MENGGUNAKAN FINE-TUNING DISTILBERT Merupakan hasil karya saya sendiri bukan plagiat dari laporan karya tulis ilmiah yang ditulis oleh orang lain, dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar serta dicantumkan di Daftar Pustaka.

Jika di kemudian hari terbukti ditemukan kecurangan/penyimpangan, baik dalam pelaksanaan maupun dalam penulisan laporan karya tulis ilmiah, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk mata kuliah yang telah saya tempuh.

Tangerang, 4 Juni 2025



Christopher Darren

HALAMAN PENGESAHAN

Skripsi dengan judul

ANALISA SENTIMEN APLIKASI MOBILE BANKING BCA DI PLATFORM
GOOGLE PLAY MENGGUNAKAN FINE-TUNING DISTILBERT

Oleh

Nama : Christopher Darren
NIM : 00000054804
Program Studi : Sistem Informasi
Fakultas : Teknik dan Informatika

Telah diujikan pada hari Jumat, 13 Juni 2025

Pukul 10.00 s.d 12.00 dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut.

Ketua Sidang
07/03/2025

Prof. Friska Natalia, Ph.D.,
0306128307

Penguji

Dr. Erick Fernando, S.Kom., M.S.I.,
01029118501

Pembimbing

Ir. Raymond Sinarhi Oetama,
M.CIS_328046803

Ketua Program Studi Sistem Informasi

Ririn Ikana Desanti, S.Kom., M.Kom
0313058001

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH

Yang bertanda tangan di bawah ini:

Nama : Christopher Darren

NIM : 00000054804

Program Studi : Sistem Informasi

Jenjang : S1

Judul Karya Ilmiah : ANALISA SENTIMEN APLIKASI
MOBILE BANKING BCA DI PLATFORM GOOGLE PLAY
MENGUNAKAN FINE-TUNING DISTILBERT

Menyatakan dengan sesungguhnya bahwa saya bersedia* (**pilih salah satu**):

- ☒ Saya bersedia memberikan izin sepenuhnya kepada Universitas Multimedia Nusantara untuk mempublikasikan hasil karya ilmiah saya ke dalam repositori Knowledge Center sehingga dapat diakses oleh Sivitas Akademika UMN/Publik. Saya menyatakan bahwa karya ilmiah yang saya buat tidak mengandung data yang bersifat konfidensial.
- ☐ Saya tidak bersedia mempublikasikan hasil karya ilmiah ini ke dalam repositori Knowledge Center, dikarenakan: dalam proses pengajuan publikasi ke jurnal/konferensi nasional/internasional (dibuktikan dengan *letter of acceptance*) **.
- ☐ Lainnya, pilih salah satu:
 - ☐ Hanya dapat diakses secara internal Universitas Multimedia Nusantara
 - ☐ Embargo publikasi karya ilmiah dalam kurun waktu 3 tahun.

Tangerang, 9 Juli 2025



(Christopher Darren)

* Pilih salah satu

** Jika tidak bisa membuktikan LoA jurnal/HKI, saya bersedia mengizinkan penuh karya ilmiah saya untuk dipublikasikan ke KC UMN dan menjadi hak institusi UMN.

KATA PENGANTAR

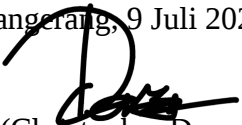
Puji dan syukur saya panjatkan kepada Tuhan Yang Maha Esa karena berkat dan rahmat-Nya saya dapat menyelesaikan pembuatan skripsi ini dan menjadi acuan untuk memperluas wawasan, pengetahuan, dan mengembangkan kemampuan penulis dalam menganalisis sebuah data yang berguna bagi saya serta orang banyak. Tujuan pembuatan laporan akhir ini adalah untuk membantu memberikan informasi dan pengetahuan mengenai penerapan *fine-tuning* dan *hyperparameter tuning transformer model* menggunakan Optuna khususnya pada algoritma DistilBERT.

