

SKRIPSI

**PENERAPAN ALGORITMA *K-NEAREST NEIGHBOR* (KNN) DALAM
KLASIFIKASI SUARA BERDASARKAN BAHASA DAERAH TIDORE**



OLEH
Muh. Irfan Mansyur
07352011055

PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS KHAIRUN
TERNATE
2024

LEMBAR PENGESAHAN

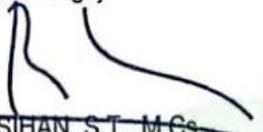
PENERAPAN ALGORITMA *K-NEAREST NEIGHBOR* DALAM KLASIFIKASI SUARA BERDASARKAN BAHASA DAERAH TIDORE

Oleh
Muh. Irfan Mansyur
07352011055

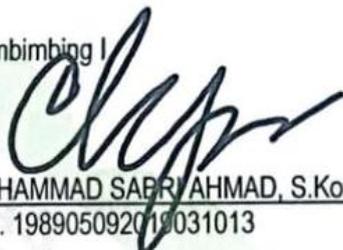
Skripsi ini telah disahkan
Tanggal 29 Juli 2024

Menyetujui
Tim Penguji

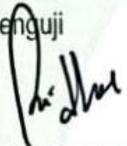
Ketua Penguji


ROSIHAN, S.T., M.Cs.
NIP. 197607192010121001

Pembimbing I


MUHAMMAD SABRI AHMAD, S.Kom., M.Kom.
NIP. 198905092019031013

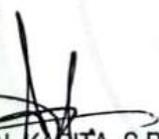
Anggota Penguji


Dr. MUHAMMAD RIDHA ALBAAR, S.Kom., M.Kom.
NIP. 198504232008031001

Pembimbing II


MUHAMMAD FHADLI, S.Kom., M.Kom.
NIP. 199611232023211012

Anggota Penguji


SYARIFUDDIN N. KAPITA, S.Pd., M.Si.
NIP. 199103122024211001

Mengetahui/Menyetujui

Koordinator Program Studi
Informatika


ROSIHAN, S.T., M.Cs.
NIP. 197607192010121001

Dekan Fakultas Teknik
Universitas Khairun


U. ENDANG HARISUN, S.T., M.T., CRP.
NIP. 197511302005011013



LEMBAR PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Muh. Irfan Mansyur
NPM : 07352011055
Fakultas : Teknik
Jurusan/Program Studi : Informatika
Judul Skripsi : Penerapan Algoritma K-Nearest Neighbor Dalam Klasifikasi Suara Berdasarkan Bahasa Daerah Tidore

Dengan ini menyatakan bahwa penulisan Skripsi yang telah saya buat ini merupakan hasil karya sendiri dan benar keasliannya. Apabila ternyata di kemudian hari penulisan Skripsi ini merupakan hasil plagiat atau penjiplakan terhadap karya orang lain, maka saya bersedia mempertanggung jawabkan sekaligus bersedia menerima sanksi berdasarkan aturan tata tertib di Universitas Khairun.
Demikian pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan.

Penulis



Muh. Irfan Mansyur

HALAMAN PERSEMBAHAN

Tiada lembar skripsi yang paling indah dalam laporan skripsi ini kecuali lembar persembahan, Bismillahirrahmannirrahim, dengan rasa bangga penulis persembahkan skripsi ini kepada:

Allah *Subhanaahu Wata'ala* yang telah memberikan kemudahan dan pertolongan sehingga penulis dapat menyelesaikan skripsi ini dengan baik.

Teristimewa kedua orang tua saya tercinta Bapak Mansyur Rasyid dan Ibu Dana Sabate yang sangat saya cintai dan banggakan karena tiada hentinya melangitkan doa-doa tulus, terima kasih juga karena selalu memberikan dukungan, nasehat dan pengorbanan yang dilakukan sehingga menjadi motivasi untuk saya dalam menyelesaikan skripsi ini. Terima kasih sudah mengantarkan saya sampai ditempat ini, saya persembahkan karya tulis sederhana ini dan gelar untuk mama dan papa.

