

Klasifikasi Emosi Manusia Berdasarkan Suara dengan Deep Learning

TUGAS AKHIR

**Diajukan untuk memenuhi Salah Satu Persyaratan Mencapai Derajat
Sarjana Komputer**



Dibuat oleh:

William Juvent

210711067

**PROGAM STUDI INFORMATIKA
FAKULTAS TEKNOLOGI INDUSTRI
UNIVERSITAS ATMA JAYA
YOGYAKARTA
2025**

HALAMAN PENGESAHAN

Tugas Akhir Berjudul

KLASIFIKASI EMOSI MANUSIA BERDASARKAN SUARA DENGAN DEEP
LEARNING

yang disusun oleh

William Juvent

210711067

dinyatakan telah memenuhi syarat pada tanggal 10 Juni 2025

		Keterangan
Dosen Pembimbing 1	: Aloysius Gonzaga Pradnya Sidhawara, S.T., M.Eng.	Telah Menyetujui
Dosen Pembimbing 2	: Prof. Dr. Pranowo, S.T., M.T.	Telah Menyetujui
Tim Penguji		
Penguji 1	: Aloysius Gonzaga Pradnya Sidhawara, S.T., M.Eng.	Telah Menyetujui
Penguji 2	: Prof. Dr. Ir. Alb. Joko Santoso, M.T.	Telah Menyetujui
Penguji 3	: Dra. Ernawati, M.T., Ph.D.	Telah Menyetujui

Yogyakarta, 10 Juni 2025

Universitas Atma Jaya Yogyakarta

Teknologi Industri

Dekan

ttd.

Dr. Ir. Parama Kartika Dewa SP., S.T., M.T.

Dokumen ini merupakan dokumen resmi UAJY yang tidak memerlukan tanda tangan karena dihasilkan secara elektronik oleh Sistem Bimbingan UAJY. UAJY bertanggung jawab penuh atas informasi yang tertera di dalam dokumen ini

HALAMAN PERSEMBAHAN

Veni, Vidi, Vici

KATA PENGANTAR

Dalam Kesempatan ini, penulis ingin menyampikan puji dan syukur kepada tuhan yang maha esa atas rahmat-nya, sehingga penulis dapat menyelesaikan Tugas Akhir berjudul "Klasifikasi Emosi Berdasarkan Suara manusia dengan Deep Learning". Tugas Akhir ini merupakan salah satu persyaratan untuk memperoleh gelar sarjana dari Program Studi Informatika, Universitas Atma Jaya Yogyakarta. Dalam proses pengerjaan Tugas Akhir ini, penulis tidak dapat berjalan sendiri. Oleh karena itu, penulis ingin mengucapkan terima kasih yang sebesar besarnya kepada:

- A. Bapak A. Gonzaga Pradnya Sidhawara, S.T., M.Eng. dan Prof. Dr. Pranowo, S.T., M.T. selaku dosen pembimbing yang telah memberikan arahan, masukan, bimbingan dan bantuan mulai dari penulisan hingga tahap pengujian tugas akhir ini. Sehingga, tugas akhir ini dapat diselesaikan dengan baik oleh penulis.
- B. Seluruh saudara dan kedua Orang Tua yang selalu memberikan dukungan dan doa selama pengerjaan tugas akhir ini.
- C. Teman-teman terdekat yang selalu hadir memberi dukungan dan semangat selama penulisan tugas akhir ini.
- D. Member Lesserafim, Heart2hearts, Illit, Aespa, New Jeans, NMIXX, IVE, Hearts2Hearts dan BabyMonster yang selalu menemani selama pengerjaan tugas akhir.

Penulis menyadari bahwa tugas akhir ini masih jauh dari kata sempurna. Penulis memohon maaf apabila terdapat kesalahan maupun kekurangan pada penulisan tugas akhir ini. Penulis berharap tugas akhir ini dapat bermanfaat bagi pembaca.