Dalam menyelesaikan laporan ini penulis banyak mendapatkan dukungan, bimbingan, pengarahan, dan bantuan dari berbagai pihak. Pada kesempatan ini penulis ingin menyampaikan terima kasih kepada:

1. Bapak Dr. Ir. Andrey Andoko, selaku Rektor Universitas Multimedia Nusantara.
2. Bapak Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
3. Ibu Ririn Ikana Desanti, S.Kom., M.Kom., selaku Ketua Program Studi Sistem Informasi Universitas Multimedia Nusantara.
4. Bapak Ir. Raymond Sunardi Oetama, M.CIS, sebagai Pembimbing pertama yang telah memberikan bimbingan, arahan, dan motivasi atas terselesainya skripsi ini.
5. Keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan skripsi ini dengan lancar.
6. Teman teman jurusan Sistem Informasi yang memberikan dukungan dan semangat kepada penulis sehingga skripsi ini bisa selesai dengan baik.

Semoga skripsi ini dapat memberikan wawasan dan ilmu yang berguna bagi peneliti dan para pembaca semuanya serta menjadi referensi penelitian kedepannya. Penulis juga menyadari bahwa skripsi ini jauh dari kata sempurna, oleh karena itu peneliti mohon maaf sebesar-besarnya apabila terjadi kesalahan kata, ejaan dalam penulisan, dan isi dari skripsi ini. Peneliti berharap dengan adanya skripsi ini dapat memberikan dampak positif kedalam bidang yang relevan. Terima kasih.

Tangerang, 9 Juli 2025



(Christopher Darren)



UMN

UNIVERSITAS
MULTIMEDIA
NUSANTARA

ANALISA SENTIMEN APLIKASI MOBILE BANKING BCA DI PLATFORM GOOGLE PLAY MENGGUNAKAN *FINE- TUNING* DISTILBERT

Christopher Darren

ABSTRAK

Di era transformasi digital, layanan mobile banking seperti BCA Mobile menjadi pilihan utama masyarakat dalam melakukan transaksi secara cepat dan fleksibel. Namun, banyaknya keluhan dari pengguna dapat memengaruhi citra aplikasi. Ulasan di Google Play Store menjadi sumber penting untuk memahami pengalaman pengguna, namun analisis manual terhadap data berskala besar tidak efisien. Penelitian ini bertujuan menerapkan analisis sentimen otomatis berbasis Deep Learning untuk mengevaluasi persepsi pengguna terhadap aplikasi BCA Mobile.

Model DistilBERT digunakan karena mampu memahami konteks bahasa secara mendalam dengan arsitektur yang lebih ringan dibandingkan BERT. Dataset ulasan diproses melalui tahap pra-pemrosesan dan pelabelan rule-based, kemudian dilakukan fine-tuning terhadap model pretrained DistilBERT agar sesuai dengan konteks bahasa Indonesia dan label sentimen positif, netral, serta negatif. Untuk menangani ketidakseimbangan data, diterapkan metode oversampling. Evaluasi dilakukan menggunakan classification report dan confusion matrix untuk mengukur akurasi, presisi, recall, dan f1-score.

Model akhir diuji menggunakan dataset pengujian dan diimplementasikan ke dalam prototipe berbasis web menggunakan platform Streamlit. Hasil menunjukkan bahwa DistilBERT menghasilkan akurasi dan performa yang baik dalam klasifikasi sentimen serta lebih efisien dibanding metode tradisional. Proses hyperparameter tuning juga meningkatkan efektivitas model. Penelitian ini memberikan kontribusi dalam pemanfaatan teknologi Deep Learning untuk analisis ulasan pengguna dan dapat menjadi dasar pengambilan keputusan dalam peningkatan kualitas layanan BCA Mobile.

