

**KLASIFIKASI IRAMA AZAN BERBASIS *MEL FREQUENCY*
CEPSTRAL COEFFICIENTS MENGGUNAKAN
*MULTILAYER PERCEPTRON***

SKRIPSI



**UIN SUNAN AMPEL
S U R A B A Y A**

Disusun Oleh:

**HAFID RAMADHAN
H96219046**

**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA
2023**

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : Hafid Ramadhan

NIM : H96219046

Program Studi : Sistem Informasi

Angkatan : 2019

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul: "KLASIFIKASI IRAMA AZAN BERBASIS MEL FREQUENCY CEPSTRAL COEFFICIENTS MENGGUNAKAN MULTILAYER PERCEPTRON". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 14 Juli 2023

Yang menyatakan,



Hafid Ramadhan
NIM. H96219046

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh:

NAMA : HAFID RAMADHAN

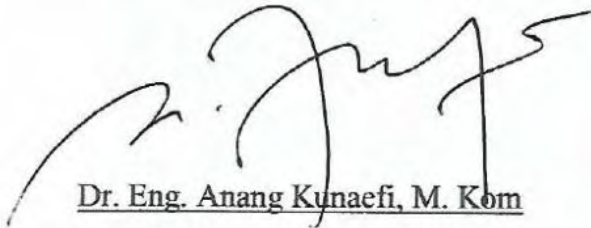
NIM : H96219046

JUDUL : KLASIFIKASI IRAMA AZAN BERBASIS *MEL*
FREQUENCY CEPSTRAL COEFFICIENTS
MENGUNAKAN *MULTILAYER PERCEPTRON*

Ini telah diperiksa dan disetujui untuk diujikan.

Surabaya, 13 Juli 2023

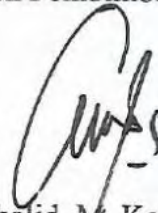
Dosen Pembimbing 1



Dr. Eng. Anang Kunaefi, M. Kom

NIP. 197911132014031001

Dosen Pembimbing 2



Khalid, M. Kom

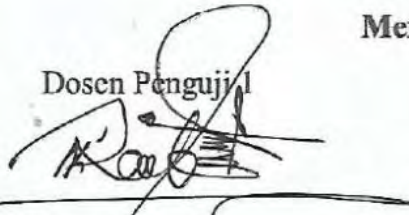
NIP. 197906092014031002

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi Hafid Ramadhan ini telah dipertahankan
di depan tim penguji skripsi
di Surabaya, 14 Juli 2023

Menyetujui,

Dosen Penguji 1



Mujib Ridwan, S.Kom., M.T

NIP. 198604272014031004

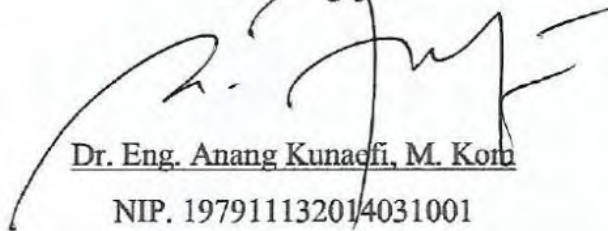
Dosen Penguji 2



Dwi Rolliawati, MT

NIP. 197909272014032001

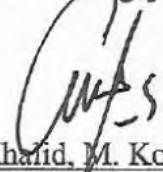
Dosen Penguji 3



Dr. Eng. Anang Kunaefi, M. Kom

NIP. 197911132014031001

Dosen Penguji 4



Khalid, M. Kom

NIP. 197906092014031002

Mengetahui,

Dekan Fakultas Sains dan Teknologi
UIN Sunan Ampel Surabaya



Hamdani, M.Pd.

NIP. 196507312000031002



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : Hafid Ramadhan
NIM : H96219046
Fakultas/Jurusan : Sains dan Teknologi / Sistem Informasi
E-mail address : h96219046@student.uinsby.ac.id

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

Sekripsi Tesis Desertasi Lain-lain (.....)
yang berjudul : KLASIFIKASI IRAMA AZAN BERBASIS MEL FREQUENCY CEPSTRAL

COEFFICIENTS MENGGUNAKAN MULTILAYER PERCEPTRON

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 26 Juli 2023

Penulis

(Hafid Ramadhan)

ABSTRAK

KLASIFIKASI IRAMA AZAN BERBASIS *MEL FREQUENCY CEPSTRAL COEFFICIENTS* MENGGUNAKAN *MULTILAYER PERCEPTRON*

Oleh:

Hafid Ramadhan

Sebagai penanda masuknya waktu salat fardhu yang dikumandangkan lima kali setiap harinya, azan menjadi syiar Islam yang paling tampak secara umum. Dari banyaknya variasi irama azan, yang terjadi di lingkungan sekitar masih banyak masyarakat atau bahkan beberapa muazin yang belum mengetahui atau mengenal perihal *maqam* azan yang dilantunkan. *Voice recognition* berbasis *Mel Frequency Cepstral Coefficient* (MFCC) dan klasifikasi menggunakan *Multilayer Perceptron* (MLP) digunakan untuk mengenali azan dalam lima irama berbeda diantaranya *Rast*, *Bayati*, *Ajam*, *Nahawand* dan *Hijaz*. Penelitian ini dilakukan untuk Mengukur performa pengenalan irama azan menggunakan metode MFCC dan MLP. Hasil yang didapatkan setelah implementasi model adalah tingkat akurasi sebesar 56%. Dari hasil tersebut diambil kesimpulan bahwa performa MFCC dan MLP dalam mengenali irama azan kurang baik dengan indikasi terjadinya *overfitting* pada model klasifikasi yang digunakan.

Kata kunci: Azan, Klasifikasi Audio, *Machine Learning*, *Mel Frequency Cepstral Coefficient*, *Multilayer Perceptron*.

ABSTRACT

CLASSIFYING MAQAMAT OF ADHAN BASED ON MEL FREQUENCY CEPSTRAL COEFFICIENT USING MULTILAYER PERCEPTRON

By:

Hafid Ramadhan

The adhan is delivered five times a day to remind for mandatory prayer, it become the most visible form of Islamic propagation in general. Of the many variations of the adhan *maqamat*, what happens in the surrounding environment is that there are still many people or even some muadhin who did not know or recognize the *maqam* of the adhan being performed. Voice recognition based on Mel Frequency Cepstral Coefficient (MFCC) and Multilayer Perceptron (MLP) classification is used to recognize the adhan in five different *maqam* including *Rast*, *Bayati*, *Ajam*, *Nahawand* and *Hijaz*. This research was conducted to measure the performance of *maqamat* of the adhan recognition using MFCC and MLP methods. The result obtained after the implementation of the model is an accuracy rate of 56%. From these results it is concluded that the performance of MFCC and MLP in recognizing the *maqamat* of the adhan was not good enough with an indication of overfitting in the used classification model.

Kata kunci: Adhan, Audio Classification, Machine Learning, Mel Frequency Cepstral Coefficient, Multilayer Perceptron.

DAFTAR ISI

PERNYATAAN KEASLIAN.....	i
LEMBAR PERSETUJUAN PEMBIMBING	ii
PENGESAHAN TIM PENGUJI SKRIPSI.....	iii
LEMBAR PERSETUJUAN PUBLIKASI.....	iv
ABSTRAK	v
ABSTRACT	vi
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR	x
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Perumusan Masalah.....	4
1.3. Batasan Masalah.....	4
1.4. Tujuan Penelitian.....	4
1.5. Manfaat Penelitian.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1. Tinjauan Penelitian Terdahulu	6
2.2. Dasar Teori	8
2.2.1. Irama Azan	8
2.2.2. <i>Voice Recognition</i>	10
2.2.3. <i>Mel Frequency Cepstral Coefficient (MFCC)</i>	11
2.2.4. <i>Machine Learning</i>	14
2.2.5. <i>Multiclass Classification</i>	15
2.2.6. <i>Artificial Neural Network (ANN)</i>	16
2.2.7. Evaluasi Model.....	18
2.3. Integrasi Keilmuan	20
BAB III METODOLOGI PENELITIAN.....	22
3.1. Desain Penelitian.....	22
3.2. Uraian Desain Penelitian	22
3.2.1. Identifikasi Masalah	22

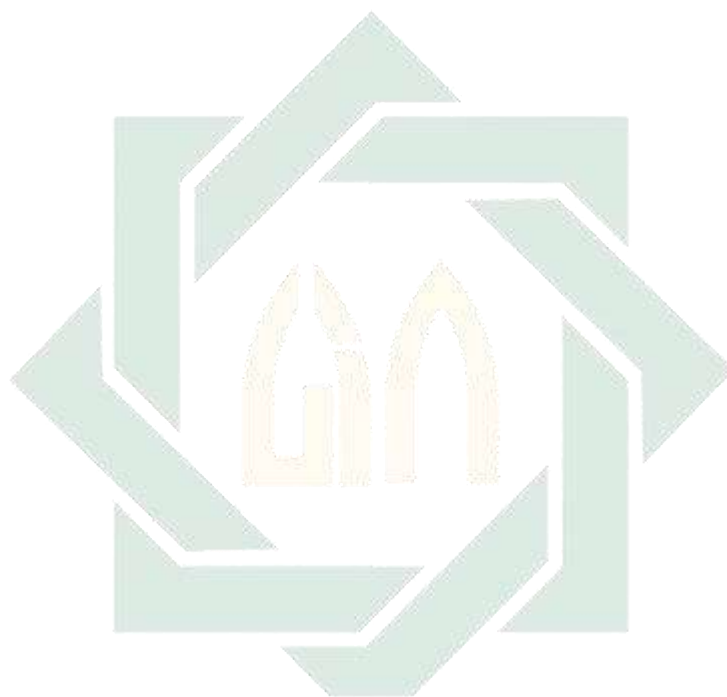
DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....	6
Tabel 3.1 Muazin Kontributor.....	23
Tabel 3.2 Parameter <i>MLPClassifier</i>	27
Tabel 4.1 Kode Memuat File Audio	29
Tabel 4.2 Kode <i>Pre-emphasis</i>	30
Tabel 4.3 Kode Ekstraksi Fitur MFCC	33
Tabel 4.4 Kode Fungsi “features_extractor”.....	34
Tabel 4.5 Kode Perintah Perulangan.....	35
Tabel 4.6 Hasil Ekstraksi Fitur MFCC	35
Tabel 4.7 Kode <i>Data Splitting</i>	37
Tabel 4.8 Parameter <i>MLPClassifier</i>	37
Tabel 4.9 Kode Klasifikasi MLP	38
Tabel 4.10 Kode Pemrosesan Hasil	38
Tabel 4.11 <i>Classification Report</i> 70:30	39
Tabel 4.12 <i>Classification Report</i> 80:20	40
Tabel 4.13 <i>Classification Report</i> 90:10	42
Tabel 4.14 Hasil Skenario Model Klasifikasi	43
Tabel 4.15 Kode Klasifikasi <i>Support Vector Machine</i>	44
Tabel 4.16 <i>Classification Report Support Vector Machine</i>	44
Tabel 4.17 Kode Klasifikasi <i>Random Forest</i>	46
Tabel 4.18 <i>Classification Report Random Forest</i>	47
Tabel 4.19 Kode Klasifikasi <i>XGBoost</i>	48
Tabel 4.20 <i>Classification Report XGBoost</i>	49
Tabel 4.21 Hasil Klasifikasi Pembanding.....	50

DAFTAR GAMBAR

Gambar 2.1 Pendekatan <i>Machine Learning</i>	15
Gambar 2.2 <i>Layer Artificial Neural Network</i>	17
Gambar 2.3 Model Jaringan <i>Backpropagation</i>	17
Gambar 2.4 <i>K-fold Cross Validation</i>	19
Gambar 2.5 <i>Confusion Matrix</i>	19
Gambar 3.1 Diagram Alur Penelitian.....	22
Gambar 3.2 Sampel <i>Hamming Window</i>	25
Gambar 3.3 Domain Waktu dan Frekuensi.....	25
Gambar 3.4 <i>Mel Filterbank</i>	26
Gambar 4.1 Folder Data Irama Azan	29
Gambar 4.2 <i>Waveform</i> Sampel Irama Azan.....	30
Gambar 4.3 <i>Pre-emphasis</i>	31
Gambar 4.4 <i>Frame Blocking</i>	31
Gambar 4.5 <i>Frame Setekah Windowing</i>	32
Gambar 4.6 <i>Fast Fourier Transform</i>	32
Gambar 4.7 <i>Mel Frequency Spectrogram</i>	33
Gambar 4.8 MFCCs	34
Gambar 4.9 MFCCs Sampel Irama Azan	36
Gambar 4.10 <i>Confusion Matrix</i> 70:30	39
Gambar 4.11 <i>Learning Curve</i> 70:30	40
Gambar 4.12 <i>Confusion Matrix</i> 80:20	41
Gambar 4.13 <i>Learning Curve</i> 80:20	41
Gambar 4.14 <i>Confusion Matrix</i> 90:10	42
Gambar 4.15 <i>Learning Curve</i> 90:10	43
Gambar 4.16 <i>Confusion Matrix Support Vector Machine</i>	45
Gambar 4.17 <i>Learning Curve Support Vector Machine</i>	45

Gambar 4.18 <i>Confusion Matrix Random Forest</i>	47
Gambar 4.19 <i>Learning Curve Random Forest</i>	48
Gambar 4.20 <i>Confusion Matrix XGBoost</i>	49
Gambar 4.21 <i>Learning Curve XGBoost</i>	50
Gambar 4.22 <i>Learning Curve Klasifikasi</i>	51



UIN SUNAN AMPEL
S U R A B A Y A

BAB I

PENDAHULUAN

1.1. Latar Belakang

Azan adalah pengumuman tentang masuknya waktu salat fardhu dengan menggunakan lafaz-lafaz tertentu (Raya & Mulia, 2003). Azan merupakan syiar Islam terbesar, tanda agama paling nyata yang terus menerus dikekalkan oleh Rasulullah SAW., baik di kala beliau bermukim, bahkan di kala beliau dalam keadaan safar sekalipun, sejak tahun pertama hijrah ketika kali pertama disyariatkan hingga beliau wafat (Ash-Shiddieqy, 1994). Sebagai penanda masuknya waktu salat fardhu yang dikumandangkan lima kali setiap harinya, azan menjadi syiar Islam yang paling tampak secara umum. Seruan azan dapat terdengar baik oleh kaum muslim atau non muslim. Diperuntukkan pada kaum non muslim, azan dapat menjadi media dakwah ampuh yang menandakan keindahan dan ketenangan agama Islam.