Yogyakarta, April 2025



William Juvent

210711067

DAFTAR ISI

HALAMAN PENGESAHAN.....	ii
HALAMAN PERSEMBAHAN	iii
KATA PENGANTAR	iv
DAFTAR ISI.....	v
DAFTAR GAMBAR	ix
DAFTAR TABEL	xv
Daftar Persamaan	xvi
BAB I PENDAHULUAN	18
A. Latar Belakang	18
B. Rumusan Masalah	18
C. Batasan Masalah.....	21
D. Tujuan Penelitian	22
E. Metode Penelitian.....	23
1. <i>Data loading</i>	23
2. <i>Preprocessing Data</i>	23
3. <i>Train val test split</i>	24
4. Pelatihan Model	24
5. Evaluasi Model.....	25
6. Implementasi Sistem	25
6. Sistematika Penulisan	25
BAB I Pendahuluan	25
BAB II Tinjauan Pustaka	25
BAB III Landasan Teori	25
BAB IV Dataset dan Pengembangan Model	26

BAB V Implementasi Model dan Pengujian Sistem	26
BAB II TINJAUAN PUSATAKA.....	27
BAB III LANDASAN TEORI.....	36
1. Emosi	36
2. Enam Emosi Dasar	37
3. Deep Learning.....	37
<i>a. Artificial Neural Network.....</i>	<i>38</i>
4. <i>Convolutional Neural Network (CNN)</i>	<i>39</i>
5. <i>Recurrent Neural Networks (RNN)</i>	<i>47</i>
a. Long Short-Term Memory (LSTM).....	47
<i>b. Gate Reccurent Unit (GRU).....</i>	<i>53</i>
6. Dataset yang digunakan	54
a. Ravdess	55
b. Crema-D.....	55
c. Savee	56
d. Tess	56
e. IEMOCAP	57
7. Augmentasi Data.....	57
<i>a. Noise.....</i>	<i>58</i>
<i>b. Pitch Shifting.....</i>	<i>59</i>
<i>c. Stretch</i>	<i>60</i>
<i>d. Time shifting.....</i>	<i>61</i>
<i>e. Speed changing</i>	<i>62</i>
8. <i>Feature Extraction</i>	<i>63</i>
<i>a. Mel Frequency Cepstral Coefficient (MFCC)</i>	<i>64</i>
<i>b. Root Mean Square (RMS) Energy</i>	<i>65</i>

c.	<i>Zero Crossing Rate (ZCR)</i>	65
d.	<i>Spectral</i>	66
e.	<i>Chroma Short-time Fourier transform (Chroma stft)</i>	69
9.	<i>Feature Scaling</i>	70
10.	<i>Hyperparameter</i>	71
a.	<i>Learning Rate</i>	72
b.	<i>Batch Size</i>	72
c.	<i>Momentum</i>	73
d.	<i>Regularization Hyperparameters</i>	73
e.	<i>Number of Training Epochs</i>	75
f.	<i>Learning Rate Decay</i>	75
g.	<i>Activation Function</i>	76
h.	<i>Jumlah Lapisan (Layer)</i>	77
i.	<i>Jumlah Node per Layer</i>	77
j.	<i>Loss Function</i>	78
k.	<i>Optimizer</i>	79
11.	<i>Streamlit</i>	82
BAB IV DATASET DAN PENGEMBANGAN MODEL.....		83
A.	<i>Deskripsi Problem</i>	83
B.	<i>Dataset</i>	83
1.	<i>Analisis Data</i>	83
2.	<i>Preprocessing Data</i>	91
C.	<i>Pengembangan Model</i>	95
1.	<i>Pelatihan dan Evaluasi Model</i>	101
2.	<i>Pengujian Model</i>	104
BAB V IMPLEMENTASI MODEL DAN PENGUJIAN SISTEM.....		116

A. Implementasi Model.....	116
B. Implementasi Sistem	117
C. Pengujian Sistem.....	121
BAB VI PENUTUP	146
A. Kesimpulan	146
B. Saran.....	146
DAFTAR PUSTAKA	148