Kata kunci: Analisa Sentimen, *DistilBERT*, Fine-Tuning, Streamlit, Ulasan

SENTIMENT ANALYSIS OF BCA MOBILE BANKING APPLICATION ON GOOGLE PLAY PLATFORM USING DISTILBERT FINE-TUNING

Christopher Darren

ABSTRACT (English)

In the era of digital transformation, mobile banking services such as BCA Mobile have become the primary choice for people to conduct transactions quickly and flexibly. However, numerous complaints from users can affect the image of the application. Reviews on Google Play Store are an important source for understanding user experience, but manual analysis of large-scale data is inefficient. This study aims to apply automatic sentiment analysis based on Deep Learning to evaluate user perceptions of the BCA Mobile application.

The DistilBERT model was chosen because it can understand language context deeply with a lighter architecture compared to BERT. The review dataset was processed through pre-processing and rule-based labeling stages, followed by fine-tuning the pretrained DistilBERT model to align with the Indonesian language context and positive, neutral, and negative sentiment labels. To address data imbalance, the oversampling method was applied. Evaluation was conducted using a classification report and confusion matrix to measure accuracy, precision, recall, and F1-score.

The final model was tested using a test dataset and implemented into a web-based prototype using the Streamlit platform. The results show that DistilBERT achieves good accuracy and performance in sentiment classification and is more efficient than traditional methods. The hyperparameter tuning process also enhances the model's effectiveness. This research contributes to the application of Deep Learning technology for user review analysis and can serve as a basis for decision-making in improving the quality of BCA Mobile services.

Keywords: *DistilBERT, Fine-Tuning, Review, Sentiment Analysis, Streamlit*

DAFTAR ISI

HALAMAN PERNYATAAN TIDAK PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH	iv
KATA PENGANTAR	v
ABSTRAK	vii
ABSTRACT (<i>English</i>)	viii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiii
DAFTAR RUMUS	xvi
DAFTAR LAMPIRAN	xvii
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	6
1.3 Batasan Masalah	6
1.4 Tujuan dan Manfaat Penelitian	7
1.4.1 Tujuan Penelitian	7
1.4.2 Manfaat Penelitian	7
1.5 Sistematika Penulisan	8
BAB II LANDASAN TEORI	10
2.1 Penelitian Terdahulu	10
2.2 Tinjauan Teori	13
2.2.1 Ulasan	13
2.2.2 Analisis Sentimen	14
2.2.3 <i>Text Mining</i>	14
2.2.4 <i>Heuristic-based labeling</i>	15
2.2.5 <i>Latent Dirichlet Allocation</i>	15
2.2.6 <i>Transformer Model</i>	17
2.3 Teori tentang <i>Framework</i> , <i>Algoritma</i> , dan <i>Metode Evaluasi</i>	19
2.3.1 <i>Framework</i>	19

2.3.1.1	CRISP-DM	19
2.3.1.2	Optuna.....	21
2.3.2	Algoritma.....	21
2.3.2.1	BERT	21
2.3.2.2	DISTILBERT.....	24
2.3.2.3	LSTM.....	26
2.3.3	Metrik Evaluasi.....	27
2.3.3.1	<i>Accuracy</i>	28
2.3.3.2	<i>Precision</i>	29
2.3.3.3	<i>F1-Score</i>	29
2.3.3.4	<i>Recall</i>	29
2.4	Tools Penelitian.....	29
2.4.1	Python	30
2.4.2	Jupyter Notebook.....	31
2.4.3	Visual Studio Code.....	31
2.4.4	HuggingFace	32
2.4.5	Streamlit.....	33
2.4.6	Windows.....	34
BAB III METODOLOGI PENELITIAN		35
3.1	Gambaran Umum Objek Penelitian.....	35
3.2	Metode Penelitian	36
3.3	Teknik Pengumpulan Data.....	51
3.4	Teknik Analisis Data	53
BAB IV ANALISIS DAN HASIL PENELITIAN		55
4.1	<i>Business Understanding</i>	55
4.2	<i>Data Understanding</i>	57
4.2.1	<i>Exploration Data Analysis (EDA)</i>	60
4.3	<i>Data Preparation</i>	75
4.3.1	<i>Checking Missing Values</i>	76
4.3.2	<i>Rule-based labelling</i>	77
4.3.3	<i>Balancing data</i>	78
4.3.4	<i>Split data</i>	81

