

**PENERAPAN ALGORITMA CNN UNTUK TRANSLASI WICARA
MENJADI TEKS**



Thenardhi Syechlo

00000036583

UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA
PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK DAN INFORMATIKA
UNIVERSITAS MULTIMEDIA NUSANTARA
TANGERANG
2023

**PENERAPAN ALGORITMA CNN UNTUK TRANSLASI WICARA
MENJADI TEKS**



Diajukan sebagai salah satu syarat untuk memperoleh
Gelar Sarjana Komputer (S.Kom.)



HALAMAN PERNYATAAN TIDAK PLAGIAT

Dengan ini saya,

Nama : Thenardhi Syechlo

Nomor Induk Mahasiswa : 00000036583

Program Studi : Informatika

Skripsi dengan judul:

Penerapan Algoritma CNN untuk Translasi Wicara Menjadi Teks

merupakan hasil karya saya sendiri bukan plagiat dari karya 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 Skripsi maupun dalam penulisan laporan Skripsi, saya bersedia menerima konsekuensi dinyatakan TIDAK LULUS untuk Tugas akhir yang telah saya tempuh.

Tangerang, 6 Oktober 2023



(Thenardhi Syechlo)

UNIVERSITAS
MULTIMEDIA
NUSANTARA

HALAMAN PENGESAHAN

Skripsi dengan judul

PENERAPAN ALGORITMA CNN UNTUK TRANSLASI WICARA MENJADI TEKS

oleh

Nama : Thenardhi Syechlo
NIM : 00000036583
Program Studi : Informatika
Fakultas : Fakultas Teknik dan Informatika

Telah diujikan pada hari Selasa, 17 Oktober 2023

Pukul 13.00 WIB s/d 15.00 WIB dan dinyatakan

LULUS

Dengan susunan penguji sebagai berikut

Ketua Sidang

(Marlinda Vasty Overbeek, S.Kom.,
M.Kom.)
NIDN: 0818038501

Penguji

(Yaman Khaeruzzaman, M.Sc.)
NIDN: 0413057104

Pembimbing

(Moeljono Widjaja, B.Sc., M.Sc., Ph.D)
NIDN: 0311106903

Ketua Program Studi Informatika,

(Marlinda Vasty Overbeek, S.Kom., M.Kom.)
NIDN: 0818038501

HALAMAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademik Universitas Multimedia Nusantara, saya yang bertanda tangan di bawah ini:

Nama	:	Thenardhi Syechlo
NIM	:	00000036583
Program Studi	:	Informatika
Fakultas	:	Teknik dan Informatika
Jenis Karya	:	Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada **Universitas Multimedia Nusantara** hak Bebas Royalti Non-eksklusif (*Non-exclusive Royalty-Free Right*) atas karya ilmiah saya yang berjudul:

PENERAPAN ALGORITMA CNN UNTUK TRANSLASI WICARA MENJADI TEKS

Beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non eksklusif ini Universitas Multimedia Nusantara berhak menyimpan, mengalih media / format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis / pencipta dan sebagai pemilik Hak Cipta. Demikian pernyataan ini saya buat dengan sebenarnya.

Tangerang, 6 Oktober 2023

Yang menyatakan



Thenardhi Syechlo

UNIVERSITAS
MULTIMEDIA
NUSANTARA

Halaman Persembahan / Motto

”A good name is to be more desired than great wealth, Favor is better than silver and gold.”

Proverbs 22:1 (NASB)



KATA PENGANTAR

Puji Syukur atas berkat dan rahmat kepada Tuhan Yang Maha Esa, atas selesainya penulisan laporan Skripsi ini dengan judul: Penerapan Algoritma CNN untuk Translasi Wicara Menjadi Teks dilakukan untuk memenuhi salah satu syarat untuk mencapai gelar Sarjana Komputer Jurusan Informatika Pada Fakultas Teknik dan Informatika Universitas Multimedia Nusantara. Saya menyadari bahwa, tanpa bantuan dan bimbingan dari berbagai pihak, dari masa perkuliahan sampai pada penyusunan skripsi ini, sangatlah sulit bagi saya untuk menyelesaikan skripsi ini. Oleh karena itu, saya mengucapkan terima kasih kepada:

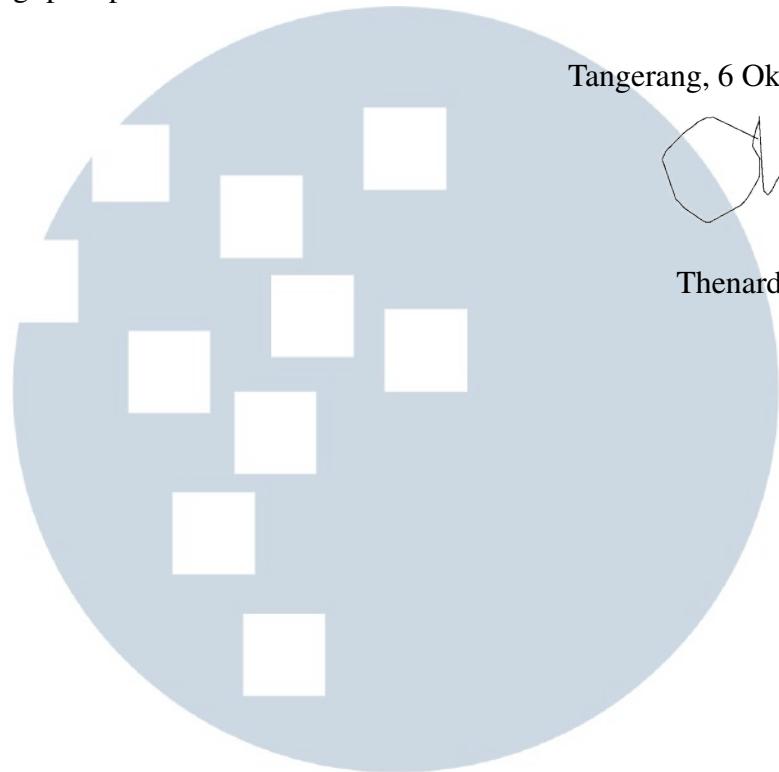
1. Orang Tua dan keluarga saya yang telah memberikan bantuan dukungan material dan moral, sehingga penulis dapat menyelesaikan skripsi ini.
2. Bapak Dr. Ninok Leksono, selaku Rektor Universitas Multimedia Nusantara.
3. Dr. Eng. Niki Prastomo, S.T., M.Sc., selaku Dekan Fakultas Teknik dan Informatika Universitas Multimedia Nusantara.
4. Ibu Marlinda Vasty Overbeek, S.Kom., M.Kom., selaku Ketua Program Studi Informatika Universitas Multimedia Nusantara.
5. Bapak Moeljono Widjaja, B.Sc., M.Sc., Ph.D, sebagai Pembimbing pertama yang telah banyak meluangkan waktu untuk memberikan bimbingan, arahan dan motivasi atas terselesainya skripsi ini.
6. Ibu Alethea Suryadibrata, S.Kom., M.Eng., selaku dosen Teknik Informatika yang telah meminjamkan komputer di ruangan laboratorium AI UMN
7. Ibu Farica Perdana Putri, S.Kom., M.Sc., dan Bapak Dennis Gunawan, S.Kom., M.Sc. selaku dosen Teknik Informatika.
8. Ibu Monika Evelin Johan, S.Kom.,M.M.S.I., selaku dosen Sistem Informasi yang telah meminjamkan komputer di ruangan laboratorium Big Data UMN
9. Bapak Dian Kurnadi, selaku warga perumahan Sutera Delima I no. 16, Alam Sutera
10. Kakak Aldo Aditya, selaku master AI di Nanyang Technological University, Singapura

Semoga skripsi ini bermanfaat, baik sebagai sumber informasi maupun sumber inspirasi, bagi para pembaca.