Kepada kedua orang adikku, Alfira Damayanti dan Muh. Ilham. Terima kasih sudah ikut serta dalam proses penulis menempuh Pendidikan selama ini, terima kasih atas semangat, doa dan cinta yang selalu diberikan kepada penulis

Terima kasih kepada diri saya sendiri atas segala kerja keras dan semangatnya sehingga tidak pernah menyerah dalam mengerjakan skripsi ini. Terima kasih kepada diri saya sendiri yang sudah kuat melewati lika liku kehidupan hingga sekarang dan masih tetap tegar dan ikhlas menjalani semuanya, Saya bangga pada diri saya sendiri! Kedepannya mari bekerja sama untuk lebih berkembang lagi menjadi pribadi yang lebih baik.

MOTTO

“Orang lain gak akan bisa paham struggle dan masa sulit nya kita yang mereka ingin tahu hanya bagian success stories, Berjuanglah utuk diri sendiri walaupun tidak ada yang tepuk tangan. Kelak diri kita dimasa depan akan sangat bangga dengan apa yang kita perjuangkan hari ini, tetap berjuang ya!”

KATA PENGANTAR

Puji syukur saya panjatkan kepada Allah *Subhanahu Wata'ala* yang telah melimpahkan rahmat, taufik dan hidayah-Nya sehingga penulis dapat menyusun skripsi ini dengan judul “Penerapan Algoritma K-Nearest Neighbor (KNN) Dalam Klasifikasi Suara Berdasarkan Bahasa Daerah Tidore”, ini dapat diselesaikan untuk memenuhi salah satu persyaratan penyelesaian pendidikan sarjana Informatika Strata Satu (S1) pada Program Studi Informatika Fakultas Teknik Universitas Unkhair.

Untuk menyelesaikan skripsi ini penulis sepenuhnya mendapat dukungan dari banyak pihak, oleh karena itu dengan rendah hati penulis mengucapkan terimakasih kepada:

1. Bapak Dr. Ridha Ajam, M.Hum., selaku Rektor Universitas Khairun Ternate.
2. Bapak Ir. Endah Harisun, S.T., M.T., CRP., selaku dekan Fakultas Teknik Universitas Khairun.
3. Bapak Rosihan, S.T., M.Cs., selaku Koordinator Program Studi Informatika Fakultas Teknik Universitas Khairun dan sekaligus sebagai Penguji I, terima kasih atas bimbingannya, serta dukungan dalam penyelesaian skripsi ini.
4. Bapak Muhammad Sabri Ahmad, S.Kom., M.Kom., selaku Pembimbing I, yang telah bersedia meluangkan waktunya untuk memberikan bimbingan, dukungan serta saran juga kritik yang sangat berarti bagi penulis dalam proses penyelesaian skripsi penelitian ini.
5. Bapak Muhammad Fhadli, S.Kom., M.Sc., selaku Pembimbing II, yang telah bersedia meluangkan waktunya untuk memberikan bimbingan, dukungan serta saran juga kritik yang sangat berarti bagi penulis dalam proses penyelesaian laporan skripsi penelitian ini.
6. Bapak Dr. Muhammad Ridha Albaar. S.Kom., M.Kom., selaku Penguji II yang telah bersedia meluangkan waktunya untuk menguji dan memberikan masukan perbaikan, sehingga penulis dapat menyelesaikan skripsi ini.
7. Bapak Syarifuddin N. Kapita, S.Pd., M.Si., selaku Penguji III yang telah bersedia meluangkan waktunya untuk menguji dan memberikan masukan perbaikan, sehingga penulis dapat menyelesaikan skripsi ini.
8. Bapak Mansyur Rasyid ayah dari penulis terima kasih telah membentuk penulis

menjadi orang yang bisa bertanggung jawab terhadap diri sendiri. Walaupun beliau tidak sempat merasakan Pendidikan sampai bangku perkuliahan, tetapi beliau mampu mendidik penulis dan memberikan dukungan yang penuh sehingga penulis dapat menyelesaikan skripsi ini.