4.4	<i>Modeling</i>	81
4.4.1	<i>Model Tuning dengan LSTM</i>	87
4.4.2	<i>Model Fine-Tuning dengan Distilbert</i>	88
4.4.3	<i>Model Hyperparameter tuning dengan Distilbert-Optuna</i>	90
4.5	<i>Evaluation</i>	93
4.5.1	<i>Model Fine-tuning dengan LSTM</i>	93
4.5.2	<i>Model Fine-tuning dengan DistilBert</i>	95
4.5.3	<i>Model Hyperparameter tuning Distilbert dengan Optuna</i>	96
4.6	<i>Deployment</i>	97
4.6.1	<i>Saving Model</i>	98
4.6.2	<i>Pengembangan website</i>	99
4.6.3	<i>Calling Function</i>	100
4.7	Hasil Penelitian	102
4.8	Pembahasan hasil penelitian	103
BAB V SIMPULAN DAN SARAN		109
5.1	Simpulan	109
5.2	Saran	110
DAFTAR PUSTAKA		111
LAMPIRAN		121



DAFTAR TABEL

Tabel 2. 1 Tabel Penelitian Terdahulu	10
Tabel 3. 1 Tabel dari berbagai Perbandingan framework dalam data mining	45
Tabel 3. 2 Tabel alasan algoritma terpilih.....	48
Tabel 3. 3 Tabel Kolom aplikasi BCA Mobile pada Google Playstore	51
Tabel 4. 1 Teks representasi dari kata kunci negatif yang sering dibahas	69
Tabel 4. 2 Teks representasi dari kata kunci positif yang sering dibahas	71
Tabel 4. 3 Spesifikasi perangkat yang digunakan dalam penelitian	82
Tabel 4. 4 Skenario Parameter yang digunakan untuk proses train model Distilbert	82
Tabel 4. 5 Parameter yang digunakan pada model LSTM.....	87
Tabel 4. 6 Parameter yang digunakan pada model Distilbert	88
Tabel 4. 7 Tabel hasil perbandingan model antar model	91
Tabel 4. 8 Perbandingan hasil dengan penelitian terdahulu dan penelitian serupa	103



DAFTAR GAMBAR

Gambar 1. 1 Chart urutan mobile banking terpopuler di Indonesia tahun 2024.....	3
Gambar 2. 1 Komponen dari LDA.....	16
Gambar 2. 2 Transformer model Architecture	17
Gambar 2. 3 Alur kerangka CRISP-DM	19
Gambar 2. 4 Gambar representasi input cara kerja tokenisasi BERT	22
Gambar 2. 5 Bagan BERT yang terpecah ke tahap pre training dan fine tuning.	22
Gambar 2. 6 Arsitektur model DistilBERT dan komponennya	24
Gambar 2. 7 Arsitektur model LSTM dan komponennya	26
Gambar 2. 8 Confusion Matrix	28
Gambar 2. 9 Logo bahasa pemrograman Python	30
Gambar 2. 10 Logo Jupyter Notebook	31
Gambar 2. 11 Logo Visual Studio Code	32
Gambar 2. 12 Logo HuggingFace	32
Gambar 2. 13 Logo dari Streamlit	33
Gambar 2. 14 Logo Windows	34
Gambar 3. 1 Aplikasi BCA Mobile pada platfrom Google Play Store	36
Gambar 3. 2 Flowchart alur utama penelitian analisa sentimen	37
Gambar 3. 3 Highlight link id source yang akan diambil untuk scraping data.....	38
Gambar 3. 4 Contoh ulasan dari aplikasi BCA Mobile	39
Gambar 3. 5 Flowchart Preprocessing Data.....	40
Gambar 3. 6 Flowchart Split data	40
Gambar 3. 7 Flowchart training data	41
Gambar 3. 8 Flowchart Evaluate Model	42
Gambar 3. 9 Flowchart Inference Test.....	43
Gambar 3. 10 Design dari deployment website	43
Gambar 3. 11 Flowchart deployment website	44
Gambar 3. 12 Proses Pembentukan Script	50
Gambar 4. 1 Fitur-fitur yang tersedia pada BCA Mobile	55
Gambar 4. 2 Komentar negatif di Google Play (1)	56
Gambar 4. 3 Komentar negatif di Google Play (2)	56
Gambar 4. 4 Komentar positif dengan rating rendah (1)	56
Gambar 4. 5 Komentar positif dengan rating rendah (2)	57
Gambar 4. 6 Kode scraping data ulasan BCA Mobile.....	58
Gambar 4. 7 Tampilan data yang disimpan ke dalam Pandas.....	59
Gambar 4. 8 Hasil mensortir data berdasarkan tanggal	59
Gambar 4. 9 Dataset tersimpan dalam bentuk .csv di dalam VS code.....	60
Gambar 4. 10 Tipe Data ulasan bca	60
Gambar 4. 11 Kode membuat kategorikal menggunakan barchart.....	61
Gambar 4. 12 Barchart Distribusi score pada ulasan BCA Mobile	61
Gambar 4. 13 Impor library untuk eksplorasi topik.....	62