DAFTAR GAMBAR

Gambar 1. 1 <i>Waveform</i> suara dengan sumbu-Y menunjukkan amplitudo, dan sumbu-X menunjukkan jumlah sampel.....	19
Gambar 1. 2 <i>Waveform</i> suara dengan sumbu-Y menunjukkan amplitudo, dan sumbu-X menunjukkan waktu (dalam detik).....	20
Gambar 1. 3 <i>Spectrogram</i> suara dengan sumbu-Y menunjukan frekuensi (Hz), sumbu-X menunjukan waktu (dalam detik), dan warna menunjukan energi dari frekuensi.....	20
Gambar 1. 4 alur pengerjaan	23
Gambar 3. 1 arsitektur CNN sederhana, sumber: Modern Convolutional Neural Network Architectures	41
Gambar 3. 2 Arsitektur LSTM, sumber: Insights into LSTM architecture Thorir Mar Ingolfsson	48
Gambar 3. 3 arsitektur BiLSTM, sumber: BiLSTM Explained Papers With Code	51
Gambar 3. 4 arsitektur GRU, sumber: The Architecture of the gated recurrent unit (GRU) cell Download Scientific Diagram	53
Gambar 3. 5 Perbedaan suara Original dan suara dengan tambahan <i>Noise</i> , sumber: Data Augmentation for Audio. Data Augmentation by Edward Ma Medium ...	59
Gambar 3. 6 Perbedaan suara Original dan suara dengan tambahan <i>Pitch Shifting</i> , sumber: Data Augmentation for Audio. Data Augmentation by Edward Ma Medium	60
Gambar 3. 7 Perbedaan suara Original dan suara setelah di <i>stretch</i>	61
Gambar 3. 8 Perbedaan suara Original dan suara setelah <i>time shifting</i> , sumber: Data Augmentation for Audio. Data Augmentation by Edward Ma Medium ...	62
Gambar 3. 9 Perbedaan suara Original dan suara setelah <i>Speed Changing</i> , sumber: Data Augmentation for Audio. Data Augmentation by Edward Ma Medium ...	63
Gambar 4. 1 <i>Identifier</i> pada <i>dataset</i> Ravdess, sumber: Speech Emotion Recognition (en)	84
Gambar 4. 2 <i>Identifier</i> pada dataset CREMA-D dan TESS, sumber: Speech Emotion Recognition (en).....	85

Gambar 4. 3 <i>Identifier</i> pada dataset SAVEE, sumber: Speech Emotion Recognition (en)	85
Gambar 4. 4 Implementasi kode untuk Filtrasi dataset.....	86
Gambar 4. 5 <i>Path dataset</i> sebelum difiltrasi.....	87
Gambar 4. 6 <i>Path dataset</i> setelah difiltrasi.....	87
Gambar 4. 7 Implementasi kode pemanggilan fungsi <i>filter_dataset</i>	87
Gambar 4. 8 Grafik jumlah sebaran dataset setelah dilakukan filtrasi.....	88
Gambar 4. 9 implementasi kode augmentasi data.....	89
Gambar 4. 10 Implementasi kode untuk <i>feature extraction</i>	90
Gambar 4. 11 Implementasi kode untuk feature extraction pada data augmentasi.....	90
Gambar 4. 12 implementasi kode tambahan untuk <i>feature extraction</i>	91
Gambar 4. 13 implementasi kode load data untuk <i>preprocessing data</i>	92
Gambar 4. 14 implementasi kode fungsi pemanggilan kode preprocessing data	92
Gambar 4. 15 implementasi kode <i>feature extraction</i> di dalam preprocessing data	93
Gambar 4. 16 implementasi kode <i>label encoder</i> di dalam <i>preprocessing data</i>	94
Gambar 4. 17 Implementasi kode pemanggilan fungsi untuk <i>preprocessing data</i>	94
Gambar 4. 18 keseluruhan implementasi <i>kode preprocessing</i>	95
Gambar 4. 19 keseluruhan implementasi <i>kode preprocessing</i>	96
Gambar 4. 20 implementasi kode untuk <i>optimizer</i>	97
Gambar 4. 21 implementasi kode untuk input layer dan layer model CNN	98
Gambar 4. 22 implementasi kode untuk <i>layer</i> model LSTM.....	99
Gambar 4. 23 implementasi kode untuk <i>layer</i> model GRU.....	99
Gambar 4. 24 implementasi kode untuk <i>layer</i> model BiLSTM.....	99
Gambar 4. 25 implementasi kode untuk <i>Fully Connected Layer</i> atau <i>dense</i>	100
Gambar 4. 26 implementasi kode untuk <i>output layer</i>	100
Gambar 4. 27 implementasi kode <i>optimizer</i> dan <i>compile model</i>	101
Gambar 4. 28 implementasi kode untuk <i>load data</i> dan pembagian data sebelum <i>training</i>	101
Gambar 4. 29 implementasi kode untuk load <i>feature scaler</i>	101
Gambar 4. 30 implementasi kode untuk <i>feature scaling</i> terhadap data yang sudah di-load	102
Gambar 4. 31 percobaan <i>training</i> dengan <i>epoch 75</i>	103
Gambar 4. 32 implementasi kode untuk <i>training</i> menggunakan GPU	103
Gambar 4. 33 implementasi kode untuk evaluasi model	104
Gambar 4. 34 hasil <i>Precision</i> , <i>Recall</i> dan <i>f1-score</i> CNN x BiLSTM+ADAM+Robust.	107