9. Ibu Dana Sabate terima kasih sebesar-besarnya penulis sampaikan kepada beliau pintu surga dari penulis atas dukungan, semangat dan doa yang tiada henti dikirimkan, terima kasih sudah mendidik dan mengingatkan penulis untuk selalu beribadah, beribu-ribu penulis ucapkan karena selalu memberikan kepercayaan dan penguat disaat penulis merasa down, dan selalu memberikan motivasi sehingga penulis mampu menyelesaikan skripsi ini.
10. Teruntuk Kedua adikku Alfira Damayanti dan Muh. Ilham Teriam kasih karena telah memberikan semangat, dorongan serta doa yang baik kepada penulis sehingga penulis bisa sampai pada tahap ini dan menyelesaikan skripsi ini
11. Terima kasih kepada teman-teman seperjuangan Ketar Ketir yaitu: Ubaidinnur Syah, Dani Kurniawan Zakarias, Fauzan Gani Sahupala, Rickal Nurmidin, Nurul Sara Dewi, Cinta Naloni, Fadira Badar, Maulidia Tiradhita Ladette dan Anissa Dwi Sudirno yang telah memberikan saran dan terus membantu penulis, terima kasih karena telah membuat penulis dapat merasakan hal baru selama penulis berada di ternate dan selalu ada disaat susah maupun senang, terima kasih telah mau berkembang Bersama penulis selama 4 tahun yang berharga ini.
12. Terima kasih kepada teman-teman Grup S.Kom yang selalu membantu penulis selama perkuliahan dan memberikan semangat dan memberikan dorongan positif sehingga penulis dapat menyelesaikan skripsi ini.
13. Terima kasih kepada teman-teman pejuang S.Kom yang selalu membantu dan mengarahkan penulis selama proses pembuatan skripsi ini bisa selesai.
14. Dan tak lupa juga ucapan terima kasih dan semangat kepada teman-teman seperjuangan Kosan yaitu: Rivai Maradjabessy, Muhammad. Isra M. Abudu, Asis wally, Fajar Apriyaldi I. Rejeb, Amar Nazirun, Riskal H. Saleh, Syahrul Ibrahim Rahmat Fabanyo yang saling bekerja sama, memberikan saran dan membantu satu sama lain. Dan susah senang telah dilewati bersama hingga penulis dapat menyelesaikan skripsi ini.

15. Dan semua pihak yang tidak dapat penulis sebutkan yang telah membantu baik secara langsung maupun tidak langsung dalam menyelesaikan laporan skripsi ini. Walaupun demikian dalam penelitian ini, penulis menyadari masih belum sempurna. Oleh karena itu harapan penulis dalam memberikan kritik dan saran dari semua pihak yang bersifat membangun.

Ternate, 29 Juli 2024

Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
HALAMAN PENGESAHAN	ii
HALAMAN PERNYATAAN KEASLIAN	iii
HALAMAN PERSEMBAHAN	iv
KATA PENGANTAR	v
DAFTAR ISI	vii
DAFTAR GAMBAR	ix
DAFTAR TABEL	x
ABSTRAK	xi
BAB I PENDAHULUAN	
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah.....	2
1.4. Tujuan Penelitian.....	2
1.5. Manfaat Penelitian.....	3
1.6. Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	
2.1. Penelitian Terkait.....	4
2.2. Algoritma <i>K-Nearest Neighbor</i> (KNN)	9
2.3. <i>Python</i>	10
2.4. Suara	10
2.5. <i>Mel Frequency Cepstrum Coefficients</i> (MFCC).....	11
2.6. <i>Klasifikasi</i>	11
2.7. <i>Flowchart</i>	12
2.8. <i>White Box Testing</i>	13
2.9. Bahasa Daerah.....	13
2.10. Bahasa Indonesia.....	14
2.11. Bahasa Tidore	15

BAB III METODE PENELITIAN

3.1	Objek dan Waktu Penelitian	17
3.2.	Metode Pengumpulan Data.....	17
3.3.	Alur Penelitian	17
3.4.	Ekstraksi Ciri <i>Mel Frequency Cepstrals Coefficients</i>	18
3.5.	Algoritma <i>K-Nearest Neighbor</i>	20
3.6.	Contoh Perhitungan <i>K-Nearest Neighbor</i>	21
3.7.	Perencanaan Pengujian Sistem	23
3.8.	Alat Penelitian	23
3.8.1.	Perangkat Keras (<i>Hardware</i>).....	23
3.8.2.	Perangkat Lunak (<i>Software</i>).....	23

BAB IV HASIL DAN PEMBAHASAN

4.1	Pengumpulan Data.....	25
4.2.	Ekstraksi Ciri MFCC	26
4.3.	Membuat Dan Melatih Model KNN.....	27
4.4.	Evaluasi Model <i>K-Nearest Neighbor</i>	29
4.5.	<i>White Box</i>	31
4.6.	Pembahasan.....	35