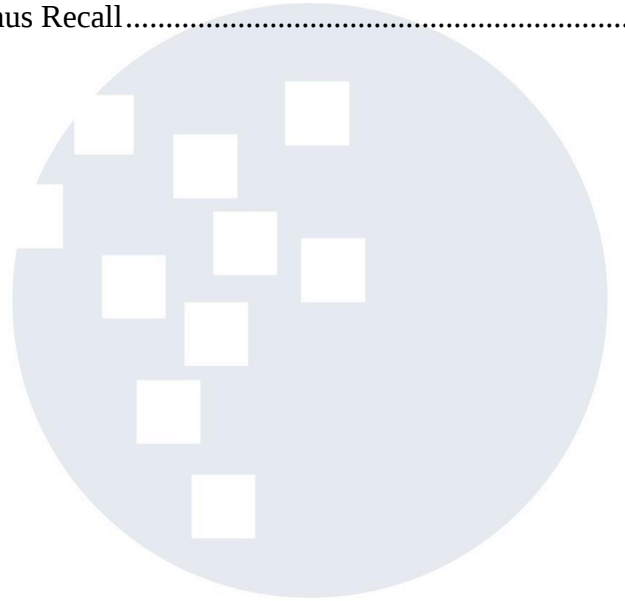
Gambar 4. 14 Kode preprocessing menghapus stop words dan pembuatan dictionary.....	62
Gambar 4. 15 Kode pencarian topik menggunakan LDA.....	63
Gambar 4. 16 Hasil dari topik pemodelan LDA	64
Gambar 4. 17 Kode untuk membuat visualisasi dari topik LDA.....	64
Gambar 4. 18 Kode untuk membuat visualisasi dari topik LDA lanjutan (2)	65
Gambar 4. 19 Barchart Distribusi persebaran topik ulasan teks BCA Mobile	66
Gambar 4. 20 Kode untuk membuat visualisasi distribusi sentimen per topic BCA Mobile	67
Gambar 4. 21 Kode untuk membuat visualisasi sentimen per topic BCA Mobile lanjutan (2)	67
Gambar 4. 22 Barchart distribusi sentimen per topic BCA Mobile.....	68
Gambar 4. 23 Barchart kata kata yang sering muncul berdasarkan salah satu topik "Update & Force Close issues"	69
Gambar 4. 24 Kode untuk membuat word cloud	72
Gambar 4. 25 WordCloud positif dari BCA Mobile.....	72
Gambar 4. 26 Wordcloud negatif dari BCA Mobile.....	73
Gambar 4. 27 Wordcloud netral dari BCA Mobile.....	73
Gambar 4. 28 Kode untuk menjalankan fungsi membuat bidirectional chart.....	74
Gambar 4. 29 Kode untuk menjalankan fungsi membuat bidirectional chart lanjutan (2)	74
Gambar 4. 30 Barchart kata kata yang sering muncul berdasarkan ulasan BCA Mobile di Google PlayStore	75
Gambar 4. 31 Pengecekan missing value pada dataset	76
Gambar 4. 32 Kode untuk menghapus missing data.....	77
Gambar 4. 33 Kode untuk melihat jumlah rows dan shape	77
Gambar 4. 34 Hasil klasifikasi sentiment berdasarkan Rules based labelling.....	78
Gambar 4. 35 Kode untuk melakukan pengecekan balance data.....	79
Gambar 4. 36 Kode untuk balancing data.....	79
Gambar 4. 37 Kode untuk balancing data lanjutan (2)	80
Gambar 4. 38 Check dataset setelah proses balancing data selesai	80
Gambar 4. 39 Kode untuk train and test split	81
Gambar 4. 40 Kode Fine-Tuning LSTM	84
Gambar 4. 41 Kode fine-tuning Distilbert	85
Gambar 4. 42 Kode Hyperparameter tuning Distilbert dengan Optuna.....	86
Gambar 4. 43 Tabel hasil test model Classification report dengan LSTM.....	88
Gambar 4. 44 Tabel hasil test model Classification report dengan DistilBERT ..	89
Gambar 4. 45 Hyperparameter Importance menggunakan Optuna	90
Gambar 4. 46 Tabel hasil test model Classification report dengan DistilBERT Optuna	91
Gambar 4. 47 Confusion matrix LSTM	94
Gambar 4. 48 Visualisasi Training dan Validation Loss LSTM.....	94
Gambar 4. 49 Confusion matrix Distilbert.....	95