Tangerang, 6 Oktober 2023



Thenardhi Syechlo



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

PENERAPAN ALGORITMA CNN UNTUK TRANSLASI WICARA MENJADI TEKS

Thenardhi Syechlo

ABSTRAK

Wicara adalah keterampilan umum yang diterapkan sejak usia dini. Wicara adalah aktivitas keseharian manusia yang digunakan dalam berbicara, berkomunikasi, dan melakukan interaksi. 90% kegiatan tersebut mesti dilakukan dengan wicara. Pada kesempatan kali ini, kemampuan berbicara manusia akan diuji dengan metode pembelajaran mendalam. Metode ini merupakan metode pembelajaran mesin berdasarkan jaringan saraf tiruan. Metode ini bertujuan agar aplikasi komputer dapat mengetahui kata demi kata yang diucapkan oleh manusia. Aplikasi yang akan digunakan adalah PyCharm dengan menggunakan bahasa pemrograman Python. Selain itu, aplikasi Tensorflow juga digunakan dalam menguji konteks pembicaraan yang dilakukan oleh manusia. Algoritma yang akan digunakan adalah *Convolutional Neural Network* (CNN). Algoritma ini akan mengukur ketepatan himpunan data yang diuji sebanyak beberapa periode. Selain itu, himpunan data dari Kaggle juga digunakan dalam proses pelaksanaan penelitian kali ini. Himpunan data tersebut digunakan untuk mengukur tingkat akurasi data suara yang diucapkan melalui mulut manusia. Kesimpulan dari penelitian ini adalah kemampuan berbicara seseorang dapat ditingkatkan dengan metode pembelajaran mendalam melalui pembuatan kode program menggunakan bahasa pemrograman Python pada aplikasi PyCharm dengan pengujian model menggunakan pengujian model yang berupa pengujian spektogram dan pengujian *Mel Frequency Cepstral Coefficients* (MFCC) pada himpunan data yang digunakan.

Kata kunci: *CNN, MFCC, Pembelajaran Mendalam, Python, Wicara*

UNIVERSITAS
MULTIMEDIA
NUSANTARA

Application of the CNN Algorithm for Translation of Speech into Text

Thenardhi Syechlo

ABSTRACT

Speech is a common skill that is applied from an early age. Speech is a daily human activity that is used in speaking, communicating, and interacting. 90% of these activities must be done with speech. On this occasion, the ability to speak humans will be tested using deep learning methods. This method is a machine learning method based on artificial neural networks. This method aims to allow computer applications to find out word for word spoken by humans. The application to be used is PyCharm using the Python programming language. In addition, the Tensorflow application is also used to test the context of human speech. The algorithm to be used is a Convolutional Neural Network (CNN). This algorithm will measure the accuracy of the data set tested for several periods. Apart from that, the data set from Kaggle was also used in the process of conducting this research. The data set is used to measure the level of accuracy of voice data uttered by the human mouth. The conclusion from this study is that one's speaking ability can be improved with deep learning methods through creating program code using the Python programming language in the PyCharm application by testing models using model testing in the form of spectrogram testing and Mel Frequency Cepstral Coefficients (MFCC) testing on the data sets used.

