

**KLASIFIKASI MOOD MUSIK BERBASIS SUARA
MENGUNAKAN ALGORITMA CONVOLUTIONAL
NEURAL NETWORK (CNN)**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

SIWAN YESARELA HAREFA

22.11.5179

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**KLASIFIKASI MOOD MUSIK BERBASIS SUARA
MENGUNAKAN ALGORITMA CONVOLUTIONAL
NEURAL NETWORK (CNN)**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

SIWAN YESARELA HAREFA

22.11.5179

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**KLASIFIKASI MOOD MUSIK BERBASIS SUARA MENGGUNAKAN
ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN)**

yang disusun dan diajukan oleh

Siwan Yesarela Harefa

22.11.5179

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 22 Mei 2026

Dosen Pembimbing,



Theopilus Bayu Sasongko, S.Kom, M.Eng

NIK. 190302375

HALAMAN PENGESAHAN
SKRIPSI
KLASIFIKASI MOOD MUSIK BERBASIS SUARA MENGGUNAKAN
ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN)

yang disusun dan diajukan oleh

Siwan Yesarela Harefa

22.11.5179

Telah dipertahankan di depan Dewan Penguji
pada tanggal 22 Mei 2026

Susunan Dewan Penguji

Nama Penguji

Nuri Cahyono, M.Kom.
NIK. 190302278

Nafiatun Sholihah, S.Kom., M.Cs.
NIK. 1190302524

Theopilus Bayu Sasongko, S.Kom, M.Eng
NIK. 190302375

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 22 Mei 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Siwan Yesarela Harefa
NIM : 22.11.5179

Menyatakan bahwa Skripsi dengan judul berikut:

KLASIFIKASI MOOD MUSIK BERBASIS SUARA MENGGUNAKAN ALGORITMA CONVOLUTIONAL NEURAL NETWORK (CNN)

Dosen Pembimbing : Theopilus Bayu Sasongko, S.Kom, M.Eng

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 22 Mei 2026

Yang Menyatakan,

Siwan Yesarela Harefa

HALAMAN PERSEMBAHAN

Puji syukur saya panjatkan kepada Tuhan Yang Maha Esa atas segala rahmat, kasih, dan penyertaan-Nya sehingga saya dapat menyelesaikan skripsi ini dengan baik. Karya sederhana ini saya persembahkan dengan penuh rasa hormat dan terima kasih kepada:

Kedua orang tua tercinta, yang selalu memberikan kasih sayang, doa, dukungan, perhatian, serta pengorbanan yang tidak ternilai harganya.

Keluarga tercinta, yang senantiasa memberikan semangat, motivasi, dan dukungan selama proses perkuliahan hingga penyusunan skripsi ini.

Dosen pembimbing, yang dengan penuh kesabaran telah meluangkan waktu, tenaga, dan pikiran untuk memberikan bimbingan, arahan, kritik, serta saran yang membangun selama proses penelitian dan penyusunan skripsi ini.

Seluruh dosen dan staf Program Studi, yang telah memberikan ilmu pengetahuan, pengalaman, pelayanan, serta bantuan selama saya menjalani masa studi di perguruan tinggi.

Sahabat dan teman-teman seperjuangan, yang telah menemani perjalanan perkuliahan, memberikan dukungan, bantuan, motivasi, serta menjadi tempat berbagi suka dan duka selama proses penyelesaian skripsi ini.

Diri saya sendiri, yang telah berusaha, bertahan, dan tidak menyerah dalam menghadapi berbagai tantangan selama proses perkuliahan dan penyusunan skripsi ini. Terima kasih karena telah terus berjuang hingga mencapai tahap ini.