Gambar 4. 50 Visualisasi Training dan Loss Validation DistilBert.....	96
Gambar 4. 51 Confusion Matrix DistilBert Optuna.....	97
Gambar 4. 52 Visualisasi Training dan Loss Validation DistilBert Optuna.....	97
Gambar 4. 53 Kode untuk save model Distilbert.....	98
Gambar 4. 54 Halaman dari website sederhana sentiment analyzer.....	99
Gambar 4. 55 Halaman ketika tombol "Browse File" ditekan.....	100
Gambar 4. 56 Kode API untuk memuat model ke dalam fungsi script Python..	100
Gambar 4. 57 Kode API untuk memanggil fungsi classifier dengan cara input text	101
Gambar 4. 58 Kode API dari fungsi Upload file untuk melakukan klasifikasi .	101
Gambar 4. 59 Kode API untuk menyimpan hasil klasifikasi dalam bentuk file csv	102
Gambar 4. 60 Hasil klasifikasi sentimen positif (Bahasa Indonesia).....	105
Gambar 4. 61 Hasil klasifikasi sentimen negatif (Bahasa Indonesia).....	105
Gambar 4. 62 Hasil klasifikasi sentimen positif (Bahasa English).....	106
Gambar 4. 63 Hasil klasifikasi sentimen negatif (Bahasa English).....	106
Gambar 4. 64 Proses melakukan opsi 2 dengan Bulk File Upload.....	107
Gambar 4. 65 Hasil klasifikasi sentimen dengan bulk upload file 1.....	108
Gambar 4. 66 Hasil klasifikasi sentimen dengan bulk upload file 1(netral).....	108



DAFTAR RUMUS

Rumus 2. 1 Rumus dari Forget Gate LSTM	27
Rumus 2. 2 Rumus dari Input Gate LSTM	27
Rumus 2. 3 Rumus dari Output Gate LSTM.....	27
Rumus 2. 4 Rumus Accuracy	28
Rumus 2. 5 Rumus Precision	29
Rumus 2. 6 Rumus f1-score	29
Rumus 2. 7 Rumus Recall.....	29



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran A Turnitin Similarity Report	121
Lampiran B Form Konsultasi Bimbingan.....	131