Karakteristik dan kriteria muadzin terbaik adalah mereka yang memiliki suara keras, musikalitas, kerapian, merdu, lembut dan keaslian suara yang menarik perhatian pendengar (Wan Abdullah & Mat Razali, 2015). Azan dilantunkan atas dasar *maqamat* (bentuk jamak dari *maqam*) Arab (Al Hushari, 2002). Koleksi repertoar mencakup sekitar 50 *maqamat*, tujuh diantaranya bersifat mendasar. Tujuh *maqam* tersebut adalah *Ajam*, *Bayati*, *Hijaz*, *Nahawand*, *Rast*, *Saba* dan *Sikah* (Qasim, 1989). Muazin sebagai orang yang mengumandangkan salah satu syiar agama harus memenuhi kriteria tertentu serta memperhatikan pelafalan lafaz azan. Selain pelafalan lafaz azan yang benar, irama yang digunakan saat melantunkan azan menjadi faktor penting untuk menarik perhatian pendengar. Dari banyaknya variasi irama azan yang telah disebutkan, yang terjadi di lingkungan sekitar masih banyak masyarakat atau bahkan beberapa muazin yang belum mengetahui atau mengenal perihal *maqam* azan yang dilantunkan.

Machine learning adalah sebuah cabang dari sains komputer yang secara umum bertujuan untuk memungkinkan komputer “belajar” tanpa diprogram secara langsung (Samuel, 2000). *Voice recognition* adalah proses mengenali sebuah sinyal suara secara otomatis dengan cara melakukan perbandingan sinyal

suara yang menjadi referensi atau acuan terhadap pola karakteristiknya (Setiawan & Ariyanto, 2015). Proses pengenalan suara dilakukan melalui beberapa tahapan yaitu input suara, *pre-processing* suara, *feature extraction*, dan klasifikasi (Tirumala et al., 2017). Dengan pola pembelajaran mesin yang digunakan, *voice recognition* merupakan salah satu bentuk implementasi metode *machine learning*. Dalam penerapan *voice recognition* dibutuhkan data uji yang merupakan hasil ekstraksi fitur data suara untuk melakukan pengenalan atau disebut tahap klasifikasi.

Berdasarkan hasil penelitian, analisis spektral dalam waktu singkat mencakup *Linear Prediction Cepstral Coefficients* (LPCC), *Mel Frequency Cepstral Coefficients* (MFCC) dan *Perceptual Linear Prediction* (PLP) merupakan metode yang umum digunakan untuk ekstraksi informasi dari sinyal suara (López-Ruiz, 2018). Menurut jenis data yang diproses, algoritma dibagi menjadi algoritma data terstruktur dan algoritma data tidak terstruktur. Data tersusun mencakup algoritma *Artificial Neural Network* (ANN) dan *FM-Deep Learning*. Algoritma data tidak terstruktur mencakup *Recurrent Neural Network* (RNN), *Convolutional Neural Network* (CNN) (Yu et al., 2022). Beberapa percobaan berhasil dilakukan dalam pengembangan sistem yang dapat mengenali sinyal suara. Beberapa percobaan tersebut dilakukan dengan berbagai macam kombinasi metode ekstraksi ciri dan klasifikasi yang telah disebutkan.

Dari berbagai metode ekstraksi fitur spektral berbasis *Discrete Fourier Transform* (DFT), MFCC adalah algoritma populer dan telah banyak digunakan untuk aplikasi pengenalan pembicara (Sahidullah & Saha, 2012). MFCC memposisikan pita frekuensi secara logaritmik, posisi tersebut mendekati sistem respon manusia dibanding sistem lainnya (Janse et al., 2014). ANN memiliki keunggulan berdaya guna untuk klasifikasi pola statis (Khamlich et al., 2021). *Multilayer Perceptron* merupakan salah satu arsitektur ANN yang paling banyak digunakan (Sen et al., 2020). Dengan rekam jejak dan karakteristik metode MFCC dan ANN dalam berbagai penelitian, tahap ekstraksi ciri atau pola sinyal yang digunakan pada penelitian ini adalah MFCC. Sedangkan untuk tahap klasifikasi hasil ekstraksi pola suara digunakan metode ANN dengan arsitektur *Multilayer Perceptron* (MLP).

Beberapa penelitian *voice recognition* dengan metode ekstraksi pola MFCC dan klasifikasi ANN telah dilakukan sebelumnya. Penelitian sistem identifikasi pembicara Thailand dilakukan dengan membandingkan metode ekstraksi pola suara LPCC, PFL, MFCC. Dengan sistem pencocokan pola DTW dengan 5-NN, MFCC mendapatkan nilai rata-rata tingkat identifikasi tertinggi sebesar 86.74%. Penelitian dilanjutkan dengan menerapkan metode MFCC untuk perbandingan metode klasifikasi GMM, DTW dan MLP, dengan hasil menunjukkan nilai rata-rata tingkat identifikasi tertinggi sebesar 99% untuk GMM (Tanprasert & Achariyakulporn, 2000). Pada penelitian Shahriar & Tariq mengenai klasifikasi *maqam* pembacaan Al Quran dengan *deep learning*, ekstraksi MFCC dan klasifikasi ANN dengan 5 *hidden layers* menunjukkan persentase akurasi 95,7% menjadi yang tertinggi dibanding dengan metode CNN, *Long Short Term Memory* (LSTM) dan kombinasi CNN-LSTM (Shahriar & Tariq, 2021).

Penelitian sebelumnya mengenai pengenalan *nagham azan* dilakukan menggunakan metode *Mellin Transform* dan DWT (*Discrete Wavelet Transform*). Hasil penelitian menunjukkan perbandingan kinerja kedua metode dengan nilai *true detection* sebesar 73,33% untuk metode DWT dan 80% untuk *Mellin Transform* (Rizal et al., 2018). Penelitian tersebut hanya berfokus pada perbandingan dua metode ekstraksi suara yakni DWT dan *Mellin Transform*. Dalam laporan penelitian tidak dijelaskan mengenai metode yang digunakan untuk klasifikasi hasil ekstraksi pola suara. Belum diterapkan metode klasifikasi *machine learning* pada hasil ekstraksi pola suara pada penelitian tersebut. Nilai sensitivitas yang sangat tipis menimbulkan kelemahan pada sistem dengan sampel suara latih yang beragam.

Dengan memanfaatkan perpaduan ekstraksi fitur MFCC dan klasifikasi MLP, penelitian ini diharapkan dapat memprediksi dan mengenalkan beragam irama *azan*. Semakin dikenalnya setiap irama *azan* dengan warna nada, irama dan emosi masing-masing dapat meluaskan syiar agama islam. Berdasarkan latar belakang yang telah diuraikan, maka dilakukan penelitian berjudul **“KLASIFIKASI IRAMA AZAN BERBASIS MEL FREQUENCY**

CEPSTRAL COEFFICIENTS MENGGUNAKAN MULTILAYER PERCEPTRON

1.2. Perumusan Masalah

Berdasarkan uraian latar belakang, maka rumusan masalah dari penelitian ini adalah:

1. Bagaimana ekstraksi fitur suara irama azan menggunakan metode MFCC?
2. Bagaimana klasifikasi irama azan menggunakan MLP?
3. Bagaimana performa pengenalan irama azan menggunakan metode MFCC dan MLP?
4. Bagaimana performa pengenalan irama azan menggunakan MLP dibandingkan dengan menggunakan metode klasifikasi lain?

1.3. Batasan Masalah

Pada laporan ini, diperlukan batasan masalah agar pembahasan jelas dan tidak bertentangan dari judul yang sudah ditentukan. Batasan masalah dalam laporan ini adalah:

1. Data yang digunakan adalah data primer berupa file audio azan dalam irama *Rast, Bayati, Ajam, Nahawand* dan *Hijaz* yang dikumpulkan dengan ekstraksi audio dari sumber yang ada di situs internet dan perekaman suara azan oleh enam muazin masjid di Kota Surabaya.
2. Setiap irama azan memiliki 50 data audio dengan lafaz yang ditentukan dalam penelitian ini adalah *hayya alash shalah* pertama yang dilantunkan.
3. File audio memiliki durasi 20 detik dengan frekuensi sampel 44100 Hz, *channel stereo* dan disimpan dalam ekstensi WAV.
4. Metode klasifikasi pembandingan meliputi metode *Support Vector Machine, Random Forest*, dan *XGBoost*.

1.4. Tujuan Penelitian

Dari rumusan masalah yang telah ditetapkan, penelitian ini memiliki tujuan untuk:

1. Mengimplementasikan ekstraksi pola suara irama azan dengan menggunakan metode MFCC.
2. Mengimplementasikan klasifikasi irama azan dengan menggunakan MLP.

3. Mengukur performa pengenalan irama azan menggunakan metode MFCC dan MLP.
4. Mengukur performa pengenalan irama azan menggunakan metode MLP dengan perbandingan metode klasifikasi lain.

1.5. Manfaat Penelitian

Berdasarkan tujuan penelitian, penelitian ini diharapkan dapat memberikan manfaat, di antaranya:

1. Secara Akademis
 - a. Sebagai bentuk kontribusi dalam ilmu pengetahuan sebagai referensi mengenai ekstraksi pola suara menggunakan metode MFCC dan klasifikasi menggunakan MLP.
 - b. Sebagai bentuk pengintegrasian ilmu pengetahuan teknologi dengan keislaman.
2. Secara Aplikatif
 - a. Implementasi nilai akurasi pengenalan irama azan menggunakan metode MFCC dan MLP pada penelitian pemrosesan suara berikutnya.
 - b. Mampu memprediksi azan dalam irama Rast, Bayati, Ajam, Nahawand dan Hijaz dengan karakteristik dan keragaman masing-masing.