Akhir kata, skripsi ini saya persembahkan sebagai wujud rasa syukur dan terima kasih kepada semua pihak yang telah menjadi bagian dari perjalanan akademik saya. Semoga karya ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan menjadi langkah awal untuk meraih cita-cita di masa depan.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke Tuhan YME atas segala rahmat, karunia, serta hidayah-Nya sehingga skripsi ini dapat terselesaikan dengan baik. Skripsi ini disusun sebagai salah satu syarat kelulusan dalam menempuh program studi Sarjana (S-1) di Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Yogyakarta. Penulis menyadari bahwa keberhasilan dalam menyelesaikan skripsi ini tidak lepas dari dukungan, dan bantuan yang sangat berarti dari berbagai pihak. Oleh karena itu, dengan segala kerendahan hati, penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M. sebagai Rektor Universitas Amikom Yogyakarta.
2. Bapak Hanif Al Fatta, M.Kom. Ph.D. Sebagai Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
3. Ibu Eli Pujastuti, M.Kom. sebagai Kaprodi S1 Informatika Universitas Amikom Yogyakarta.
4. Bapak Theopilus Bayu Sasongko, S.Kom, M.Eng. sebagai Dosen Pembimbing.
5. Bapak dan Ibu Dosen di Universitas Amikom Yogyakarta yang telah memberikan bimbingan dan ilmu pengetahuan selama proses perkuliahan.
6. Orang tua serta seluruh anggota keluarga yang selalu memberikan doa, dukungan, semangat dan kasih sayang yang tiada henti kepada penulis dalam menyelesaikan studi hingga tahap akhir ini.
7. Kepada seluruh teman-teman dan pihak-pihak yang tidak bisa disebutkan satu persatu

Yogyakarta, 22 Mei 2026



Siwan Yesarela Harefa

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	x
DAFTAR GAMBAR.....	xi
DAFTAR LAMBANG DAN SINGKATAN	xii
DAFTAR ISTILAH	xiii
INTISARI	xiv
<i>ABSTRACT</i>	xv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	5
BAB II TINJAUAN PUSTAKA	7
2.1 Studi Literatur.....	7
2.2 Dasar Teori	14

2.2.1 Musik dan Mood Musik	14
2.2.2 Audio Digital	15
2.2.3 Mel-Spectrogram	15
2.2.4 Mel-Frequency Cepstral Coefficients (MFCC).....	16
2.2.5 Machine Learning dalam Klasifikasi Audio.....	16
2.2.6 Deep Learning	17
2.2.7 Class Weight	19
2.2.8 Convolutional Neural Network (CNN)	20
2.2.9 Evaluasi Model	24
BAB III METODE PENELITIAN	29
3.1 Objek Penelitian	29
3.2 Alur Penelitian.....	30
3.2.1 Pengumpulan Data.....	32
3.2.2 Pra-pemrosesan Data Audio	32
3.2.3 Ekstraksi Fitur Audio	34
3.2.4 Pembagian Data.....	35
3.2.5 Pelatihan Model.....	36
3.2.6 Augmentasi data	37
3.2.7 Perancangan Model CNN.....	37
3.2.8 Pengujian dan Evaluasi Model	38
3.2.9 Perbandingan Hasil 2 Pelatihan Model	39
3.2.10 Analisis Hasil.....	39
3.3 Alat dan Bahan	40
3.3.1 Alat.....	40
3.3.2 Bahan.....	41

BAB IV HASIL DAN PEMBAHASAN	42
4.1 Pengumpulan Data	42
4.2 Pra-Pemrosesan dan Ekstraksi Fitur Audio	43
4.2.1 Pra-Pemrosesan Sinyal Audio	43
4.2.2 Ekstraksi Fitur	44
4.3 Pembagian Dataset dan Augmentasi Data	48
4.3.1 Skema Pembagian Dataset	48
4.3.2 Augmentasi Data	48
4.4 Arsitektur Model CNN	50
4.5 Hasil Pelatihan Model.....	56
4.5.1 Konfigurasi Pelatihan.....	56
4.5.2 Kurva Training dan Validasi	57
4.6 Evaluasi Model pada Data Uji	59
4.6.1 Metrik Evaluasi Keseluruhan	59
4.6.2 Analisis Per Kelas (Classification Report).....	59
4.7 Analisis Confusion Matrix.....	60
4.8 Visualisasi Fitur Audio per Kelas Mood	62
4.9 Perbandingan Komprehensif Program 1 dan Program 2	65
4.10 Deploymentasi Program.....	68
BAB V PENUTUP	73
5.1 Kesimpulan.....	73
5.2 Saran.....	74
REFERENSI	76
LAMPIRAN.....	80