Keywords: *CNN, Deep Learning, MFCC, Python, Speech*



DAFTAR ISI

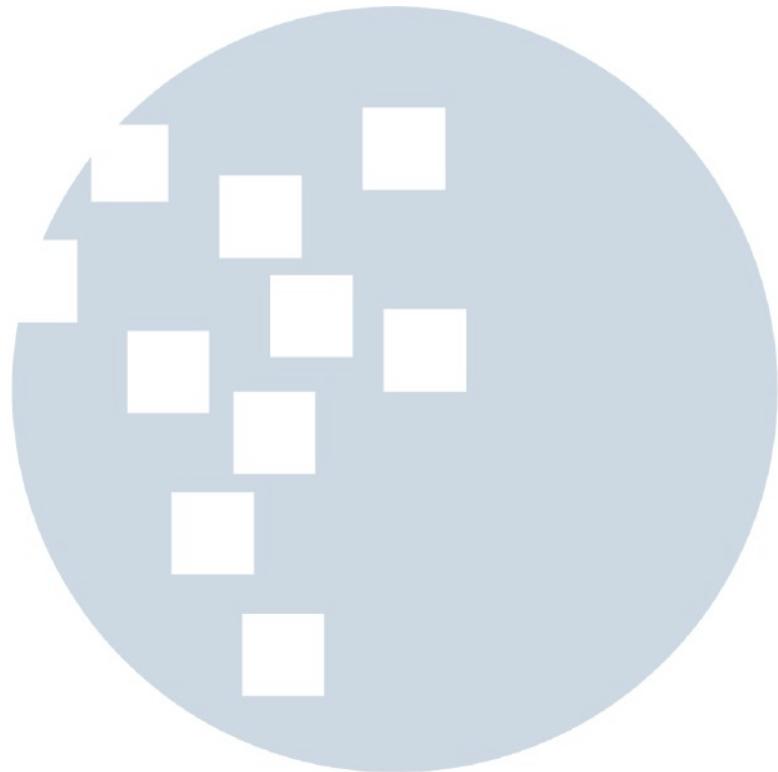
HALAMAN JUDUL	i
PERNYATAAN TIDAK MELAKUKAN PLAGIAT	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN PUBLIKASI ILMIAH	iv
HALAMAN PERSEMPERBAHAN/MOTO	v
KATA PENGANTAR	vi
ABSTRAK	viii
ABSTRACT	ix
DAFTAR ISI	x
DAFTAR GAMBAR	xii
DAFTAR TABEL	xiv
DAFTAR KODE	xvi
DAFTAR LAMPIRAN	xvii
BAB 1 PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Rumusan Masalah	3
1.3 Batasan Permasalahan	3
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	4
1.6 Sistematika Penulisan	4
1.7 Spesifikasi Sistem	5
BAB 2 LANDASAN TEORI	7
2.1 Berbicara	7
2.2 Wicara	7
2.3 Convolutional Neural Network (CNN)	8
2.3.1 Spektogram	12
2.3.2 Mel Frequency Cepstral Coefficients (MFCC)	13
2.3.3 Zero crossing data	16
2.3.4 Pembuatan Model	16
2.4 Tensorflow	16
2.5 Audio	18
2.6 One Hot Encoder	19
2.7 Model h5	21
BAB 3 METODOLOGI PENELITIAN	22
3.1 Studi Pustaka	22
3.2 Tempat dan Waktu Penelitian	22
3.3 Metode Pengumpulan Data	22
3.4 Metode Pengambilan Sampel	23
3.5 Metode Perancangan	24
3.6 Metode Evaluasi	26
3.7 Metode Dokumentasi	26
3.8 Metode Menampilkan Hasil Wicara	26
3.9 Metode Penggunaan Aplikasi	27
3.9.1 Visual Studio Code	27
3.9.2 PyCharm	27
3.9.3 Windows Media Player	29
3.9.4 Tensorflow	29
3.10 Himpunan data yang digunakan	30

3.11	Proses implementasi data Kaggle	30
3.11.1	Melihat semua berkas audio	31
3.11.2	Impor pustaka	32
3.11.3	Proses menggunakan himpunan data pribadi	32
3.11.4	Menghitung jumlah sampel	32
3.11.5	Proses Distribusi Data	33
3.11.6	Proses Pengujian Kdeplot	33
3.11.7	Proses Perhitungan Sampel	34
3.11.8	Pengujian fitur ekstraksi audio	35
3.11.9	Pengujian semua himpunan data	40
3.11.10	Pemisahan data	41
3.12	Proses Pembuatan Model	42
3.12.1	Pelatihan Data	42
3.12.2	Pengujian Data	45
3.12.3	Pemeriksaan Model	46
3.13	Menyimpan Model	47
3.14	Metode Melihat Model	48
3.14.1	Himpunan data	48
3.14.2	Jenis himpunan data	49
3.14.3	Ukuran dari himpunan data	49
3.14.4	Isi dari himpunan data	50
3.14.5	Kunci pada himpunan data	50
3.14.6	Membandingkan Akurasi Data	51
BAB 4	HASIL DAN DISKUSI	53
4.1	Pengujian Suara	53
4.2	Hasil Implementasi Data Kaggle	53
4.3	Hasil Penyimpanan audio	54
4.4	Hasil Pengujian Model	54
4.4.1	Pengenalan Model	54
4.4.2	Hasil Distribusi Data	55
4.4.3	Hasil Pengujian Kdeplot	56
4.4.4	Fitur ekstraksi audio	58
4.4.5	Hasil pengujian semua himpunan data	65
4.4.6	Hasil Pemisahan Data	66
4.4.7	Hasil Pembuatan Model	66
4.4.8	Hasil dari Pemeriksaan Model	72
4.5	Hasil dari Model	73
4.5.1	Himpunan data	74
4.5.2	Jenis himpunan data	74
4.5.3	Ukuran himpunan data	75
4.5.4	Isi dari himpunan data	75
4.5.5	Kunci pada himpunan data	75
4.5.6	Hasil pengujian dan pelatihan data	76
4.5.7	Perbandingan Hasil Akurasi Data	78
BAB 5	SIMPULAN DAN SARAN	79
5.1	Kesimpulan	79
5.2	Saran	79
DAFTAR PUSTAKA	81