UIN SUNAN AMPEL
S U R A B A Y A

	sebesar 95.7%. Nilai akurasi tersebut merupakan yang terbesar dibandingkan dengan algoritma <i>deep learning</i> CNN dan LSTM.
Penelitian 3	
Judul Penelitian dan Nama Penulis	“Intelligent Arabic Letters Speech Recognition System Based On Mel Frequency Cepstral Coefficients” (Quteishat et al., 2023)
Tujuan Penelitian	Penelitian ini bertujuan untuk mendapatkan performa optimal dengan mempelajari pengaruh perubahan atribut <i>multi-layer perceptron</i> (MLP) ANN.
Hasil Penelitian dan Kesimpulan	Tahap penelitian dibagi menjadi dua tahap. Tahap pertama merubah domain waktu rekaman suara menjadi domain frekuensi dengan FFT, dilanjutkan dengan ekstraksi fitur menggunakan MFCC. Tahap kedua mengklasifikasikan hasil ekstraksi fitur menggunakan MLP ANN dengan algoritma pembelajaran <i>backpropagation</i> . Hasil menunjukkan klasifikasi menggunakan dua <i>hidden layer</i> pada <i>neural network</i> mendapatkan nilai akurasi tertinggi sekitar 86%.
Penelitian 4	
Judul Penelitian dan Nama Penulis	“Klasifikasi Mood Musik Berdasarkan Mel Frequency Cepstral Coefficients Dengan Backpropagation Neural Network” (Maulana et al., 2022)
Tujuan Penelitian	Klasifikasi musik berdasarkan empat jenis <i>mood</i> menggunakan pemrosesan sinyal audio MFCC dan algoritma klasifikasi BPNN (<i>Backpropagation Neural Network</i>).
Hasil Penelitian dan Kesimpulan	Dengan jumlah data sebanyak 200 sampel suara, klasifikasi hasil ekstraksi MFCC menggunakan BPNN dengan 1 <i>hidden layer</i> 256 <i>neuron</i> mendapatkan nilai akurasi sebesar 87,67%. Perbedaan durasi, panjang <i>frame</i> dan besar <i>overlap</i> dapat mempengaruhi hasil klasifikasi.
Penelitian 5	
Judul Penelitian dan Nama Penulis	“Footstep Recognition Using Mel Frequency Cepstral Coefficients and Artificial Neural Network” (Wulandari Siagian et al., 2020)
Tujuan Penelitian	Mengusulkan sebuah sistem pengenalan langkah kaki baru menggunakan ekstraksi fitur MFCC dan algoritma klasifikasi ANN pada dataset rekaman langkah kaki dengan total 500 rekaman suara.
Hasil Penelitian dan Kesimpulan	Hasil klasifikasi 500 data suara rekaman menunjukkan sistem usulan dapat memperoleh hasil akurasi tertinggi sebesar 92 %.

digunakan saat mengumandangkan azan, selama maulid (peringatan hari lahir Nabi Muhammad SAW.) dan begitu pula dengan bacaan Al Quran (Shokouhi & Yusof, 2013). *Maqam* atau dalam beberapa penelitian lain disebut *nagham* dapat diartikan sebagai irama. Irama yang digunakan ketika mengumandangkan azan memiliki tujuan untuk memperindah bacaan agar menarik perhatian pendengar. Selain hanya menarik perhatian disebutkan, bahwa setiap *maqam* memiliki ciri tersendiri yang dapat mempengaruhi hati orang yang mendengarkan.

Muazin mengumandangkan azan menggunakan irama yang telah diajarkan oleh para guru muazin. Seiring berjalannya waktu, perkembangan *maqam* mengarah pada munculnya corak gaya lokal yang ditandai variasi pada beberapa unsur budaya yang disesuaikan dengan lingkungan sekitar (Al Bakri et al., 2019). Koleksi repertoar mencakup sekitar 50 *maqamat*, tujuh diantaranya bersifat mendasar. Tujuh *maqam* tersebut adalah *Ajam*, *Bayati*, *Hijaz*, *Nahawand*, *Rast*, *Saba* dan *Sikah* (Qasim, 1989). Lantunan azan menggunakan irama telah menjadi ajaran tradisi turun-menurun dari guru kepada muridnya. Improvisasi, latar belakang, dan budaya yang berbeda dari masing-masing muazin dapat menghasilkan variasi irama azan baru. Namun dari banyaknya macam irama azan telah disepakati, bahwa hanya tujuh irama yang bersifat mendasar.

Bayati paling sering digunakan di timur tengah dan tersebar luas di Madinah dan Saudi Arabia, *maqam Rast* dan *Hijaz* dapat didengar dari muazin terkenal (Shokouhi & Yusof, 2013). Penelitian ini menggunakan lima dari tujuh irama azan dasar, diantaranya irama *Rast*, *Bayati*, *Ajam*, *Nahawand* dan *Hijaz*. Lima irama azan tersebut dipilih berdasarkan pertimbangan popularitas irama dan perihal ketersediaan data untuk dikumpulkan. Penjelasan mengenai lima irama yang digunakan pada penelitian ini sebagai berikut.

A. *Rast*

Irama *Rast* adalah jenis irama dasar dan paling dominan. Irama *Rast* memiliki durasi yang lebih cepat dari irama lain, sehingga banyak digunakan muazin untuk mengumandangkan azan (Salim, 2004). Irama ini memiliki karakteristik yang kuat, maskulin dan mewakili kebersihan pikiran. Dengan membawakan irama *Rast* secara tulus, perasaan kagum akan kebesaran Allah SWT. dapat dibagikan dan dirasakan pendengar.

B. Bayati

Irama *Bayati* merupakan *maqam* terluas dalam bacaan Al Quran, *Bayati* termasuk dalam bagian turunan dari 12 *maqam* Iran (Shokouhi & Yusof, 2013). Irama *Bayati* sangat populer di dunia tilawati Al Quran sehingga menjadi standar *maqam* yang ditempatkan sebagai irama pertama dalam tradisi melagukan bacaan Al Quran oleh *qurra* senior Mesir (Misbachul Munir, 1997). Irama *Bayati* mewakili perasaan yang sedih, kelembutan, kenyamanan, keagungan dan sedikit kegembiraan.

C. Ajam

Irama *Ajam* dinamakan dari bahasa Arab untuk “orang Iran”, Irama ini digunakan pada acara bahagia seperti liburan, pernikahan dan acara gembira lainnya (Shokouhi & Yusof, 2013). Irama ini sering dilantunkan pada saat takbir lebaran Idulfitri dan Iduladha (Masrurin, 2018). Irama *Ajam* atau *Jiharkah* mewakili perasaan senang yang mendalam dengan lantunan yang merdu untuk didengar.

D. Nahawand

Karakteristik sedih irama *Nahawand* sesuai untuk melagukan syair atau ayat Al Quran yang memiliki nuansa kesedihan (Salim, 2004). Irama *Nahawand* membawakan perasaan keseriusan dilantunkan tidak terlalu kasar dan tidak terlalu ringan.

E. Hijaz

Irama *Hijaz* merepresentasikan tarik suara khas ketimuran, memiliki kesan indah dan irama yang bersifat dasar. Irama ini sering dikumandangkan para penggembala unta di gurun menurut sebagian orang (Hasan, 2021). Irama *Hijaz* dinamakan dari tempat asalnya yakni wilayah sebelah barat laut Arab Saudi yang mencakup kota Mekkah dan Madinah. Irama memiliki ciri alunan tegas dan bersemangat yang sesuai dengan para *qari* bersuara tinggi dan lantang.

2.2.2. Voice Recognition

Voice recognition merupakan pengembangan sistem dan teknik yang menjadikan mungkin sebuah komputer untuk menerima *input* berupa kata yang diucapkan (Rusdi & Yani, 2018). Perangkat mengenali suara melalui proses digitalisasi kata dan mengklasifikasikannya dengan pola tertentu yang telah

α = konstanta filter *pre-emphasis*

B. *Frame Blocking*

Frame blocking membagi suara ke dalam beberapa *frame* dengan tujuan memudahkan perhitungan dan analisa suara, setiap *frame* mencakup beberapa sampel sesuai dengan panjang durasi sampel suara dan besaran frekuensi *sampling* (Basuki, 2009). Durasi *frame* yang membagi setiap sampel ke dalam beberapa *frame* berdasarkan waktu terletak di antara 20 sampai 40 ms (Chamidy, 2016). *Frame* panjang dimaksudkan untuk mendapat resolusi frekuensi yang baik, sedangkan *frame* pendek ditujukan untuk mendapat ranah waktu terbaik (Heriyanto et al., 2018).

C. *Windowing*

Windowing merupakan proses mengalihkan *frame* menjadi suara kontinu setelah suara terbagi pada proses *frame blocking*. Tahap *windowing* dilakukan untuk menghindari kesalahan data pada proses *fourier transform* (Riyanto & Sutejo, 2014). Representasi fungsi *windowing* dapat digunakan menggunakan persamaan 2.2.

$$X(n) = f_i(n) w(n) \quad (2.2)$$

Keterangan:

$X(n)$ = hasil *windowing*

f_i = hasil *frame blocking*

n = jumlah sampel dalam masing-masing *frame*

$w(n)$ = fungsi *windowing*

Hamming windowing merupakan fungsi *window* yang paling sering digunakan dalam *voice recognition* berfungsi untuk menghasilkan *side lobe level* rendah, selain itu *noise* yang ditimbulkan tidak terlampaui besar (Adriana, 2013). Persamaan 2.3 berikut merupakan rumus *hamming windowing*.

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{(N-1)}\right) \quad (2.3)$$

Keterangan:

$w(n)$ = fungsi *windowing*

$n = 1, 2, 3, \dots$

N = panjang *frame*

D. Fast Fourier Transform (FFT)

Fast Fourier Transform (FFT) merupakan cara cepat melakukan metode *Discrete Fourier Transform* (DFT) yang digunakan untuk konversi dari *spatial domain* ke bentuk *frequency domain* (Gupta et al., 2013). Tahap FFT mengubah setiap sampel *frame* N dalam bentuk domain waktu menjadi domain frekuensi berfungsi untuk mengubah respon gelombang saluran suara dan konvolusi getaran celah suara dalam domain waktu (Chamidy, 2016). Persamaan 2.4 dapat digunakan untuk metode FFT.

$$T(k) = \sum_{n=0}^{N-1} X(n) \cos\left(\frac{2\pi kn}{N}\right) - \sum_{n=0}^{N-1} X(n) \sin\left(\frac{2\pi kn}{N}\right) \quad (2.4)$$

Keterangan:

$T(k)$ = hasil perhitungan FFT ke- k

$X(n)$ = hasil perhitungan *windowing* ke- n

n = nomor urut sinyal

k = indeks frekuensi

E. Mel Frequency Wrapping (Filterbank)

Mel Frequency Wrapping adalah sebuah *filter* berupa *filterbank* untuk mengetahui besaran energi dari *band* frekuensi tertentu dalam sinyal suara (Nasution, 2012). *Filterbank* mempunyai reaksi frekuensi melalui jalur segitiga dengan jarak dan besaran yang ditetapkan interval frekuensi konstan. Tahap ini digunakan untuk memperoleh *mel spectrum* (Marwala, 2012). *Mel spectrum* didapatkan menggunakan persamaan 2.5.

$$Y[i] = \sum_{j=1}^G T[j] H_i[j] \quad (2.5)$$

Keterangan:

$Y[i]$ = hasil perhitungan *Mel Frequency Wrapping* ke- i

$T[j]$ = hasil FFT

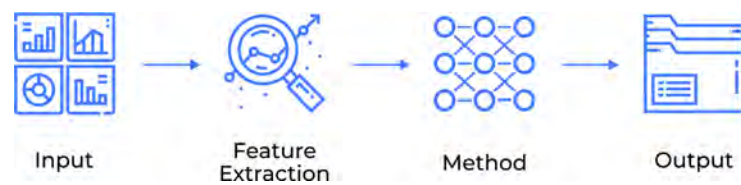
$H_i[j]$ = koefisien *filterbank* pada frekuensi j ($1 \leq i \leq E$)

G = jumlah *magnitude spectrum* ($G \in N$)

E = jumlah *channel* dalam *filterbank*

Setiap *output* filter adalah jumlah dari komponen *spectral*, berikutnya digunakan persamaan 2.6 untuk menghitung *mel* dalam satuan frekuensi Hz.

dalam *machine learning*. Pendekatan *machine learning* digambarkan pada Gambar 2.1.



Gambar 2.1 Pendekatan *Machine Learning*

Model *machine learning* dapat dibagi berdasarkan sifat pelabelan data menjadi *supervised*, *unsupervised*, dan *semi-supervised*. *Supervised learning* ditujukan untuk memperkirakan *input-output* yang belum diketahui berdasarkan sampel yang telah diketahui, dimana *output* diberi label (misalnya regresi dan klasifikasi). Pada *Unsupervised learning* yang diberikan ke sistem pembelajaran hanya sampel *input* (misalnya *clustering* dan estimasi probabilitas). *Semi-supervised learning* merupakan gabungan kedua pembelajaran sebelumnya, dimana sebagian saja data yang memiliki label (misalnya sistem pencarian teks/gambar) (El Naqa & Murphy, 2015).

2.2.5. *Multiclass Classification*

Klasifikasi merupakan proses mengelompokkan objek dengan ciri atau karakteristik sama ke dalam beberapa kelas (Kolcz et al., 2001). Klasifikasi adalah metode *machine learning* guna membagi atau mengklasifikasikan objek berdasar pada ciri tertentu, seperti ketika manusia membedakan suatu benda dengan yang lain (Ahmad, 2017). Algoritma *supervised learning* membutuhkan data *train* berlabel dalam jumlah besar untuk membangun model dengan performa prediksi tinggi (Hady & Schwenker, 2013). Proses klasifikasi membutuhkan data yang memiliki label dikarenakan klasifikasi termasuk ke dalam algoritma *supervised learning*. Proses pelabelan data bertujuan untuk membagi data ke dalam beberapa kelas. Dalam proses pelabelan data diperlukan keterlibatan ahli atau pakar yang sesuai dengan data yang akan diberi label. Keterlibatan ahli ditujukan agar data dikelompokkan ke dalam kelas sesuai dengan ketentuan kelas tersebut.