Gambar 4. 35 grafik <i>accuracy training</i> dan <i>Validation</i> CNN x BiLSTM+ADAM+Robust	108
Gambar 4. 36 grafik <i>Loss training</i> dan <i>Validation</i> CNN x BiLSTM+ADAM+Robust..	109
Gambar 4. 37 hasil <i>Precision</i> , <i>Recall</i> dan <i>f1-score</i> CNN x LSTM+SGD+Minmax.....	109
Gambar 4. 38 grafik <i>accuracy training</i> dan <i>Validation</i> CNN x LSTM+SGD+Minmax	110
Gambar 4. 39 grafik <i>loss training</i> dan <i>Validation</i> CNN x LSTM+SGD+Minmax	111
Gambar 4. 40 hasil <i>Precision</i> , <i>Recall</i> dan <i>f1-score</i> CNN x GRU+SGD+Minmax	111
Gambar 4. 41 grafik <i>accuracy training</i> dan <i>Validation</i> CNN x GRU+SGD+Minmax ..	112
Gambar 4. 42 grafik <i>Loss training</i> dan <i>Validation</i> CNN x GRU+SGD+Minmax.....	113
Gambar 4. 43 hasil <i>Precision</i> , <i>Recall</i> dan <i>f1-score</i> CNN x BiLSTM+SGD+Minmax....	113
Gambar 4. 44 grafik <i>accuracy training</i> dan <i>Validation</i> CNN x BiLSTM+SGD+Minmax	114
Gambar 4. 45 grafik <i>Loss training</i> dan <i>Validation</i> CNN x BiLSTM+SGD+Minmax....	115
Gambar 5. 1 implementasi kode untuk load model.....	116
Gambar 5. 2 implementasi kode untuk memastikan panjang fitur konsisten	116
Gambar 5. 3 implementasi kode <i>extract_feature</i> yang sama seperti pada proses preprocessing.....	117
Gambar 5. 4 implementasi kode untuk <i>upload</i> file suara.....	118
Gambar 5. 5 implementasi kode untuk menyimpan suara yang telah <i>upload</i> ..	118
Gambar 5. 6 implementasi kode untuk menggunakan <i>extract_feature</i>	118
Gambar 5. 7 implementasi kode untuk menggunakan <i>scaler</i>	119
Gambar 5. 8 implementasi kode untuk melakukan prediksi	119
Gambar 5. 9 implementasi kode untuk menampilkan hasil prediksi	120
Gambar 5. 10 tampilan utama website	120
Gambar 5. 11 tampilan halaman dokumentasi	121
Gambar 5. 12 prediksi pada data Ses01F_impro01_F012 representatif emosi marah.....	123
Gambar 5. 13 prediksi pada data Ses01F_impro01_M011 representatif emosi marah.....	123
Gambar 5. 14 prediksi pada data Ses01F_impro04_F030 representatif emosi marah.....	124
Gambar 5. 15 prediksi pada data Ses01F_impro05_F017 representatif emosi	