BAB V PENUTUP

5.1.	Kesimpulan	36
5.2.	Saran	36

DAFTAR PUSTAKA

LAMPIRAN

DAFTAR GAMBAR

	Halaman
Gambar 3.1 Diagram Alir Penelitian	18
Gambar 3.2 Tahapan Ekstraksi Ciri MFCC	18
Gambar 3.3 Diagram Proses Algoritma KNN	21
Gambar 4.1 Folder Penyimpanan Dataset Suara	25
Gambar 4.2 Library yang Digunakan	26
Gambar 4.3 Mendefinisikan Direktori	26
Gambar 4.4 Ekstrak dan Menyimpan Fitur	26
Gambar 4.5 Hasil Running	27
Gambar 4.6 Membagi Dataset	27
Gambar 4.7 Mengubah Bentuk Data Pelatihan	28
Gambar 4.8 Membuat Model	28
Gambar 4.9 Menggunakan 3 Tetangga Terdekat	29
Gambar 4.10 Memprediksi dan Evaluasi Data	29
Gambar 4.11 Laporan Klasifikasi	29
Gambar 4.12 Hasil Laporan Klasifikasi	30
Gambar 4.13 Hasil Laporan Klasifikasi	30
Gambar 4.14 Listing Program Ekstraksi dan Penyimpanan Fitur	32
Gambar 4.15 Ekstraksi Fitur	33
Gambar 4.16 <i>Flowgraph</i> Pengujian	34

DAFTAR TABEL

	Halaman
Tabel 2.1 Penelitian Terkait	4
Tabel 2.2 <i>Flowchart</i>	12
Tabel 2.2 Kosakata Tidore.....	15
Tabel 3.1 Spesifikasi Perangkat Keras.....	23
Tabel 3.2 Spesifikasi Perangkat Lunak	24
Tabel 4.1 Laporan Klasifikasi.....	31

ABSTRAK

PENERAPAN ALGORITMA *K-NEAREST NEIGHBOR* (KNN) DALAM KLASIFIKASI SUARA BERDASARKAN BAHASA DAERAH TIDORE

Muh. Irfan Mansyur¹, Muhammad Sabri Ahmad, ² Muhammad Fhadli³
Program Studi Informatika, Fakultas Teknik, Universitas Khairun
Jl. Jati Metro, Kota Ternate

E-mail : irfanmnsyr55@gmail.com¹, Muhammadsabri@gmail.com², muhammadfhadli@ymail.com³

Bahasa adalah atribut manusiawi yang membedakan manusia dari makhluk lain. Studi ini dilakukan dengan mengklasifikasikan suara berdasarkan bahasa daerah Tidore menggunakan algoritma *K-Nearest Neighbor* (KNN). Data suara dalam bahasa Tidore dan Indonesia dikumpulkan melalui rekaman, kemudian diekstraksi fiturnya menggunakan *Mel Frequency Cepstral Coefficients* (MFCC). Algoritma KNN diterapkan untuk klasifikasi dengan jumlah tetangga (k) yang berbeda. Hasil penelitian menunjukkan bahwa KNN dapat mengklasifikasikan suara ke dalam dua kelas (bahasa Tidore dan Indonesia) dengan cukup baik. Akurasi model bervariasi tergantung jumlah tetangga yang digunakan, dengan akurasi tertinggi 97,01% pada $k=1$, 76,62% pada $k=3$, dan 66,67% pada $k=5$. Penggunaan 3 tetangga memberikan keseimbangan yang lebih baik antara bias dan *varians*. Metode pengujian *white box* juga diterapkan untuk menguji proses ekstraksi dan penyimpanan fitur. Penelitian ini berkontribusi pada upaya pelestarian bahasa daerah Tidore melalui teknologi pengenalan suara.

Kata kunci : Klasifikasi Suara, Bahasa Daerah Tidore, *K-Nearest Neighbor*, *Mel Frequency Cepstral Coefficients*, Ekstraksi Fitur

APPLICATION OF THE *K-NEAREST NEIGHBOR* (KNN) ALGORITHM IN VOICE CLASSIFICATION BASED ON THE TIDORE REGIONAL LANGUAGE

Abstract

Language is a human attribute that distinguishes humans from other creatures. This study was conducted by classifying speech based on the Tidore regional language using the K-Nearest Neighbor (KNN) algorithm. Voice data in Tidore and Indonesian languages were collected through recordings, then their features were extracted using Mel Frequency Cepstral Coefficients (MFCC). The KNN algorithm was applied for classification with different numbers of neighbors (k). The results showed that KNN can classify speech into two classes (Tidore and Indonesian languages) quite well. The model's accuracy varied depending on the number of neighbors used, with the highest accuracy of 97.01% at $k=1$, 76.62% at $k=3$, and 66.67% at $k=5$. Using 3 neighbors provided a better balance between bias and variance. White box testing methods were also applied to test the feature extraction and storage process. This research contributes to efforts to preserve the Tidore regional language through speech recognition technology.