Berdasarkan jumlah dan sifat kelas, klasifikasi dapat dibagi menjadi *binary*, *multiclass*, *multilabel*, dan *multilevel/hierarchical*. *Multiclass classification* adalah kondisi apabila terdapat lebih dari dua kelas. Pada *multiclass*

classification data yang telah diklasifikasikan ke kelas tertentu tidak mungkin berada pada kelas lainnya (Putra, 2020). *Multiclass classification* dapat dilakukan dengan merepresentasikan *output unit* dari *multilayer perceptron*. *Output unit multilayer perceptron* yang dapat berjumlah lebih dari satu unit, dapat direpresentasikan sesuai dengan jumlah kelas yang dibutuhkan.

2.2.6. *Artificial Neural Network (ANN)*

Dapat dipertimbangkan *Artificial Neural Network (ANN)* sebagai model sederhana dari struktur biologi jaringan *neural*. Sebuah ANN terdiri dari unit pemrosesan yang saling terhubung. Model umum dari sebuah unit pemrosesan terdiri dari bagian penjumlahan diikuti bagian *output* (Yegnanarayana, 2009). Di antara berbagai metode *machine learning*, *artificial neural network* dengan *hidden layer* menunjukkan keunggulan kinerja untuk regresi dan klasifikasi data gambar besar (Kwon et al., 2020). ANN merupakan sebuah metode yang terinspirasi dari cara otak manusia bekerja. ANN dapat digunakan untuk menyelesaikan tugas kompleks seperti pemahaman pola, regresi dan klasifikasi. Terdapat tiga hal yang menentukan ANN diantaranya komponen, arsitektur jaringan dan fungsi aktivasi.

Secara umum terdapat tiga layer dalam ANN yakni *input*, *hidden* dan *output* digambarkan seperti Gambar 2.2. *Input layer* menerima masukan kemudian diteruskan ke *hidden unit*, di dalam *hidden units*, dilakukan pemrosesan dan perhitungan hasil fungsi aktivasi setiap *neuron*, hasil pemrosesan kemudian diteruskan menuju *output layer* (Putra, 2020). *Neuron* pada ANN memiliki fungsi yang sama dengan 10 miliar *neuron* yang terhubung dalam otak manusia, *neuron* berfungsi untuk menerima, memproses dan mengirimkan informasi (Sydenham & Thorn, 2005). Terdapat beberapa bentuk arsitektur dari ANN diantaranya *Single-Layer Feedforward*, *Multiple-Layer Feedforward*, *Recurrent/Feedback* dan *Mesh Architecture* (da Silva et al., 2017). Dari arsitektur ANN yang disebutkan, arsitektur *Multiple-Layer Feedforward* digunakan pada penelitian ini.

Fase pertama sinyal masukan diteruskan ke *hidden layer*, *output* setiap unit *hidden layer* Z_j dipropagasikan maju menuju *hidden layer* di atas. Propagasi dijalankan dengan fungsi aktivasi yang telah ditetapkan. Tahapann ini dilakukan terus hingga *output* jaringan Y_k dihasilkan, hasil *output* dibandingkan dengan target T_k . Fase kedua berdasarkan perbandingan $Y_k T_k$ dilakukan perhitungan faktor δ_k ($k=1, 2, \dots, m$) yang digunakan untuk distribusi kesalahan kembali ke *hidden layer*. Perhitungan dilakukan di setiap unit *hidden layer* hingga faktor δ pada *hidden layer* terhubung dengan unit *input*. Pada tahap ketiga dilakukan modifikasi pada bobot garis secara bersamaan. Perubahan bobot garis berdasar pada faktor δ *neuron* pada *layer* atas (Jumarwanto et al., 2009).

2.2.7. Evaluasi Model

Salah satu proses dalam *machine learning* adalah melatih model dengan data yang tersedia. Sebuah model mungkin dapat menampilkan kinerja yang cukup baik pada data pelatihan. Namun, kinerja dan hasil yang diperoleh dapat berbeda ketika model dilatih menggunakan set data yang lain (Yadav & Shukla, 2016). Untuk mengatasi masalah tersebut *K-fold Cross Validation* dapat digunakan, dengan membagi set data menjadi beberapa bagian (*fold*) dan memastikan bahwa setiap bagian menjadi set pengujian (Peryanto et al., 2020). Data dibagi menjadi data *training* dan *testing* dengan perbandingan 80:20, kemudian dilakukan pembagian data *training* sejumlah bagian (*fold*) K . Pengujian dilakukan berulang sesuai dengan jumlah K dimana setiap bagian bergantian menjadi set data pengujian. Gambar 2.4 merupakan ilustrasi implementasi *K-fold Cross Validation* dengan jumlah K sebanyak lima.

Apabila sistem memprediksi sebuah kelas positif ketika kelas sebenarnya adalah kelas negatif, kondisi *false positive* (FP) dinyatakan. *False negative* (FN) adalah kondisi sistem tidak berhasil memprediksi kelas negatif ketika kelas aktual adalah kelas positif. Nilai dalam *confusion matrix* dapat digunakan untuk evaluasi klasifikasi dengan menggunakan rumus presisi, *recall*, akurasi dan *f-measure* pada persamaan 2.8, 2.9, 2.10, 2.11 berikut:

$$Precision = \frac{TP}{TP+FP} \quad (2.8)$$

$$Recall = \frac{TP}{TP+FN} \quad (2.9)$$

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (2.10)$$

$$F-Measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (2.11)$$

2.3. Integrasi Keilmuan

Integrasi keilmuan menjelaskan bagaimana korelasi antara penelitian yang dilakukan dengan sudut pandang kajian ayat Al Quran dan hadis. Integrasi keilmuan ditulis berdasarkan hasil wawancara dengan Ustaz H. Musyafa', M.Pd.I, M.Pd, yang merupakan tokoh masyarakat sekaligus salah satu pengurus Yayasan Lembaga Pendidikan Islam Al Hikmah Surabaya. Wawancara langsung dilakukan pada hari Rabu tanggal 8 Maret 2023.

Berdasarkan hasil wawancara, narasumber mengingatkan kembali mengenai fungsi dan tujuan dasar mengumandangkan azan. Azan adalah panggilan ibadah umat muslim atau pertanda masuk waktu salat dan termasuk dalam syariat agama. Muazin sebagai orang yang mengumandangkan azan sebaiknya orang yang memiliki suara lantang dan merdu, seperti halnya kisah Bilal bin Robah sebagai seorang muazin Nabi Muhammad SAW. Hal ini didasarkan pada hadis riwayat Tirmidzi nomor 147 mengenai sejarah azan sebagai berikut.

لَمَّا أَصْبَحْنَا أَتَيْنَا رَسُولَ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ فَأَخْبَرْتُهُ بِالرُّؤْيَا فَقَالَ إِنَّ هَذِهِ لَرُّؤْيَا حَقٌّ فَقُمَّ مَعَ بِلَالٍ فَإِنَّهُ أَنْدَى وَأَمْدٌ صَوْتًا مِنْكَ فَالِقِ عَلَيْهِ مَا قِيلَ لَكَ وَلِيُنَادِ بِدَلِّكَ

Terjemahan: “Pada suatu pagi kami mendatangi Rasulullah shallallahu ‘alaihi wasallam, aku lalu mengabarkan kepadanya perihal mimpi yang aku alami, beliau

lalu bersabda: “Ini mimpi yang benar maka berdirilah bersama Bilal, karena suaranya lebih lantang dan panjang daripada kamu, sampaikanlah padanya apa yang dikatakan padamu (impikan), suruhlah agar ia menyeru dengannya.” (HR. Tirmidzi no. 147)

Penggunaan salah satu bentuk perkembangan teknologi untuk mengenali irama azan dengan tujuan mengenalkan keragaman irama azan merupakan salah satu bentuk dakwah dalam syiar agama Islam. Dakwah merupakan perbuatan dengan tujuan menyerukan ajaran agama Islam dan mengingatkan orang untuk beriman dan takwa kepada Allah SWT. Diperintahkan kepada umat muslim untuk berdakwah, Dalam Al Quran surat An Nahl ayat 125 Allah SWT. berfirman.

أَدْعُ إِلَى سَبِيلِ رَبِّكَ بِالْحِكْمَةِ وَالْمَوْعِظَةِ الْحَسَنَةِ وَجَادِبْهُمْ بِالَّتِي هِيَ أَحْسَنُ إِنَّ رَبَّكَ هُوَ
أَعْلَمُ بِمَنْ ضَلَّ عَنْ سَبِيلِهِ وَهُوَ أَعْلَمُ بِالْمُهْتَدِينَ (١٢٥)

Terjemahan: “Serulah (manusia) ke jalan Tuhanmu dengan hikmah dan pengajaran yang baik serta debatlah mereka dengan cara yang lebih baik. Sesungguhnya Tuhanmu Dialah yang paling tahu siapa yang tersesat dari jalan-Nya dan Dia (pula) yang paling tahu siapa yang mendapat petunjuk.” (Q.S An Nahl 16:125)

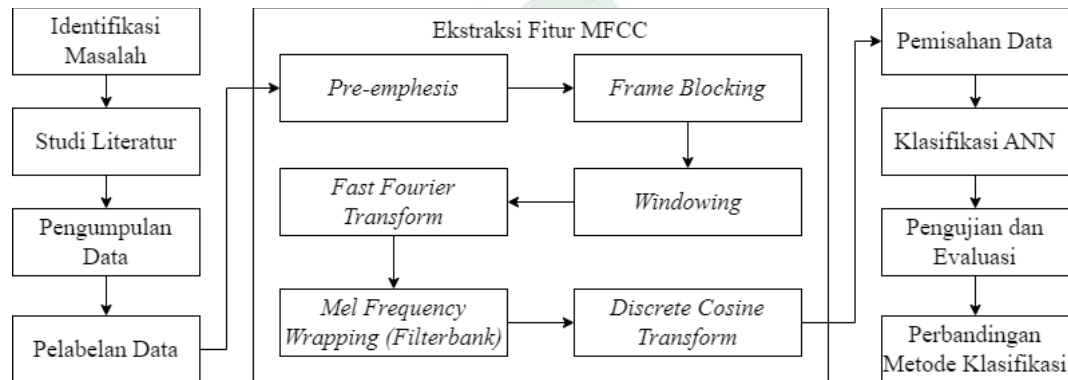
UIN SUNAN AMPEL
S U R A B A Y A

BAB III

METODOLOGI PENELITIAN

3.1. Desain Penelitian

Alur penelitian “Klasifikasi Irama Azan Berbasis *Mel Frequency Cepstral Coefficient* Menggunakan Metode *Artificial Neural Network*” digambarkan pada diagram alur Gambar 3.1.



Gambar 3.1 Diagram Alur Penelitian

3.2. Uraian Desain Penelitian

Sub bab ini menguraikan setiap tahap yang dilakukan pada penelitian ini mulai dari identifikasi masalah sampai dengan analisa hasil.

3.2.1. Identifikasi Masalah

Permasalahan yang mendasari penelitian ini selaras dengan penjabaran latar belakang dan rumusan masalah pada bab satu. Masalah yang dirumuskan menjadi landasan dalam melaksanakan penelitian. Masalah yang diangkat dalam penelitian ini adalah bagaimana ekstraksi fitur MFCC, klasifikasi MLP pada sampel audio irama azan. Kemudian dilakukan pengukuran tingkat akurasi pengenalan irama azan menggunakan MFCC dan MLP.

3.2.2. Studi Literatur

Dalam tahap studi literatur, pencarian dan pembelajaran literatur terdahulu dilakukan. Pelaksanaan tahap studi literatur merupakan upaya untuk mendapat referensi dan pemahaman mengenai konsep, metode dan algoritma relevan yang digunakan pada penelitian ini. Literatur didapatkan dari beberapa buku dan jurnal

Setiap sampel data audio irama azan dipindahkan ke dalam lima folder sesuai dengan kelasnya. Kelas data dan folder penyimpanan dinamakan sesuai dengan nama jenis irama azan yang digunakan yakni *rast*, *bayati*, *ajam*, *nahawand* dan *hijaz*. Nama folder kemudian digunakan dalam pelabelan hasil ekstraksi menggunakan pengulangan *python* dan fungsi *os.listdir*.

3.2.5. Ekstraksi Fitur MFCC

Proses ekstraksi fitur dilakukan menggunakan metode MFCC untuk memperoleh ciri suatu sampel suara. Ekstraksi fitur dilakukan pada semua data suara irama azan yang telah dikelompokkan berdasarkan kelas atau label yang ditentukan. Proses ekstraksi fitur MFCC melalui beberapa tahap diantaranya *pre-emphasis*, *frame blocking*, *windowing*, FFT, *Mel Frequency Wrapping* dan DCT. Tahapan MFCC dilakukan menggunakan *library* pemrosesan musik dan audio *Librosa*.