DAFTAR TABEL

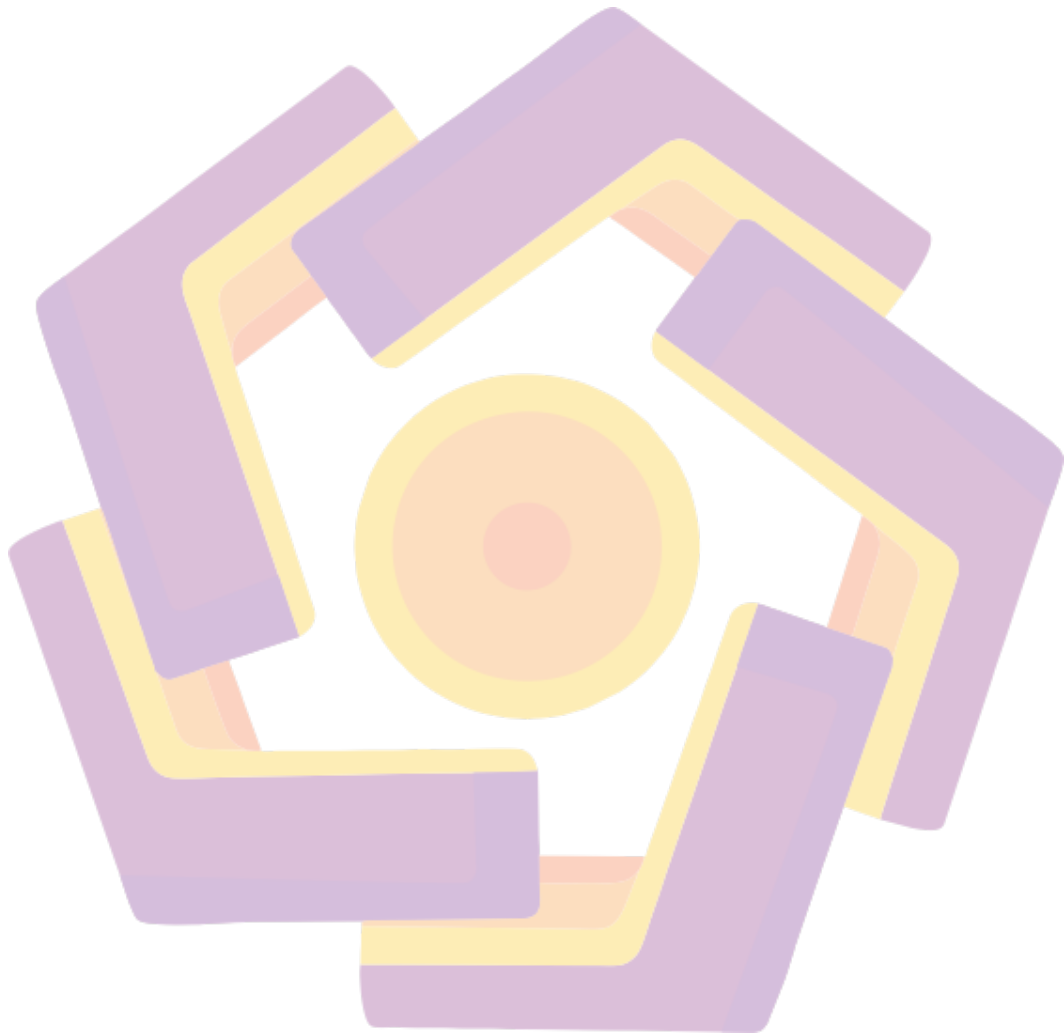
Tabel 2. 1 Keaslian Penelitian.....	10
Tabel 3. 1 Dataset Music Classification.....	41
Tabel 4. 1 Distribusi Kelas Dataset Audio Musik.....	43
Tabel 4. 2 Tabel Ringkasan Fitur	45
Tabel 4. 3 Skema Pembagian Dataset	48
Tabel 4. 4 Teknik Augmentasi yang Diterapkan.....	50
Tabel 4. 5 Arsitektur Detail Model CNN	54
Tabel 4. 6 Konfigurasi Hyperparameter Pelatihan.....	57
Tabel 4. 7 Perbandingan Hasil Terbaik pada Data Validasi.....	58
Tabel 4. 8 Perbandingan Hasil Evaluasi Model pada Data Testing	59
Tabel 4. 9 Classification Report Per Kelas - Best model	60
Tabel 4. 10 Confusion Matrix Program 2 pada Data Testing	60
Tabel 4. 11 Karakteristik Visual Fitur Audio per Kelas Mood	64
Tabel 4. 12 Perbandingan Komprehensif Program 1 dan Program 2.....	65