Keywords: *Voice classification, Tidore Regional Language, K-Nearest Neighbor, Mel Frequency Cepstral Coefficients, Feature Extraction*

BAB I

PENDAHULUAN

1.1. Latar Belakang

Bahasa adalah atribut manusiawi yang membedakan manusia dari makhluk lain. Penggunaan bahasa sebagai alat komunikasi adalah komponen paling penting dari kebudayaan, yang juga menggambarkan semua aspek budaya. Banyak fungsi bahasa berkaitan dengan kebudayaan, seperti sebagai alat untuk pengembangan, pembinaan, jalur pembinaan, dan inventarisasi. karena itu, bahasa tidak dapat dipisahkan dari budaya manusia karena bahasa berinteraksi satu sama lain dan budaya berinteraksi satu sama lain atau kausalitas (Juniar, 2019). Di Indonesia memiliki banyak bahasa dan budaya. Sebanyak 719 bahasa daerah digunakan, 707 di antaranya masih ada, tetapi Unesco baru mencatatkan 143 bahasa daerah (Momole, 2022).

Maluku Utara adalah salah satu provinsi di Indonesia yang memiliki beragam suku, budaya, dan Bahasa yang terdiri dari delapan kabupaten dan dua kotamadya. Kota tidore dan Kota Ternate merupakan salah satu kota yang ada di provinsi Maluku Utara, Kemiripan antara bahasa Tidore dan Ternate jika didengar sekilas antara kedua bahasa tersebut terdapat kemiripan baik leksikon maupun fonologi. Kemiripan dalam bidang leksikon seperti kata gia 'tangan' dan fika 'debu' (Duwila, 2022).

Berdasarkan masalah yang ada perlu adanya klasifikasi suara dengan menggunakan Algoritma *K-Nearest Neighbor* (KNN). KNN adalah metode yang mengelompokkan data baru berdasarkan jarak data baru itu ke beberapa data atau tetangga (*neighbor*) terdekat.

Selanjutnya untuk mengatasi masalah yang sudah disampaikan sebelumnya, maka penulis mengusulkan untuk melakukan penelitian guna mengklasifikasikan suara bahasa

menggunakan Algoritma *K-Nearest Neighbor* ini memberikan kemudahan mengenali bahasa Tidore. Maka dari itu penulis melakukan penelitian dengan judul "Klasifikasi Suara Berdasarkan Bahasa Daerah Menggunakan Algoritma *K-Nearest Neighbor*".

1.2. Rumusan Masalah

Berdasarkan Latar Belakang uraian masalah di atas, peneliti merumuskan permasalahan terkait penelitian ini adalah bagaimana Penerapan Algoritma *K-Nearest Neighbor* (KNN) Dalam Klasifikasi Suara Berdasarkan Bahasa Daerah Tidore dalam mengklasifikasikan suara berdasarkan bahasa daerah tidore.

1.3. Batasan Masalah

Adapun batasan masalah yang ada dalam penelitian ini dibatasi beberapa masalah sebagai berikut:

1. Penelitian ini akan berfokus pada Bahasa Daerah Tidore sebagai bahasa yang menjadi objek implementasi algoritma *K-Nearest Neighbor*. Bahasa lainnya tidak akan dicakup dalam penelitian ini.
2. Penelitian ini hanya mengklasifikasikan suara dalam Bahasa Tidore.
3. Penelitian ini bersumber dari Desa Jaya di Kota Tidore.
4. Bentuk implementasi pada penelitian ini hanya berupa model program.

1.4. Tujuan Penelitian

Berdasarkan dari rumusan masalah diatas, penelitian ini bertujuan melakukan klasifikasi pada Bahasa Daerah berdasarkan suara.