A. *Pre-emphasis*

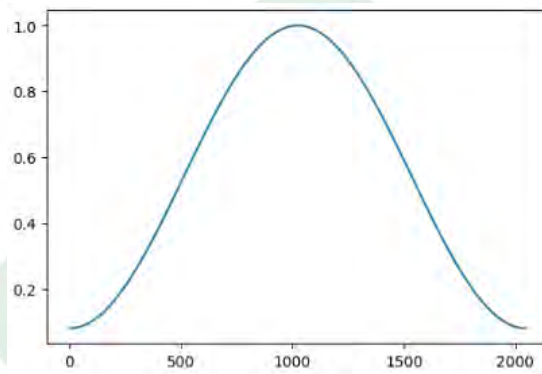
Pre-emphasis dilakukan untuk memproses suara melalui filter penekanan frekuensi dengan meningkatkan amplitudo frekuensi tinggi dan mengurangi amplitudo pada frekuensi yang lebih rendah. *Pre-emphasis* bertujuan untuk mengurangi efek *noise* dengan mengurangi nilai sinyal yang akan datang dan operasi log di bagian akhir. *Pre-emphasis* dijalankan menggunakan fitur *librosa.effects.preemphasis* dengan *default* nilai konstanta *pre-emphasis* sebesar 0.97. Besaran konstanta *pre-emphasis* sesuai dengan implementasi *Hidden Markov Model Toolkit* yang digunakan pada MFCCs.

B. *Frame Blocking*

Tahap *frame blocking* memiliki tujuan untuk memecah suara ke dalam beberapa *frame*. *Frame blocking* memudahkan perhitungan serta analisa suara dan menghindari timbulnya efek distorsi pada saat proses FFT. Jumlah *frame* sesuai dengan durasi suara sampel dan besaran frekuensi dengan ukuran yang sama dengan jumlah FFT. Panjang *frame* ditentukan dengan perhitungan nilai *sample rate* audio dan *hop length*. Dengan menjalankan fitur pada *Librosa*, didapatkan nilai bawaan FFT sebanyak 2048 dan *hop length* sebesar 512.

C. Windowing

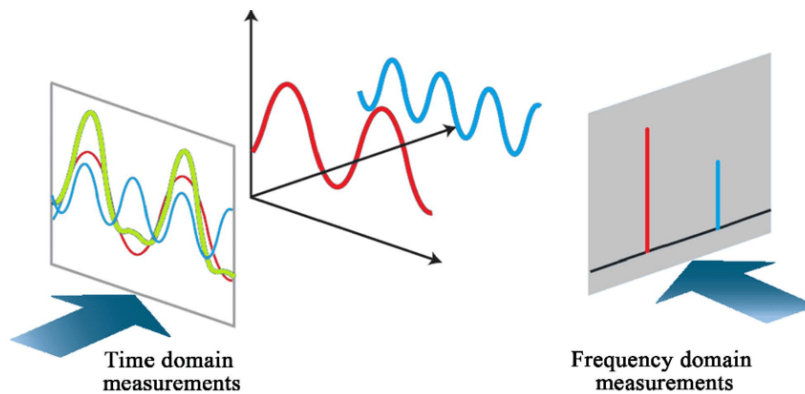
Windowing dilakukan guna mengubah *frame* yang terbagi pada tahap *frame blocking* menjadi suara kontinu. *Windowing* merupakan proses lanjutan dari *frame blocking* untuk menghindari kesalahan atau efek distorsi pada saat proses FFT. *Windowing* dijalankan menggunakan fungsi *hamming window* (Gambar 3.2). Penggunaan *hamming window* ditujukan untuk mendapatkan hasil *side lobe cancelling* yang lebih baik.



Gambar 3.2 Sampel *Hamming Window*

D. Fast Fourier Transform (FFT)

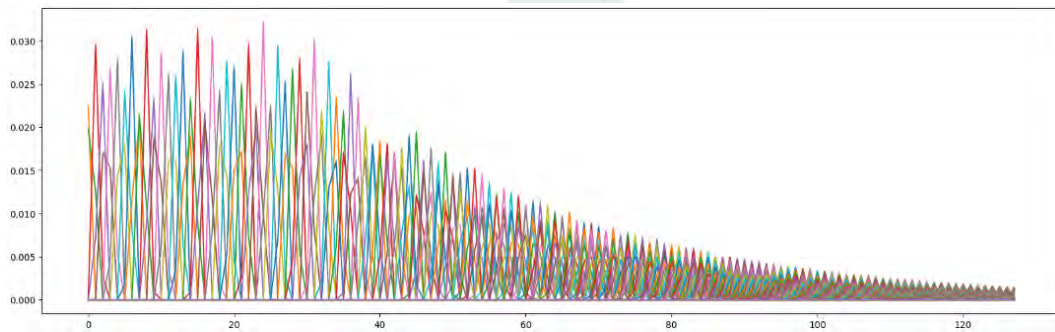
Tahap FFT bertujuan untuk menukar bentuk *spatial domain* atau domain waktu dalam data audio ke dalam bentuk *frequency domain*. Domain waktu audio telah ditampilkan dalam tampilan gelombang sinyal suara, sementara proses ekstraksi fitur diperlukan domain frekuensi. Dimensi antara domain waktu dan domain frekuensi digambarkan pada Gambar 3.3. Pada fitur MFCC *library Librosa*, tahap FFT dijalankan secara otomatis sebelum proses *mel frequency wrapping*



Gambar 3.3 Domain Waktu dan Frekuensi (Mastriani, 2018)

E. *Mel Frequency Wrapping (Filterbank)*

Mel Frequency Wrapping mengaplikasikan sebuah *filter* pada sampel audio untuk memperoleh ukuran energi dari setiap *band* frekuensi. *Filter* ini melakukan metode penskalaan untuk mengubah frekuensi ke dalam skala *mel*. *Mel Filterbank* menggunakan *filter* berbentuk segitiga yang digambarkan pada Gambar 3.4. *Filter* yang digunakan merupakan bentuk *default* dari fitur *librosa.filters.mel* pada *library Librosa*.



Gambar 3.4 *Mel Filterbank*

F. *Discrete Cosine Transform (DCT)*

Output dari tahap sebelumnya yakni *Mel Frequency Spectrogram* menjadi input pada proses ini. Tahap DCT dilakukan untuk mendapatkan indeks MFCCs dengan mengubah *log mel spectrum* ke dalam bentuk domain waktu. DCT mengekstrak perubahan frekuensi rendah dan tinggi dalam sinyal. *Output* dari pemrosesan DCT ini yang kemudian disebut sebagai MFCCs. Tahap DCT dilakukan menggunakan fitur *librosa.feature.mfcc* dengan pengaturan parameter jumlah MFCCs yang dihasilkan sebanyak 20 dan persamaan DCT tipe dua.

3.2.6. Pemisahan Data

Data hasil ekstraksi fitur MFCC kemudian dipisahkan (*data splitting*) menjadi *data train* dan *data test*. Dilakukan percobaan beberapa rasio perbandingan antara *data train* dan *data test* antara lain rasio 70:30, 80:20, 90:10. Proses *data splitting* dilakukan menggunakan fungsi *train_test_split* dalam *library Scikit Learn* dengan parameter *test_size* dan *stratify* pada kolom kelas. Parameter *stratify* digunakan untuk mengambil sampel *data test* dengan jumlah yang sama dari setiap kelas. Kemudian *data training* dibagi lagi ke dalam lima

fold untuk keperluan *5-fold cross validation* dengan menggunakan fungsi *StratifiedKfold*.

3.2.7. Klasifikasi MLP

Proses klasifikasi menggunakan algoritma ANN dilakukan pada nilai koefisien hasil ekstraksi fitur MFCC. ANN diimplementasikan menggunakan arsitektur MLP dengan pelatihan *Backpropagation*. Pendekatan klasifikasi menggunakan *multiclass classification* dengan lima kelas dimana apabila sebuah sampel data suara irama azan telah diklasifikasikan ke dalam satu kelas, data tersebut tidak mungkin ada pada kelas irama azan yang lain. Klasifikasi dijalankan menggunakan fungsi *MLPClassifier* dalam modul *Scikit Learn*. Pelatihan ANN dilakukan menggunakan lima *hidden layer* dengan jumlah *neuron* masing-masing *layer* secara berurutan 1024, 512, 256, 128, 64. Beberapa pengaturan parameter berikutnya dijelaskan pada Tabel 3.2 sebagai berikut.

Tabel 3.2 Parameter *MLPClassifier*

Parameter	Value
Jumlah <i>hidden layer</i>	5
Jumlah <i>neuron</i> dalam <i>hidden layer</i>	1024, 512, 256, 128, 64
Jenis aktivasi	ReLU
Besaran <i>batch size</i>	64
Jenis <i>learning rate</i>	<i>Constant</i>
Besaran <i>learning rate</i>	0.001
Jenis pengoptimal	adam
Jumlah iterasi maksimal	500

3.2.8. Pengujian dan Evaluasi

Pengujian performa klasifikasi dilakukan untuk mengetahui tingkat akurasi model dalam melakukan pengenalan terhadap data suara irama azan. Sebelumnya dilakukan *5-fold cross validation* pada data *train* untuk mengetahui performa pelatihan model. Saat *cross validation* berjalan dilakukan pengukuran akurasi pelatihan pada setiap *fold*, hasil setiap *fold* tersebut dijumlahkan untuk

diambil nilai akurasi rata-rata pelatihan. Pengujian model dilakukan dengan menghitung nilai *recall*, *precision*, *f-measure* dan akurasi menggunakan hasil yang didapat dari tabel *confusion matrix*. *Confusion matrix* dijalankan menggunakan fungsi *metrics confusion matrix* dalam *library Scikit Learn*.

3.2.9. Perbandingan Metode Klasifikasi

Setelah didapatkan hasil klasifikasi irama azan menggunakan MLP, dilakukan klasifikasi pada data hasil ekstraksi fitur MFCC yang sama namun menggunakan metode klasifikasi lain. Klasifikasi dengan metode lain dilakukan sebagai bentuk percobaan dan bahan evaluasi terhadap klasifikasi MLP yang menjadi metode klasifikasi utama dalam penelitian ini. Metode klasifikasi yang menjadi pembanding meliputi metode *Support Vector Machine*, *Random Forest*, dan *XGBoost*. Metode *Support Vector Machine* dan *Random Forest* dijalankan menggunakan fungsi *SVC* dan *RandomForestClassifier* dari *library Scikit Learn*, sedangkan metode *XGBoost* dijalankan menggunakan fungsi *XGBClassifier* dari *library XGBoost*.

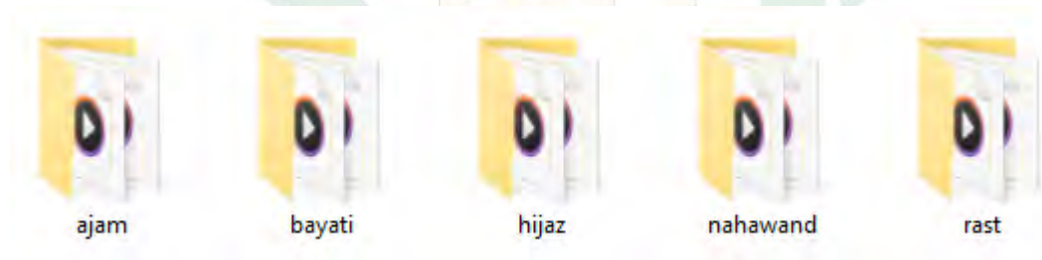
UIN SUNAN AMPEL
S U R A B A Y A

BAB IV

HASIL DAN PEMBAHASAN

4.1. Representasi Data Suara

Sampel data audio dari situs internet dan perekaman suara azan oleh enam muazin yang telah diberi label dilakukan proses ekstraksi audio. Ekstraksi audio dilakukan guna mengambil sampel data pada lafaz *hayya alash shalah*. File audio hasil ekstraksi memiliki durasi 20 detik yang kemudian disimpan dengan frekuensi sampel 44100 Hz, *channel stereo* dan dalam bentuk ekstensi WAV. Seluruh sampel data audio irama azan dengan jumlah 250 file (50 data audio di setiap jenis irama) dikelompokkan ke dalam lima folder sesuai dengan kelasnya seperti pada Gambar 4.1. Kelas data dan folder penyimpanan dinamakan sesuai dengan nama jenis irama azan yang digunakan yakni *rast*, *bayati*, *ajam*, *nahawand* dan *hijaz*.