DAFTAR GAMBAR

Gambar 2.1	Struktur dari lapisan <i>Convolutional Neural Network</i> (CNN). Sumber: [1]	11
Gambar 2.2	Menguji data dengan MFCC. Sumber: [2]	15
Gambar 3.1	Prosedur translasi hasil wicara ke teks	23
Gambar 3.2	Prosedur konversi berkas audio ke teks	24
Gambar 3.3	Prosedur konversi berkas audio ke WAV	25
Gambar 3.4	Konversi berkas audio M4A ke WAV	25
Gambar 3.5	Gambar untuk melihat hasil wicara	26
Gambar 3.6	Prosedur Implementasi Data Kaggle	31
Gambar 4.1	Hasil distribusi data	56
Gambar 4.2	Grafik Pelatihan Kdeplot	56
Gambar 4.3	Hasil pengujian fitur ekstraksi audio array ke nol	58
Gambar 4.4	Hasil pengujian MFCC array ke nol	59
Gambar 4.5	Hasil implementasi MFCC array ke nol	59
Gambar 4.6	Hasil pengujian Mel-spektogram array ke nol	60
Gambar 4.7	Hasil implementasi Mel-spektogram array ke nol	60
Gambar 4.8	Hasil pengujian zero crossing data array ke nol	61
Gambar 4.9	Hasil pengujian fitur ekstraksi audio array ke-1000	62
Gambar 4.10	Hasil pengujian MFCC array ke-1000	62
Gambar 4.11	Hasil implementasi MFCC array ke nol	63
Gambar 4.12	Hasil pengujian Mel-spektogram array ke-1000	63
Gambar 4.13	Hasil implementasi Mel-spektogram array ke nol	64
Gambar 4.14	Hasil pengujian zero crossing data mandiri array ke-1000	64
Gambar 4.15	Jumlah data untuk pelatihan dan pengujian	66
Gambar 4.16	Ringkasan pada model	67
Gambar 4.17	Akurasi data sebelum model pelatihan	67
Gambar 4.18	Hasil dari Proses Pelatihan Data	68
Gambar 4.19	Akurasi data setelah pelatihan	69
Gambar 4.20	Hasil perhitungan akurasi data dalam grafik	70
Gambar 4.21	Hasil pengujian data dengan X _{test}	70
Gambar 4.22	Hasil dari confusion matriks	72
Gambar 4.23	Hasil model dari himpunan data yang tersedia	74
Gambar 4.24	Jenis data yang tersedia	74
Gambar 4.25	Isi dari himpunan data kedua	75
Gambar 4.26	Kunci pada optimizer_weights	76
Gambar 4.27	Kunci pada model_weights	76
Gambar 4.28	Hasil dari X_train	76
Gambar 4.29	Hasil dari y_train	77
Gambar 4.30	Hasil dari X_test	77
Gambar 4.31	Hasil dari y_test	78
Gambar 4.32	Perbandingan akurasi data	78
Gambar 5.1	Lapisan konvolusi. Sumber: [1]	86
Gambar 5.2	Lapisan pooling. Sumber: [1]	86
Gambar 5.3	Lapisan Aktif. Sumber: [1]	86
Gambar 5.4	Skema <i>Discrete Fourier Transform</i> (DFT) spektogram. Sumber: [3]	87
Gambar 5.5	Proses ekstraksi MFCC. Sumber: [4]	87
Gambar 5.6	Penekanan awal pada ekstraksi MFCC. Sumber: [4]	87