DAFTAR GAMBAR

Gambar 2. 1 Machine Learning dalam Klasifikasi Audio.....	17
Gambar 2. 2 Deep Learning	18
Gambar 2. 3 Convolutional Layer	21
Gambar 3. 1 Alur Penelitian.....	31
Gambar 4. 1 Visualisasi Distribusi Kelas.....	42
Gambar 4. 2 Kode Proses Ekstraksi Fitur.....	45
Gambar 4. 3 Hasil Output Ekstraksi Fitur.....	46
Gambar 4. 4 Kode Proses Augmentasi Data	49
Gambar 4. 5 Hasil Augmentasi data.....	50
Gambar 4. 6 Kode Proses Arsitektur CNN	52
Gambar 4. 7 Hasil Arsitektur CNN	55
Gambar 4. 8 Tampilan Kurva Training dan Validasi	57
Gambar 4. 9 Tampilan Confusion Matrix	61
Gambar 4. 10 Visualisasi Fitur Audio per Kelas Mood	63
Gambar 4. 11 Halaman Utama Tampilan Program.....	68
Gambar 4. 12 Tampilan Spesifikasi Model pada Program.....	69
Gambar 4. 13 Tampilan Endpoint API.....	69
Gambar 4. 14 Tampilan Label Mood	70
Gambar 4. 15 Tampilan contoh penggunaan API	70
Gambar 4. 16 Tampilan Demo Prediksi Program	71
Gambar 4. 17 Tampilan Demo program secara live.....	72

DAFTAR LAMBANG DAN SINGKATAN

MFCC	Mel Frequency Cepstral Coefficients
CNN	Convolutional Neural Network



DAFTAR ISTILAH



Deep Learning	Metode Pembelajaran mesin berbasis jaringan saraf
Hyperparameter	Parameter yang ditentukan sebelum proses pelatihan model.
Preprocessing Data	Tahap pengolahan awal data sebelum digunakan dalam model.
Normalisasi	Proses mengubah skala data ke rentang tertentu.
Training Data	Data yang digunakan untuk melatih model.
Testing Data	Data yang digunakan untuk menguji performa model.
Overfitting	Kondisi model terlalu menyesuaikan data latih sehingga kurang baik pada data baru.
Augmentasi	Teknik memperbanyak data audio secara artifisial dengan menerapkan berbagai transformasi .
Audio	Sinyal suara dalam bentuk digital yang merepresentasikan gelombang bunyi
Perbandingan	Proses evaluasi yang dilakukan untuk membandingkan performa.
Ekstraksi	Proses pengambilan fitur-fitur relevan dari data mentah.
Class Weight	Bobot yang diberikan kepada setiap kelas selama proses pelatihan untuk menangani ketidakseimbangan data.

INTISARI

Musik merupakan media ekspresi yang memiliki kemampuan luar biasa dalam menyampaikan emosi dan memengaruhi suasana hati pendengarnya. Namun, klasifikasi mood musik secara otomatis menghadapi tantangan yang signifikan karena sifat audio yang kompleks, multidimensional, dan subjektif. Sistem klasifikasi manual yang mengandalkan penilaian manusia dinilai tidak efisien dan tidak skalabel untuk mengelola koleksi musik dalam jumlah besar, seperti pada platform streaming modern. Ketidakmampuan sistem otomatis untuk mengenali nuansa emosional dalam musik berdampak pada pengalaman pengguna yang kurang personal, rekomendasi lagu yang tidak akurat, serta keterbatasan dalam pengembangan aplikasi berbasis konteks emosional seperti terapi musik, kurasi playlist adaptif, dan sistem rekomendasi cerdas.