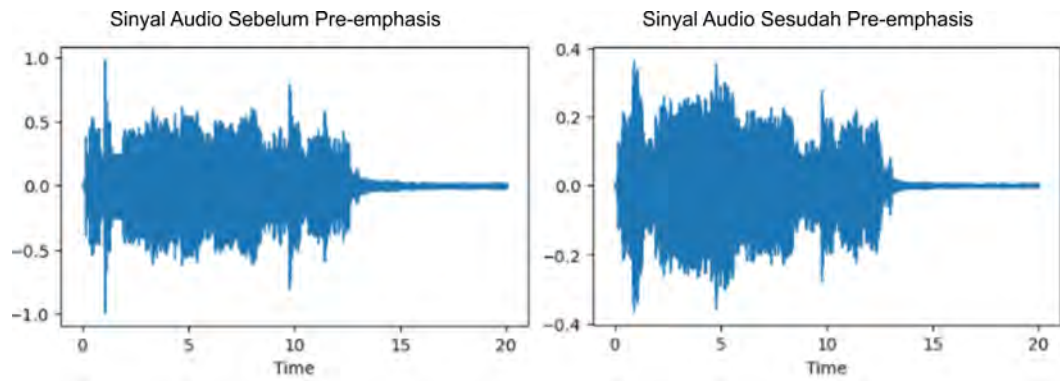
Gambar 4.1 Folder Data Irama Azan

4.2. Ekstraksi Fitur MFCC

Sebelum tahapan ekstraksi fitur MFCC dilakukan, file audio dimuat dalam bentuk *floating point time series* menggunakan fitur *librosa.load* dalam library *Librosa*. Diperlukan dua variabel untuk menyimpan *return* dari fitur *librosa.load*. Dua variabel tersebut adalah “y” yang menyimpan data audio dalam bentuk *time series* dan “sr” yang menyimpan *sampling rate* dari file audio yang dimuat. Kode program untuk menjalankan fitur *librosa.load* tercantum dalam Tabel 4.1 berikut.

Tabel 4.1 Kode Memuat File Audio

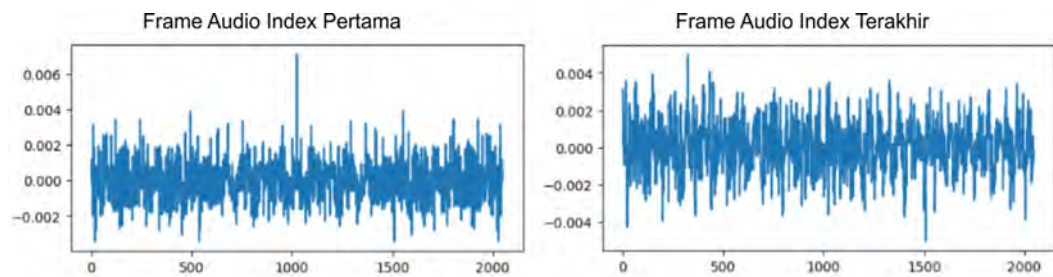
```
y, sr = librosa.load(filename)
librosa.display.waveshow(y, sr=sr)
```

Gambar 4.3 *Pre-emphasis*

4.2.2. Frame Blocking

Dalam tahap *frame blocking* diperlukan parameter jumlah FFT, perhitungan nilai *sample rate* dan *hop length*. Parameter jumlah FFT digunakan untuk menentukan jumlah *frame* dari sampel audio, sedangkan dua parameter lainnya diperlukan untuk menentukan panjang dari setiap *frame*. Dalam fitur MFCC pada *library Librosa*, tahap *frame blocking* secara otomatis dilakukan setelah menentukan besaran parameter yang telah disebutkan. Besaran parameter yang telah ditentukan adalah jumlah FFT sebanyak 2048 dan *hop length* sebesar 512. Gambar 4.4 menunjukkan contoh bentuk *frame* dari sampel file “hijaz_1.wav”.

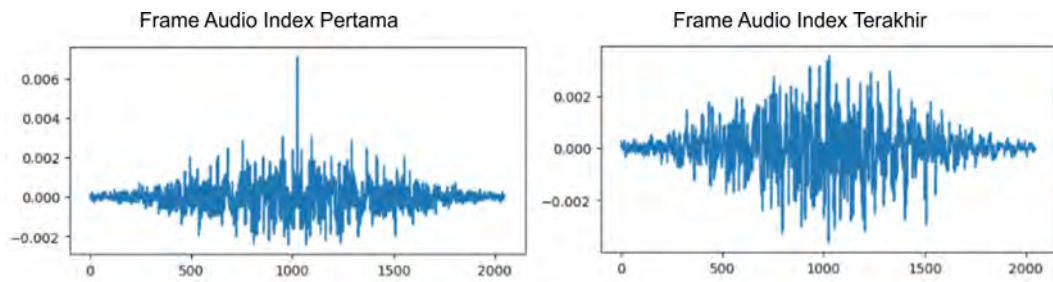


Gambar 4.4 *Frame Blocking*

4.2.3. Windowing

Sama seperti tahap *frame blocking*, tahap *windowing* berjalan secara otomatis setelah menentukan parameter jenis *window* yang digunakan. Jenis *window* yang digunakan adalah *hamming window*, sesuai dengan yang telah ditentukan sebelumnya. Tahap *windowing* mengurangi besaran nilai amplitudo

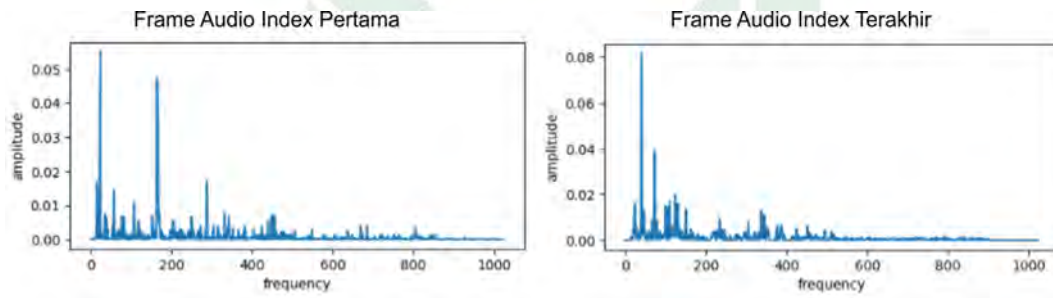
pada awal dan akhir dari setiap *frame* sampel file audio seperti yang digambarkan pada Gambar 4.5 berikut.



Gambar 4.5 *Frame* Setelah *Windowing*

4.2.4. Fast Fourier Transform (FFT)

Pada fitur MFCC *library Librosa*, tahap FFT dijalankan secara otomatis sebelum proses *mel frequency wrapping*. Hasil penerapan FFT pada sampel file audio digambarkan pada Gambar 4.6. Pada Gambar 4.6 sumbu x yang sebelumnya merupakan *domain* waktu diubah menjadi *domain* frekuensi setelah melalui tahap FFT.



Gambar 4.6 *Fast Fourier Transform*

4.2.5. Mel Frequency Wrapping (Filterbank)

Pada tahap *Mel Frequency Wrapping*, *Mel Filterbank* diterapkan pada sampel data audio hasil *FFT*. Setelah proses filtrasi, sampel data audio dalam bentuk skala *mel* dapat divisualisasikan menggunakan grafik spektrogram. Grafik spektrogram menunjukkan intensitas suara yang terjadi dalam frekuensi dan waktu tertentu. Gambar 4.7 menggambarkan hasil *Mel Filterbank* pada sampel file audio “hijaz_1.wav”.

digunakan sebagai nama kelas irama azan. Tabel 4.5 berikut memuat kode yang dijalankan untuk melakukan perintah perulangan.

Tabel 4.5 Kode Perintah Perulangan

```

audio_dataset_path = "dataset/"

extracted_features = []
for path in os.listdir(audio_dataset_path):
    for file in os.listdir(audio_dataset_path+path+"/"):
        if file.lower().endswith(".wav"):
            file_name = audio_dataset_path+path+"/"+file
            data = features_extractor(file_name)
            extracted_features.append([data, path])
    
```

Setelah semua data audio irama azan melalui proses ekstraksi fitur MFCC, hasil ekstraksi fitur yang diambil berupa 20 MFCCs indeks numerik. MFCCs kemudian disimpan dalam bentuk tabel menggunakan fungsi Pandas Dataframe, beberapa sampel data dalam tabel ditampilkan dalam Tabel 4.6.

Tabel 4.6 Hasil Ekstraksi Fitur MFCC

	feature	class
0	[-261.3098, -52.891712, -51.763306, -29.179995, -34.430443, -8.912416, -6.053541, -27.304121, 19.078766, -2.471533, -31.498371, -3.3923218, -3.1677985, -19.15501, 2.8616178, -5.9269433, 17.509476, 31.611725, 3.10002, 19.41012]	ajam
1	[-355.3325, 10.84215, -70.06596, -17.452536, -15.382382, -3.8258832, -13.067917, 21.799767, -16.524628, 19.916185, -29.346987, -9.3166485, 8.650343, -11.731097, -4.908741, 0.3091635, -2.0606182, 18.304731, 12.23881, 11.846486]	ajam
2	[-541.45667, -31.138412, -31.259012, 3.5273528, -5.5801225, -7.5414796, -0.19813256, -10.874468, 2.1270237, -7.09991, 3.0849926, 9.901076, 11.636621, 4.69884, -0.038132615, -7.220722, -7.437305, -2.8224223, 0.56086826, 4.7284636]	ajam
...
247	[-249.37192, 29.839174, -132.6499, -7.55019, -3.6860461, 4.884891, -7.124355, 22.551266, -10.360575, -20.696825, -12.629142, 12.415889, 2.5807328, 11.309815, 5.650108, -1.5864773, -13.115022, -3.447499, 8.622827, -2.0617862]	rast
248	[-342.84436, -13.319213, -51.474636, -5.6389356, -16.079361, 4.5329876, -4.2205276, -9.696367, 1.4867499, -13.668479, 2.9477475, 11.984794, -0.5685673, 3.488706, -18.33657, -8.496944, -4.599653, 1.9151404, -0.085516475, 5.172726]	rast

Tabel 4.9 Kode Klasifikasi MLP

```
clf = MLPClassifier(hidden_layer_sizes=(1024,512,256,128,64),
                    activation='relu',
                    batch_size=64,
                    verbose=0,
                    learning_rate='constant',
                    learning_rate_init=0.001,
                    solver='adam',
                    max_iter=500)

clf.fit(X_train, y_train)
y_pred=clf.predict(X_test)
```

4.5. Hasil Klasifikasi dan Evaluasi Model

Setelah proses pelatihan model klasifikasi MLP, berikutnya dilakukan pemrosesan terhadap hasil klasifikasi data audio irama azan. Pertama dilakukan pengujian *5-fold cross validation* pada *data train* model. Kemudian *confusion matrix* diadakan untuk mengetahui kelas yang diprediksi oleh model. Dari jumlah kelas positif dan negatif dalam *confusion matrix* dapat dilakukan perhitungan nilai *recall*, *precision*, *f-measure* dan akurasi. Pemrosesan hasil *5-fold cross validation* dan *confusion matrix* dijalankan menggunakan fungsi *cross_val_score*, *confusion_matrix* dan *classification_report* dalam *library Scikit Learn*. Tabel 4.10 memuat kode program yang dijalankan untuk melakukan pemrosesan hasil .

Tabel 4.10 Kode Pemrosesan Hasil

```
score = cross_val_score(clf, X_train, y_train, cv=kf,
                        scoring="accuracy")

print(classification_report(y_test, y_pred))

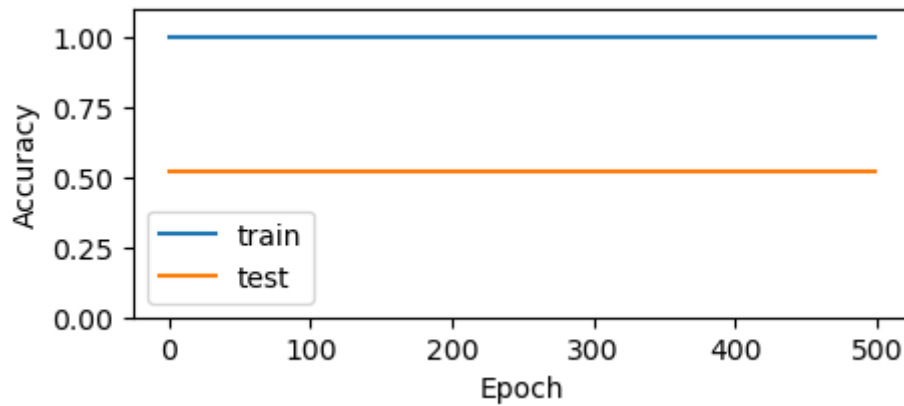
cf_matrix = confusion_matrix(y_test, y_pred)

fig, ax = plt.subplots(figsize=(4,4))
sns.heatmap(cf_matrix.astype('float')/cf_matrix.sum(axis=1),
            annot=True, fmt='.2f', cmap='Blues', cbar=False)

ax.set_ylabel('Actual Class');

ax.xaxis.set_ticklabels(['ajam', 'bayati', 'hijaz', 'nahawand',
                        'rast'], rotation=45);

ax.yaxis.set_ticklabels(['ajam', 'bayati', 'hijaz', 'nahawand',
                        'rast'], rotation=0);
```

Gambar 4.15 *Learning Curve* 90:10

Ketiga skenario klasifikasi MLP berdasarkan rasio perbandingan *data train* dan *data test* telah dilakukan. Perbedaan rasio perbandingan data menghasilkan perbedaan yang cukup signifikan. Hasil terbaik didapatkan dengan skenario rasio perbandingan *data train* dan *data test* 80:20. Hasil pengujian ketiga skenario ditampilkan pada Tabel 4.14. Dari ketiga skenario didapatkan *plotting* hasil *learning curve* yang menunjukkan nilai akurasi pada saat *train* jauh lebih tinggi dari nilai akurasi *test* secara berturut turut. Hasil *learning curve* tersebut menandakan bahwa model klasifikasi mengalami *overfitting*.