1.5. Manfaat Penelitian

Penelitian ini diharapkan dapat bermanfaat sebagai berikut:

1. Penelitian ini dapat membantu menjaga dan melestarikan Bahasa Daerah Tidore

2. dengan menggunakan teknologi.
3. Penelitian ini dapat meningkatkan pengembangan teknologi pengenalan suara.
4. Penelitian ini dapat mengklasifikasikan suara berdasarkan bahasa daerah dapat memberikan kontribusi pada pengakuan dan pemahaman budaya lokal

1.6 Sistematika Penulisan

BAB I PENDAHULUAN

Pada bab ini menjelaskan tentang latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan untuk kasus yang akan dipecah.

BAB II TINJAUAN PUSTAKA

Bab ini menjelaskan tentang teori-teori yang akan digunakan untuk penelitian.

BAB III METODE PENELITIAN

Bab ini menjelaskan langkah-langkah metode penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menjelaskan hasil penelitian

BAB V PENUTUP

Bab ini memuat kesimpulan dan saran penelitian

BAB II

TINJAUAN PUSTAKA

2.1. Penelitian Terkait

Penelitian terkait merupakan penelitian terdahulu yang berkaitan dengan penelitian yang sedang diteliti, untuk mengetahui perbedaan penelitian terkait dengan penelitian yang akan dilakukan dapat dilihat pada tabel 2.1.

Tabel 2. 1 Penelitian Terkait

No	Penulis	Judul Penelitian	Hasil Penelitian
1.	Rizki Alexander Manullang, Fricles Ariswanto Sianturi	Penerapan Algoritma <i>K-Nearest Neighbor</i> untuk Memprediksi Kelulusan Mahasiswa	penelitian yang dilakukan mengenai Penerapan Algoritma <i>K-Nearest Neighbour</i> untuk Memprediksi Kelulusan Mahasiswa, maka dapat diambil kesimpulan sebagai berikut: Dengan menguji data sebanyak 215 data dengan menerapkan metode KNN, didapatkan hasil perhitungan jumlah mahasiswa yang lulus untuk Program Studi Teknik Informatika sebanyak 148 mahasiswa, dan jumlah yang tidak lulus sebanyak 67 mahasiswa. Dari hasil penelitian di atas tentang Penerapan Algoritma <i>K-Nearest Neighbor</i> untuk memprediksi kelulusan mahasiswa, data yang digunakan sebagai data training merupakan data kelulusan mahasiswa tahun 2018, dan data yang digunakan sebagai data testing merupakan data kelulusan mahasiswa tahun 2020 jurusan Teknik Informatika dengan atribut yang digunakan adalah: IP semester 1-6. (Manullang, 2021)
2	Joni Saputra, Yuan Sa'adati, Valian Yoga	Klasifikasi Kematangan Buah Alpukat Mentega	Hasil penelitian Klasifikasi Kematangan Buah Alpukat Mentega Menggunakan Metode K-

	Pudya Ardhana, M Afriansyah	Menggunakan Metode <i>KNearest Neighbor</i> Berdasarkan Warna Kulit Buah	<i>Nearest Neighbor</i> Berdasarkan Warna Kulit Buah diperoleh dengan mengkasifikasi citra terhadap tiga objek identifikasi, diantaranya Mentah, Setengah Matang, dan Matang, dengan total sample sejumlah 20 buah alpukat alpukat mentega. Terdiri dari 7 buah alpukat mentega mentah, 6 buah alpukat mentega setengah matang, dan 7 buah alpukat mentega matang. Dimana hasil pengujian sampel buah alpukat mentega mentah di atas menunjukkan tingkat akurasi sebesar 85,71%, alpukat mentega setengah matang di atas menunjukkan tingkat akurasi sebesar 66,66%, alpukat mentega matang sebesar 80%. Sampel data pengujian citra buah alpukat setengah matang menunjukkan hasil tidak akurat lebih tinggi di bandingkan dengan buah alpukat mentah dan matang, dengan jumlah 2 ketidak akuratan pada buah setengah matang dan masing – masing 1 ketidak akuratan pada buah mentah dan matang. Optimasi algoritma <i>K-nearest neighbor</i> pada penelitian ini dapat membuat proses peran pengujian citra dalam meningkatkan hasil klasifikasi yang lebih efektif. Dapat di lihat dari hasil pengujian yang di peroleh pada nilai persentase nilai training 96% yang menunjukkan tingkat akurasi akurat pada buah alpukat mentega setengah matang. (Saputra, 2023)
3	Yovi Pratama, Anton Prayitno, Defri Nazrian, Nur Aini, Yoga Rizki R, Errissya Rasywir	Klasifikasi Penyakit Gagal Jantung Menggunakan Algoritma <i>K-Nearest Neighbor</i>	Berdasarkan tahapan penelitian yang telah dilakukan pada penerapan algoritma <i>K-Nearest Neighbor</i> pada dataset penyakit gagal jantung. Kesimpulan yang diperoleh dari penelitian ini, yaitu penerapan algoritma KNN terhadap data penyakit gagal jantung