Gambar 5.7	Proses akhir pada ekstraksi MFCC. Sumber: [4]	87
Gambar 5.8	Melihat model yang tersedia	88



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR TABEL

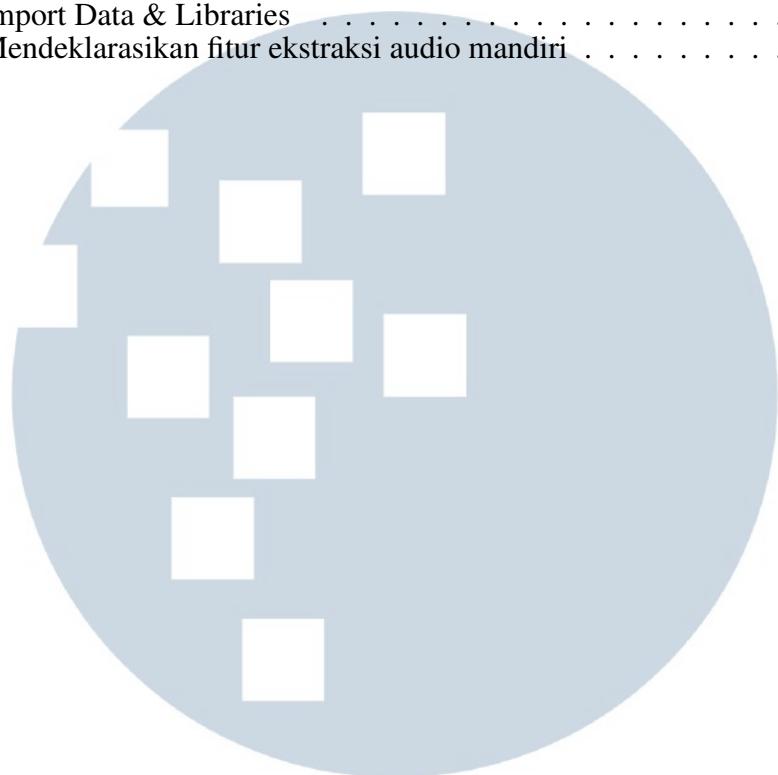
Tabel 4.1	Tabel nomor sampel setiap kosakata	54
Tabel 4.2	Tabel jumlah sampel pada kosakata	55
Tabel 4.3	Tabel visualisasi label dan sampel	57
Tabel 4.4	Tabel visualisasi sampel dan nilai	65
Tabel 4.5	Tabel visualisasi label array	66
Tabel 4.6	Hasil akurasi data satu per satu berdasarkan jumlah iterasi .	69
Tabel 4.7	Tabel model prediksi pada data	71
Tabel 4.8	Tabel y sebenarnya dan model prediksi pada data	71
Tabel 4.9	Hasil laporan klasifikasi	73
Tabel 4.10	Ukuran himpunan data satu dan himpunan data dua	75