marah.....	124
Gambar 5. 16 prediksi pada data Ses01F_1_script01_F035 representatif emosi marah.....	125
Gambar 5. 17 prediksi pada data Ses01F_script01_1_F042 representatif emosi marah.....	125
Gambar 5. 18 prediksi pada data Ses01F_script01_2_M017 representatif emosi marah.....	126
Gambar 5. 19 prediksi pada data Ses01F_script02_1_F007 representatif emosi marah.....	126
Gambar 5. 20 prediksi pada data Ses01M_script03_2_F037 representatif emosi marah.....	127
Gambar 5. 21 prediksi pada data Ses01F_impro02_2_F016 representatif emosi marah.....	127
Gambar 5. 22 prediksi pada data Ses01F_script01_3_F011 representatif emosi senang.....	128
Gambar 5. 23 prediksi pada data Ses01F_script01_3_F013 representatif emosi senang.....	129
Gambar 5. 24 prediksi pada data Ses01F_script01_3_F017 representatif emosi senang.....	129
Gambar 5. 25 prediksi pada data Ses01F_script01_3_F019 representatif emosi senang.....	130
Gambar 5. 26 prediksi pada data Ses01F_script01_3_F020 representatif emosi senang.....	130
Gambar 5. 27 prediksi pada data Ses01F_script01_3_F021 representatif emosi senang.....	131
Gambar 5. 28 prediksi pada data Ses01F_script01_3_F035 representatif emosi senang.....	131
Gambar 5. 29 prediksi pada data Ses01F_script01_3_M008 representatif emosi senang.....	132
Gambar 5. 30 prediksi pada data Ses01F_script02_2_F042 representatif emosi senang.....	132

Gambar 5. 31 prediksi pada data Ses01F_script02_2_F043 representatif emosi senang.....	133
Gambar 5. 32 prediksi pada data Ses05M_script01_1b_F022 representatif emosi netral.....	134
Gambar 5. 33 prediksi pada data Ses05M_impro08_M027 representatif emosi netral.....	134
Gambar 5. 34 prediksi pada data Ses05M_impro08_M013 representatif emosi netral.....	135
Gambar 5. 35 prediksi pada data Ses05M_impro05_M009 representatif emosi netral.....	135
Gambar 5. 36 prediksi pada data Ses05M_impro05_F020 representatif emosi netral.....	136
Gambar 5. 37 prediksi pada data Ses05M_impro04_F027 representatif emosi netral.....	136
Gambar 5. 38 prediksi pada data Ses05F_script03_2_F006 representatif emosi netral.....	137
Gambar 5. 39 prediksi pada data Ses05F_impro08_M016 representatif emosi netral.....	137
Gambar 5. 40 prediksi pada data Ses05F_impro08_F020 representatif emosi netral	138
Gambar 5. 41 prediksi pada data Ses05F_impro08_F005 representatif emosi netral	138
Gambar 5. 42 prediksi pada data Ses02M_impro06_M010 representatif emosi sedih	139
Gambar 5. 43 prediksi pada data Ses02F_impro02_F008 representatif emosi sedih	140
Gambar 5. 44 prediksi pada data Ses01M_script01_3_M020 representatif emosi sedih	140
Gambar 5. 45 prediksi pada data Ses01M_script01_1_M038 representatif emosi sedih	141
Gambar 5. 46 prediksi pada data Ses01M_script01_1_F043 representatif emosi	

sedih	141
Gambar 5. 47 prediksi pada data Ses01M_impro02_M017 representatif emosi	
sedih	142
Gambar 5. 48 prediksi pada data Ses01F_script01_3_M027 representatif emosi	
sedih	142
Gambar 5. 49 prediksi pada data Ses02M_impro06_M010 representatif emosi	
sedih	143
Gambar 5. 50 prediksi pada data Ses01F_impro02_F013 representatif emosi sedih	
.....	143
Gambar 5. 51 prediksi pada data Ses01F_impro02_F000 representatif emosi sedih	
.....	144