Tabel 4.14 Hasil Skenario Model Klasifikasi

Skenario Rasio Perbandingan		Precision	Recall	F1 Score	Acuration
Data Train	Data Test				
70%	30%	48%	47%	46%	47%
80%	20%	59%	56%	56%	56%
90%	10%	58%	44%	47%	44%

4.6. Perbandingan Performa Klasifikasi

Hasil klasifikasi menggunakan MLP telah didapatkan, di mana hasil performa terbaik didapatkan menggunakan rasio perbandingan *data train* dan *data test* 80:20. Berikutnya dengan menggunakan rasio perbandingan *data train* dan *data test* terbaik, dilakukan klasifikasi dengan data hasil ekstraksi fitur MFCC yang sama namun menggunakan metode klasifikasi lain. Metode klasifikasi yang menjadi pembanding meliputi metode *Support Vector Machine*, *Random Forest*,

secara iteratif seperti MLP. Tingkat akurasi pada saat *train* konsisten pada tingkat akurasi tertinggi, sementara tingkat akurasi *test* terus meningkat seiring dengan semakin banyaknya data yang digunakan. Tingkat akurasi pada saat *test* berada pada tingkat paling tinggi di sekitar 40% dengan ukuran data yang digunakan sebanyak 200. Perbedaan tingkat akurasi *train* dan *test* tersebut menandakan bahwa model klasifikasi terindikasi mengalami *overfitting*.

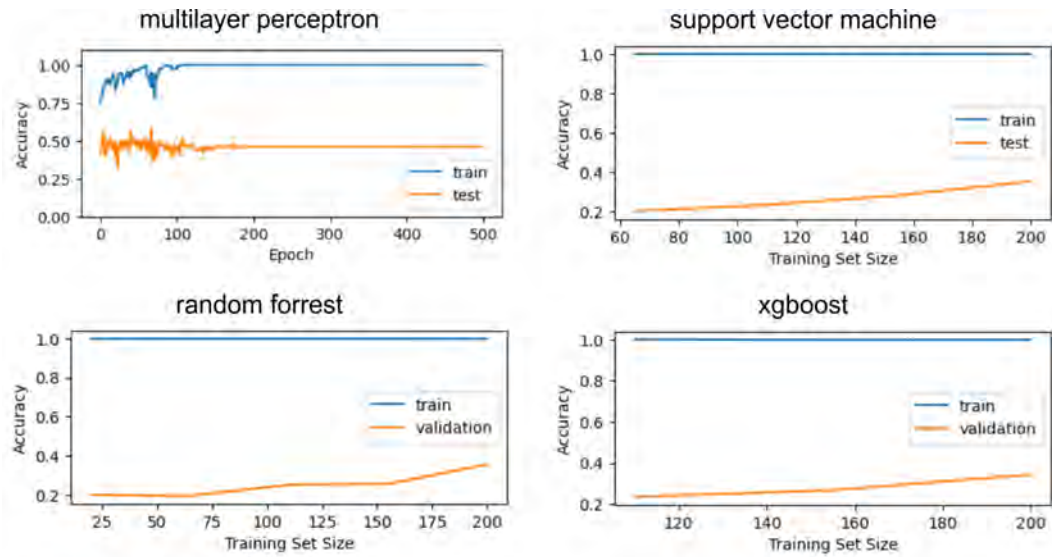
Metode kedua yang dilakukan percobaan perbandingan klasifikasi adalah *Random Forest*. Klasifikasi *Random Forest* dijalankan menggunakan fungsi *RandomForestClassifier* dari *library Scikit Learn*. Pengaturan parameter yang digunakan pada model klasifikasi merupakan pengaturan *default* fungsi *RandomForestClassifier*. Kode program mencakup parameter yang digunakan dalam klasifikasi *Random Forest* tercantum dalam Tabel 4.17 sebagai berikut.

Tabel 4.17 Kode Klasifikasi *Random Forest*

```
rfc = RandomForestClassifier(n_estimators=100,  
                             min_samples_split=2,  
                             min_samples_leaf=1)  
  
rfc.fit(X_train, y_train)  
y_pred = rfc.predict(X_test)
```

Perintah *fit* dijalankan untuk proses *fitting* model atau pelatihan model dan perintah *predict* digunakan untuk melakukan prediksi menggunakan *Random Forest*.

Setelah berhasil menjalankan kode program untuk klasifikasi *Random Forest*, berikutnya dijalankan kode untuk pemrosesan hasil klasifikasi. Didapatkan hasil rata-rata dari *5-fold cross validation* sebesar 45%. Sedangkan prediksi model mendapatkan tingkat akurasi tertinggi sebesar 54% dengan *classification report* dalam Tabel 4.18. Evaluasi model digambarkan dalam tabel *confusion matrix* pada Gambar 4.18.



Gambar 4.22 *Learning Curve* Klasifikasi

Dari seluruh *learning curve* dalam Gambar 4.22 terindikasi semua model mengalami *overfitting* dengan kondisi tingkat akurasi saat *train* jauh lebih tinggi dibandingkan saat *test*. *Overfitting* adalah kondisi dimana pelatihan *machine learning* memiliki kinerja yang baik, namun saat pengujian dilakukan kinerja menjadi kurang baik (Widhiyasana et al., 2021). *Overfitting* dapat terjadi karena keterbatasan *data training*, baik dalam keterbatasan ukuran atau terlalu banyak *noise* data yang disertakan, atau kendala terhadap algoritma yang terlalu rumit dan membutuhkan terlalu banyak parameter (Ying, 2019). Dikarenakan seluruh metode klasifikasi yang digunakan mengalami *overfitting*, maka diindikasikan penyebab *overfitting* berasal dari dataset yang digunakan.

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan dengan hasil yang didapatkan melalui beberapa tahapan yang telah dilakukan dalam penelitian klasifikasi irama azan berbasis *mel frequency cepstral coefficients* menggunakan *multilayer perceptron*, dapat diambil kesimpulan sebagai berikut:

1. Ekstraksi fitur suara MFCC irama azan dilakukan menggunakan *library librosa* dengan melalui beberapa tahapan antara lain *pre-emphasis*, *frame blocking*, *windowing*, *fast fourier transform*, *mel frequency wrapping* dan *discrete cosine transform* dengan hasil ekstraksi fitur berupa 20 koefisien MFCCs.
2. Klasifikasi MLP irama azan dilakukan menggunakan fungsi *MLPClassifier* dalam modul *Scikit Learn* dengan mempertimbangkan beberapa parameter meliputi jumlah *hidden layer*, jumlah *neuron* dalam *hidden layer*, jenis aktivasi, besaran *batch*, jenis *learning rate*, besaran *learning rate*, jenis pengoptimal dan jumlah iterasi maksimal.
3. Dilakukan tiga skenario pengenalan irama azan menggunakan metode MFCC dan MLP berdasarkan rasio perbandingan *data train* dan *data set*. Rasio perbandingan *data train* dan *data set* dengan hasil paling baik adalah rasio 80:20 dengan tingkat akurasi sebesar 56%. Hasil tersebut menunjukkan performa pengenalan irama azan yang kurang baik. Data tingkat akurasi *train* dan *test* dalam *learning curve* mengindikasikan terjadinya *overfitting* pada model yang digunakan.
4. Dilakukan klasifikasi dengan tiga metode klasifikasi lain meliputi metode *Support Vector Machine*, *Random Forest*, dan *XGBoost* sebagai bahan evaluasi. Metode klasifikasi dengan performa paling baik dihasilkan metode *Random Forest* dengan persentase akurasi tertinggi sebesar 54%. *Learning curve* dari semua metode klasifikasi yang dilakukan menunjukkan terjadinya *overfitting*. Diindikasikan penyebab *overfitting* berasal dari dataset yang digunakan.

5.2. Saran

Berdasarkan penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dipertimbangkan dalam mengembangkan penelitian ini. Beberapa saran untuk mengembangkan penelitian berdasarkan hasil temuan yang didapat adalah sebagai berikut:

1. Untuk meningkatkan generalisasi, disarankan menambahkan jumlah sampel data suara irama azan. Semakin besar ukuran dataset umumnya mengarah kepada peningkatan performa. Tidak menutup kemungkinan untuk menambah jenis irama azan yang lain.
2. Menerapkan ekstraksi fitur tambahan seperti *zero crossing*, *chroma* dan *pitch tracking* untuk meningkatkan pengenalan karakteristik dari setiap irama azan. Semakin dikenali karakteristik dari setiap irama azan dapat berimbas kepada peningkatan performa.
3. Melakukan pengembangan model dengan mencoba beberapa kombinasi parameter dalam *MLPClassifier*, atau dapat menggunakan metode klasifikasi terbaru lain pada penelitian kedepannya.

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR PUSTAKA

- Adriana, A. D. (2013). Perangkat Lunak Untuk Membuka Aplikasi Pada Komputer Dengan Perintah Suara Menggunakan Metode Mel Frequency Cepstrum Coefficients. *Komputa: Jurnal Ilmiah Komputer Dan Informatika*, 2(1).
- Ahmad, A. (2017). Mengenal artificial intelligence, machine learning, neural network, dan deep learning. *J. Teknol. Indonesia*, 3.
- Al Bakri, T., Mallah, Mohammed, & Nuserat, N. (2019). *Al Adhan: Documenting historical background, practice rules, and musicological features of the muslim call for prayer in Hashemite Kingdom of Jordan*. <https://digilib.phil.muni.cz/cs/handle/11222.digilib/141569>
- Al Hushari, M. (2002). *Akham Qira'at al Quran al Karim*. Maktabah as sunnah.
- Anishchenkova, V. (2020). *Modern Saudi Arabia (Understanding Modern Nations)*. ABC-CLIO.
- Ash-Shiddieqy, T. M. H. (1994). *Pedoman Shalat*. PT. Bulan Bintang.
- Basuki, D. K. (2009). Konversi Nada-nada Akustik Menjadi Chord Menggunakan Pitch Class Profile. *Industrial Electronic Seminar*. <http://repo.pens.ac.id/175/>
- Chamidy, T. (2016). Metode Mel Frequency Cepstral Coeffisients (MFCC) Pada klasifikasi Hidden Markov Model (HMM) Untuk Kata Arabic pada Penutur Indonesia. *MATICS: Jurnal Ilmu Komputer Dan Teknologi Informasi (Journal of Computer Science and Information Technology)*, 8(1), Article 1. <https://doi.org/10.18860/mat.v8i1.3482>
- da Silva, I. N., Hernane Spatti, D., Andrade Flauzino, R., Liboni, L. H. B., & dos Reis Alves, S. F. (2017). Artificial Neural Network Architectures and Training Processes. In I. N. da Silva, D. Hernane Spatti, R. Andrade Flauzino, L. H. B. Liboni, & S. F. dos Reis Alves (Eds.), *Artificial Neural Networks: A Practical Course* (pp. 21–28). Springer International Publishing. https://doi.org/10.1007/978-3-319-43162-8_2
- Deng, X., Liu, Q., Deng, Y., & Mahadevan, S. (2016). An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Information Sciences*, 340–341, 250–261.

<https://doi.org/10.1016/j.ins.2016.01.033>

Dwisaputra, I., Silalahi, P., Cahyawan, B., & Akbar, I. (2019). Lampu Sein Helm Sepeda Berbasis Voice Recognition. *MANUTECH: Jurnal Teknologi Manufaktur*.