DAFTAR KODE

3.1	Aplikasi melatih kemampuan berbicara	27
3.2	Mengambil dataset (himpunan data) untuk pengujian model pada Visual Studio Code	32
3.3	Menghitung jumlah sampel data	32
3.4	Distribusi data menjadi beberapa bagian	33
3.5	Menggambarkan sampel dengan grafik	33
3.6	Proses menghitung jumlah sampel mandiri	34
3.7	Menggambarkan ukuran sampel mandiri	34
3.8	Pengujian gelombang audio mandiri	35
3.9	Pengujian MFCC mandiri	35
3.10	Implementasi pengujian MFCC mandiri	36
3.11	Pengujian mel-spectrogram mandiri	36
3.12	Implementasi pengujian mel-spectrogram mandiri	36
3.13	Pengujian zero-crossing data mandiri	37
3.14	Pengujian gelombang audio array 1000	38
3.15	Pengujian MFCC array 1000	38
3.16	Implementasi pengujian MFCC array 1000	38
3.17	Pengujian mel-spektogram array 1000	39
3.18	Implementasi pengujian mel-spektogram array 1000	39
3.19	Pengujian zero-crossing data array 1000	40
3.20	Pengujian value data	40
3.21	Pengujian label data	40
3.22	Pengenalan data dengan menggunakan Label Encoder dan One Hot Encoder	41
3.23	Pemisahan data menggunakan train test split	41
3.24	Proses pembentukan model	42
3.25	Menunjukkan model yang ada	43
3.26	Menunjukkan grafik model	43
3.27	Menampilkan rata-rata akurasi data	43
3.28	Proses pelatihan data	44
3.29	Menampilkan akurasi data satu-persatu	44
3.30	Menampilkan akurasi data dalam grafik	44
3.31	Pengujian data dengan menggunakan X test	45
3.32	Proses pelatihan menggunakan model prediksi	45
3.33	Inisiasi fungsi getName	46
3.34	Pengujian data berdasarkan nilai y sebenarnya dan model prediksi	46
3.35	Proses pemeriksaan menggunakan matriks confusion	47
3.36	Proses pemeriksaan berdasarkan laporan klasifikasi	47
3.37	Menyimpan hasil pengujian model	48
3.38	Memuat hasil pengujian model	48
3.39	Melihat himpunan data yang ada pada model pelatihan	48
3.40	Melihat jenis himpunan data pada model pelatihan	49
3.41	Menampilkan ukuran dari dataset1	49
3.42	Menampilkan ukuran dari dataset2	50
3.43	Menampilkan kunci utama pada optimizer weights	50
3.44	Menampilkan kunci utama pada model weights	51
3.45	Menulis dan menyimpan akurasi data	52
3.46	Membaca akurasi data	52

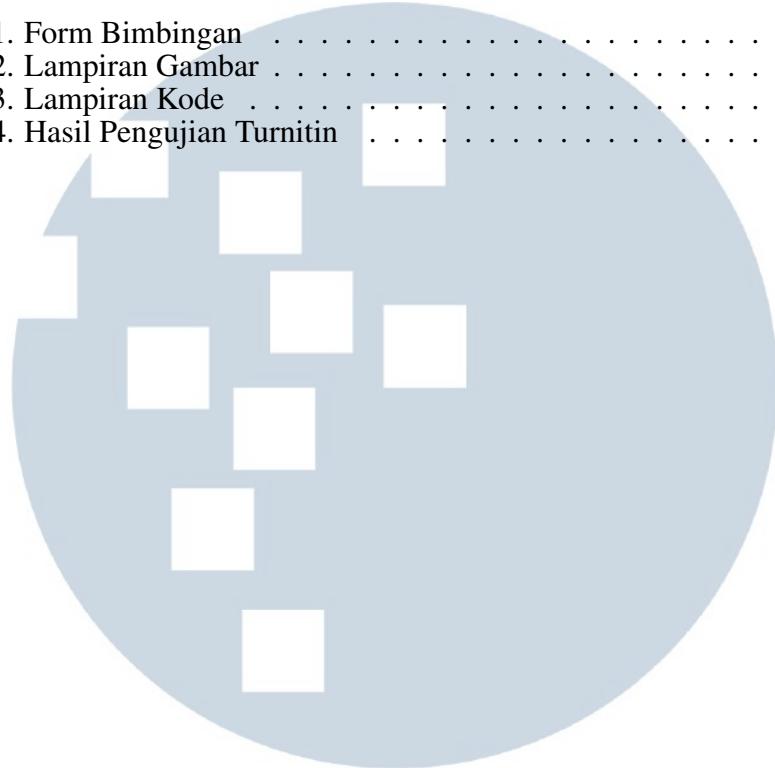
5.1	Mengambil dataset (himpunan data) untuk pengujian model pada Kaggle	89
5.2	Import Data & Libraries	89
5.3	Mendeklarasikan fitur ekstraksi audio mandiri	90



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA

DAFTAR LAMPIRAN

Lampiran 1. Form Bimbingan	83
Lampiran 2. Lampiran Gambar	86
Lampiran 3. Lampiran Kode	89
Lampiran 4. Hasil Pengujian Turnitin	91



UMN
UNIVERSITAS
MULTIMEDIA
NUSANTARA