DAFTAR TABEL

Tabel 2. 1 Tabel Perbandingan Penelitian.....	31
Tabel 4. 1 tabel hasil evaluasi model CNN dan LSTM	106
Tabel 4. 2 tabel hasil evaluasi model CNN dan GRU	106
Tabel 4. 3 tabel hasil evaluasi model CNN dan BiLSTM	106
Tabel 5. 1 tabel data uji dari dataset IEMOCAP.....	121
Tabel 5. 2 Tabel hasil uji data berlabel "marah"	128
Tabel 5. 3 Tabel hasil uji data berlabel "senang"	133
Tabel 5. 4 Tabel hasil uji data berlabel "netral"	139
Tabel 5. 5 Tabel hasil uji data berlabel "sedih"	144
Tabel 5. 6 hasil pengujian terhadap dataset	145

Daftar Persamaan

Persamaan 3. 1 Persamaan ukuran <i>Output</i>	42
Persamaan 3. 2 Persamaan <i>batch normalization layer</i>	44
Persamaan 3. 3 persamaan operasi konvolusi satu dimensi	45
Persamaan 3. 4 persamaan <i>equivariant representations</i>	46
Persamaan 3. 5 persamaan operasi konvolusi dua dimensi	46
Persamaan 3. 6 persamaan <i>cell state</i>	48
Persamaan 3. 7 <i>hidden state</i>	48
Persamaan 3. 8 persamaan <i>cell input</i>	48
Persamaan 3. 9 persamaan <i>candidate cell input</i>	48
Persamaan 3. 10 persamaan <i>input gate</i>	49
Persamaan 3. 11 persamaan <i>pre-activation</i> untuk <i>input gate</i>	49
Persamaan 3. 12 persamaan <i>forget gate</i>	49
Persamaan 3. 13 persamaan <i>Pre-activation</i> untuk <i>Forget gate</i>	49
Persamaan 3. 14 <i>Output Gate</i>	49
Persamaan 3. 15 <i>Pre-activation</i> untuk <i>output gate</i>	49
Persamaan 3. 16 persamaan fungsi aktivasi sigmoid	50
Persamaan 3. 17 persamaan <i>foward hidden state</i>	52
Persamaan 3. 18 persamaan <i>backward hidden state</i>	52
Persamaan 3. 19 persamaan <i>output</i> akhir BiLSTM	52
Persamaan 3. 20 persamaan <i>update gate</i>	54
Persamaan 3. 21 persamaan <i>reset gate</i>	54
Persamaan 3. 22 persamaan <i>candidate activation</i>	54
Persamaan 3. 23 persamaan <i>New Hidden State</i>	54
Persamaan 3. 24 persamaan gerbang <i>output</i>	54
Persamaan 3. 25 persamaan konversi frekuensi linear ke skala Mel	64
Persamaan 3. 26 persamaan koefisien MFCC.....	64
Persamaan 3. 27 persamaan nilai RMS <i>energy</i>	65
Persamaan 3. 28 persamaan ZCR.....	66
Persamaan 3. 29 persamaan <i>spectral centroid</i>	67
Persamaan 3. 30 persamaan <i>rolloff</i>	68

Persamaan 3. 31 persamaan spectral <i>bandwidth</i>	69
Persamaan 3. 32 persamaan <i>chroma short-time fourier transform</i>	70
Persamaan 3. 33 persamaan <i>Z-score normalization</i>	70
Persamaan 3. 34 persamaan minmaxscaler	71
Persamaan 3. 35 persamaan robustscaler	71
Persamaan 3. 36 persamaan momentum	73
Persamaan 3. 37 persamaan L2 <i>Regularization</i> atau <i>weight decay</i>	74
Persamaan 3. 38 persamaan ReLU	76
Persamaan 3. 39 persamaan softmax	77
Persamaan 3. 40 persamaan Categorical Cross-Entropy	78
Persamaan 3. 41 persamaan momentum pada ADAM	80
Persamaan 3. 42 persamaan RMS pada ADAM	80
Persamaan 3. 43 persamaan <i>Bias Correction</i> pada ADAM	80
Persamaan 3. 44 persamaan parameter <i>update</i> pada ADAM	80
Persamaan 3. 45 persamaan SGD	81
Persamaan 3. 46 persamaan RMSProp	81