<https://ejournal.polman-babel.ac.id/index.php/manutech/article/view/96>

El Naqa, I., & Murphy, M. J. (2015). What Is Machine Learning? In I. El Naqa, R. Li, & M. J. Murphy (Eds.), *Machine Learning in Radiation Oncology: Theory and Applications* (pp. 3–11). Springer International Publishing. https://doi.org/10.1007/978-3-319-18305-3_1

Faiz, M. N., Somantri, O., Supriyono, A. R., & Muhammad, A. W. (2022). Impact of Feature Selection Methods on Machine Learning-based for Detecting DDoS Attacks: Literature Review. *JOURNAL OF INFORMATICS AND TELECOMMUNICATION ENGINEERING*, 5(2), Article 2. <https://doi.org/10.31289/jite.v5i2.6112>

Firdousi, R., & Parveen, S. (2014). Local thresholding techniques in image binarization. *International Journal Of Engineering And Computer Science*, 3(4), 4062–4065.

Gupta, S., Jaafar, J., wan Ahmad, W. F., & Bansal, A. (2013). Feature Extraction Using Mfcc. *Signal & Image Processing : An International Journal*, 4(4), 101–108. <https://doi.org/10.5121/sipij.2013.4408>

Hady, M. F. A., & Schwenker, F. (2013). Semi-supervised Learning. In M. Bianchini, M. Maggini, & L. C. Jain (Eds.), *Handbook on Neural Information Processing* (pp. 215–239). Springer. https://doi.org/10.1007/978-3-642-36657-4_7

Hasan, M. (2021). *Korelasi Pemilihan Lagu Bacaan Al-qur'an Dengan Makna Al-qur'an*. Cipta Media Nusantara.

Heriyanto, H., Hartati, S., & Putra, A. E. (2018). Ekstraksi Ciri Mel Frequency Cepstral Coefficient (MFCC) dan Rerata Coefficient Untuk Pengecekan Bacaan Al-qur'an. *Telematika : Jurnal Informatika dan Teknologi Informasi*, 15(2), Article 2. <https://doi.org/10.31315/telematika.v15i2.3123>

Janse, P., Magre, S., Kurzekar, P., & Deshmukh, R. (2014). *A Comparative Study between MFCC and DWT Feature Extraction Technique*.

- Jian, L., Huang, Z., Zhang, J., & Hu, Z. (2022). Rapid Analysis of Cylindrical Bypass Flow Field Based on Deep Learning Model. *IOP Conference Series: Earth and Environmental Science*, 1037, 012013. <https://doi.org/10.1088/1755-1315/1037/1/012013>
- Joshi, S., & Cheeran, A. N. (2014). MATLAB Based Feature Extraction Using MFCC for ASR. *International Journal of Science, Engineering and Technology Research (IJSETR)*, 3(14).
- Jumarwanto, A., Hartanto, R., & Prastiyanto, D. (2009). Aplikasi Jaringan Saraf Tiruan Backpropagation Untuk Memprediksi Penyakit THT di Rumah Sakit Mardi Rahayu Kudus. *Jurnal Teknik Elektro*, 1(1), Article 1. <https://doi.org/10.15294/jte.v1i1.1601>
- Khamlich, S., Khamlich, F., Atouf, I., & Benrabh, M. (2021). Performance evaluation and implementations of MFCC, SVM and MLP algorithms in the FPGA board. *International Journal of Electrical and Computer Engineering Systems*, 12(3), 139–153. <https://doi.org/10.32985/ijeces.12.3.3>
- Kolcz, A., Prabakarmurthi, V., & Kalita, J. (2001). Summarization as feature selection for text categorization. *Proceedings of the Tenth International Conference on Information and Knowledge Management*, 365–370. <https://doi.org/10.1145/502585.502647>
- Kwon, O., Kim, H. G., Ham, M. J., Kim, W., Kim, G.-H., Cho, J.-H., Kim, N. I., & Kim, K. (2020). A Deep Neural Network for Classification of Melt-pool Images in Metal Additive Manufacturing. *Journal of Intelligent Manufacturing*, 31(2), 375–386. <https://doi.org/10.1007/s10845-018-1451-6>
- López-Ruiz, R. (2018). *From Natural to Artificial Intelligence: Algorithms and Applications*. BoD – Books on Demand.
- Marwala, T. (2012). *Condition Monitoring Using Computational Intelligence Methods: Applications in Mechanical and Electrical Systems*. Springer. <https://doi.org/10.1007/978-1-4471-2380-4>
- Masrurin, ‘Ainatu. (2018). Murattal dan Mujawwad Al-qur’an di Media Sosial. *Jurnal Studi Ilmu-Ilmu Al-Qur’an Dan Hadis*, 19(2), Article 2.

<https://doi.org/10.14421/qh.2018.1902-04>

- Mastriani, M. (2018). Quantum-Classical Algorithm for an Instantaneous Spectral Analysis of Signals: A Complement to Fourier Theory. *Journal of Quantum Information Science*, 08, 52–77. <https://doi.org/10.4236/jqis.2018.82005>
- Maulana, P. I., Aranta, A., Bimantoro, F., & Andika, I. G. (2022). Klasifikasi Mood Musik Berdasarkan Mel Frequency Cepstral Coefficients Dengan Backpropagation Neural Network. *Jurnal RESISTOR (Rekayasa Sistem Komputer)*, 4(1).
- Misbachul Munir, M. (1997). *Pedoman Lagu-Lagu Tilawatil Qur'an Dilengkapi dengan Tajwid dan Qosidah*. Apollo.
- Muda, L., Begam, M., & Elamvazuthi, I. (2010). *Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques*. 2(3).
- Nasution, T. (2012). Metoda Mel Frequency Cepstrum Coefficients (MFCC) untuk Mengenali Ucapan pada Bahasa Indonesia. *Jurnal Sains dan Teknologi Informasi*, 1(1), 22–31.
- Peryanto, A., Yudhana, A., & Umar, R. (2020). Klasifikasi Citra Menggunakan Convolutional Neural Network dan K Fold Cross Validation. *Journal of Applied Informatics and Computing*, 4(1), Article 1. <https://doi.org/10.30871/jaic.v4i1.2017>
- Prasetyo, E. (2012). *Data Mining Konsep dan Aplikasi Menggunakan Matlab*. Andi.
- Proakis, J. G. (2007). *Digital signal processing: Principles, algorithms, and applications*. Pearson Education India.
- Putra, J. W. G. (2020). *Pengenalan Konsep Pembelajaran Mesin dan Deep Learning (1.4)*. Self Published.
- Qasim, R. (1989). Some Islamic non-Arabic Elements of Influence on The Repertory of al Maqam al Iraqi in Bagdad. *Maqam and Raga and Zeilenmelodik, International Council for Traditional Music*, 28/2.
- Quteishat, A., Younis, M., Qtaishat, A., & Abuhamdah, A. (2023). Intelligent Arabic letters speech recognition system based on mel frequency cepstral

- coefficients. *International Journal of Electrical and Computer Engineering (IJECE)*, 13(3), 3348.
<https://doi.org/10.11591/ijece.v13i3.pp3348-3358>
- Rafiqo, D., Suyanto, Y., & Armaji, C. (2022). Klasifikasi Suara Paru-Paru Berdasarkan Ciri MFCC. *Indonesian Journal of Electronics and Instrumentation Systems (IJEIS)*, 12(1).
- Raya, A. T., & Mulia, S. M. (2003). *Menyelami Seluk Beluk Ibadah dalam Islam*. Kencana.
- Riyanto, E., & Sutejo, -. (2014). Perbandingan Metode Ekstraksi Ciri Suara MFCC, ZCPA, dan LPC. *HIMSYATECH*, 10(1), Article 1.
<https://ojs.stmik-himsya.ac.id/index.php/JHT/article/view/61>
- Rizal, R., Rosnita, L., & Ikramina, I. (2018). Sistem Pengenalan Nagham Adzan Melalui Suara Menggunakan Metode Discrete Wavelet Transform (Dwt) dan Mellin Transform. *TECHSI - Jurnal Teknik Informatika*, 10(2), Article 2. <https://doi.org/10.29103/techsi.v10i2.908>
- Rusdi, M., & Yani, A. (2018). Sistem Kendali Peralatan Elektronik Melalui Media Bluetooth Menggunakan Voice Recognition. *JET (Journal of Electrical Technology)*, 3(1), Article 1.
- Sahidullah, Md., & Saha, G. (2012). Design, analysis and experimental evaluation of block based transformation in MFCC computation for speaker recognition. *Speech Communication*, 54(4), 543–565.
<https://doi.org/10.1016/j.specom.2011.11.004>
- Salim, M. (2004). *Ilmu Nagham Al-Qur'an: Belajar Membaca Al-Qur'an dengan Lagu (Metode SBA Teotik)* (2nd ed.). PT. Kebayoran Widya Ripta.
- Samuel, A. L. (2000). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 44(1.2), 206–226.
<https://doi.org/10.1147/rd.441.0206>
- Sen, S., Sugiarto, D., & Rochman, A. (2020). Komparasi Metode Multilayer Perceptron (MLP) dan Long Short Term Memory (LSTM) dalam Peramalan Harga Beras. *Ultimatics : Jurnal Teknik Informatika*, 12(1), 35–41. <https://doi.org/10.31937/ti.v12i1.1572>
- Setiawan, S. A., & Ariyanto, Y. (2015). Voice Recognition: Pengenalan Chord

- Ukulele 4 Senar dengan Menggunakan Metode Back Propagation Neural Network. *Seminar Informatika Aplikatif Polinema*.
<http://jurnalti.polinema.ac.id/index.php/SIAP/article/view/132>
- Shahriar, S., & Tariq, U. (2021). Classifying Maqams of Qur'anic Recitations Using Deep Learning. *IEEE Access*, 9, 117271–117281.
<https://doi.org/10.1109/ACCESS.2021.3098415>
- Shokouhi, M. A., & Yusof, A. (2013). *The influence of Islamic culture and holy Quran on performing arts: Relating to sacred vocal music (LAHN)*. 363–382.
- Siang, J. J. (2005). *Jaringan syaraf tiruan dan pemrogramannya menggunakan Matlab*. Penerbit ANDI.
- Siregar, A. C., & Octariadi, B. C. (2021). Perbandingan Metode Jaringan Syaraf Tiruan Pada Klasifikasi Motif Kain Tenun Sambas. *CYBERNETICS*, 4(02), Article 02. <https://doi.org/10.29406/cbn.v4i02.2489>
- Sydenham, P. H., & Thorn, R. (Eds.). (2005). *Handbook of measuring system design*. Wiley.
- Tanprasert, C., & Achariyakulporn, V. (2000). Comparative study of GMM, DTW, and ANN on Thai speaker identification system. *6th International Conference on Spoken Language Processing (ICSLP 2000)*, vols. 2, 234-237–0. <https://doi.org/10.21437/ICSLP.2000-252>
- Tirumala, S. S., Shahamiri, S. R., Garhwal, A. S., & Wang, R. (2017). Speaker identification features extraction methods: A systematic review. *Expert Systems with Applications*, 90, 250–271.
<https://doi.org/10.1016/j.eswa.2017.08.015>
- Wan Abdullah, W. H., & Mat Razali, A. (2015). Modular Construction Shouts AdhanQari Based Egyptian Quran Reciter Mosque Management System. *Journal of Islamic Studies and Culture*, 3.
<https://doi.org/10.15640/jisc.v3n2a11>
- Widhiyasana, Y., Transmissia Semiawan, Ilham Gibran Achmad Mudzakir, & Muhammad Randi Noor. (2021). Penerapan Convolutional Long Short-Term Memory untuk Klasifikasi Teks Berita Bahasa Indonesia. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, 10(4), 354–361.

<https://doi.org/10.22146/jnteti.v10i4.2438>

- Wulandari Siagian, T. N., Nuha, H. H., & Yasirandi, R. (2020). Footstep Recognition Using Mel Frequency Cepstral Coefficients and Artificial Neural Network. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 4(3), 497–503. <https://doi.org/10.29207/resti.v4i3.1964>
- Yadav, S., & Shukla, S. (2016). Analysis of k-Fold Cross-Validation over Hold-Out Validation on Colossal Datasets for Quality Classification. *2016 IEEE 6th International Conference on Advanced Computing (IACC)*, 78–83. <https://doi.org/10.1109/IACC.2016.25>
- Yegnanarayana, B. (2009). *Artificial Neural Networks*. PHI Learning Pvt. Ltd.
- Ying, X. (2019). An Overview of Overfitting and its Solutions. *Journal of Physics: Conference Series*, 1168(2), 022022. <https://doi.org/10.1088/1742-6596/1168/2/022022>
- Yu, Z., Wang, K., Wan, Z., Xie, S., & Lv, Z. (2022). Popular deep learning algorithms for disease prediction: A review. *Cluster Computing*. <https://doi.org/10.1007/s10586-022-03707-y>

UIN SUNAN AMPEL
S U R A B A Y A