Penelitian ini mengembangkan sistem klasifikasi mood musik berbasis sinyal suara menggunakan algoritma Convolutional Neural Network (CNN) untuk lima kelas mood, yaitu *aggressive*, *dramatic*, *romantic*, *sad*, dan *happy*. Dataset yang digunakan adalah "Music Classification" dari Kaggle, terdiri dari 10.133 sampel audio berformat WAV dengan durasi 10 detik dan *sampling rate* 22.050 Hz. Fitur audio diekstraksi melalui empat *channel*, yakni MFCC, Delta MFCC, Delta-Delta MFCC, dan Mel-Spectrogram, yang kemudian digabungkan menjadi tensor berukuran $(40 \times 130 \times 4)$ sebagai input model. Arsitektur CNN yang dirancang terdiri dari empat blok konvolusi, *Batch Normalization*, *GlobalAveragePooling2D*, serta dua lapisan *fully connected* yang dilengkapi regularisasi *dropout* dan L2. Guna menangani ketidakseimbangan kelas dalam dataset, diterapkan *balanced class weight* selama proses pelatihan model.

Model terbaik yang dihasilkan mencapai akurasi sebesar 79,93% dan *precision* 80,88% pada data uji, menunjukkan kemampuan generalisasi yang baik terhadap data yang belum pernah dilihat sebelumnya. Penerapan *balanced class weight* terbukti berkontribusi signifikan dalam meningkatkan performa model secara keseluruhan. Sistem kemudian di-deploy menggunakan *framework* Flask sebagai REST API dengan antarmuka web interaktif, sehingga memungkinkan prediksi mood musik secara langsung dan mudah diakses. Hasil penelitian ini dapat dimanfaatkan oleh pengembang aplikasi musik, platform *streaming*, peneliti di bidang *Music Information Retrieval* (MIR), serta praktisi terapi musik berbasis teknologi. Untuk penelitian lebih lanjut, disarankan eksplorasi arsitektur yang lebih dalam seperti ResNet atau penggunaan model *transformer* berbasis audio seperti AST (*Audio Spectrogram Transformer*), perluasan kelas mood, serta pemanfaatan teknik augmentasi data audio guna meningkatkan akurasi dan ketahanan model pada kondisi audio dunia nyata yang lebih beragam.

Kata kunci: Klasifikasi Mood Musik, MFCC, Mel-Spectrogram, *Convolutional Neural Network*, *Class Weight*

ABSTRACT

Music is an expressive medium with a remarkable ability to convey emotions and influence the mood of its listeners. However, automatic music mood classification faces significant challenges due to the complex, multidimensional, and subjective nature of audio signals. Manual classification systems that rely on human judgment are considered inefficient and unscalable for managing large music collections, such as those found on modern streaming platforms. The inability of automated systems to recognize emotional nuances in music negatively impacts user experience, resulting in inaccurate song recommendations and limitations in the development of emotion-aware applications such as music therapy, adaptive playlist curation, and intelligent recommendation systems.

This study develops a music mood classification system based on audio signals using a Convolutional Neural Network (CNN) algorithm for five mood classes: *aggressive*, *dramatic*, *romantic*, *sad*, and *happy*. The dataset used is the "Music Classification" dataset from Kaggle, consisting of 10,133 WAV audio samples with a duration of 10 seconds and a sampling rate of 22,050 Hz. Audio features were extracted through four channels — MFCC, Delta MFCC, Delta-Delta MFCC, and Mel-Spectrogram — which were then combined into a tensor of size $(40 \times 130 \times 4)$ as model input. The CNN architecture was designed with four convolutional blocks, Batch Normalization, GlobalAveragePooling2D, and two fully connected layers equipped with dropout and L2 regularization. To address class imbalance within the dataset, balanced class weights were applied during the model training process.

The best-performing model achieved an accuracy of 79.93% and a precision of 80.88% on the test data, demonstrating strong generalization capability toward previously unseen data. The application of balanced class weights proved to contribute significantly to the overall improvement of model performance. The system was subsequently deployed using the Flask framework as a REST API with an interactive web interface, enabling real-time music mood prediction that is straightforward and accessible. The findings of this study can be utilized by music application developers, streaming platforms, researchers in the field of Music Information Retrieval (MIR), and practitioners of technology-based music therapy. For future research, it is recommended to explore deeper architectures such as ResNet or audio-based transformer models like the Audio Spectrogram Transformer (AST), expand the number of mood classes, and leverage audio data augmentation techniques to further improve model accuracy and robustness across more diverse real-world audio conditions.

Keywords : Music Mood Classification, MFCC, Mel-Spectrogram, Convolutional Neural Network, Class Weight