			<p>dilakukan dengan pengujian akurasi yang dilakukan dengan menggunakan aplikasi rapidminer. Dataset yang digunakan sebanyak 918 data dan 12 atribut. Kemudian hasil implementasi algoritma KNN pada aplikasi ripedminer dilakukan dengan menggunakan pergantian nilai k, dan hasil akurasi tertinggi didapat pada nilai k=9 dengan akurasi sebesar 70,65%, nilai presesinya adalah 75 %, sedangkan recall menghasilkan 70,73%. (Pratama, 2022).</p>
4	<p>Musliha Syam, Dr. Ir. Jangkung Raharjo, Raditiana Patmasari, S.T., M.T.</p>	<p>Identifikasi Asal Daerah Berdasarkan Suara Manusia Dengan Metode <i>Linier Predictive Coding</i> (LPC)</p>	<p>Perancangan sistem yang dibuat mampu mendeteksi asal daerah dengan menggunakan metode ekstraksi ciri Linier Predictive Coding (LPC) dan klasifikasi Jaringan Saraf Tiruan (JST) dengan metode Propagasi Balik yang didapatkan nilai akurasi 80.76% dengan orde LPC 12 dan menggunakan parameter ciri mean, 3 hidden layer sebagai parameter di JST-PB. Pemakaian jumlah hidden layer pada JST-PB mempengaruhi kinerja sistem. Dimana semakin banyak jumlah hidden layer maka semakin banyak node yang terhubung antar neuron dan semakin banyak jumlah node maka semakin banyak bobot yang digunakan yang mengakibatkan jaringan dapat menyimpan pengetahuan yang lebih baik mengenai pola ciri dari proses pelatihan. Akan tetapi sistem dapat mengalami overfitting. Saat menggunakan 1 hidden layer mendapatkan akurasi 63.30%, saat 2 hidden layer mendapatkan akurasi yang lebih baik bernilai 71.15% dengan waktu komputasi 0.0474 detik. Saat menggunakan 3 hidden layer mendapatkan nilai</p>

			akurasi 80.7692% dengan waktu komputasi lebih efisien bernilai 0.0456 detik. Oleh karena itu, menambahkan jumlah hidden layer dapat meningkatkan nilai akurasi melalui proses trial and error (Syam, 2019)
5	Faiza R. Irawan, Ahmad Jazuli, Tutik Khotimah	Analisis Sentimen Terhadap Pengguna Gojek Menggunakan Metode <i>K-Nearest Neighbors</i>	Berdasarkan hasil penelitian menggunakan metode <i>K-Nearest Neighbor</i> mendapatkan akurasi sebesar 79,43% pada k=15, dengan sentiment positif sebesar 47,59%, negative 8,66% dan netral 43,75% penelitian ini diawali dengan crawling data selanjutnya data mentah diolah terlebih dahulu pada proses <i>preprocessing</i> . Setelah itu data bersih dilakukan <i>labelling</i> (pembobolan kata), kemudian tahap klasifikasi menggunakan metode <i>K-Nearest Neighbor</i> dengan perhitungan jarak <i>Euclidean</i> . Tahap akhir yaitu tahap evaluasi terhadap metode yang digunakan dengan <i>confusion matrix</i> dengan <i>split</i> data 80:20 data <i>training</i> dan <i>data testing</i> .. (Faiza R. Irawan, Ahmad, 2022).
6.	Revanto Alif Nawasta, Nur Heri Cahyana, Heriyanto	Implementasi Ekstraksi Ciri <i>Mel-Frequency Cepstral Coefficient</i> Menggunakan <i>K-Nearest Neighbor</i> Untuk Deteksi Emosi Berdasarkan Intonasi Suara	Berdasarkan penelitian yang telah dilakukan, didapatkan kesimpulan yang dapat dinyatakan sebagai berikut. Kombinasi antara metode ekstraksi ciri MFCC dan metode KNN mampu dalam mendeteksi emosi berdasarkan intonasi suara pada pria dengan akurasi sebesar 88,8% terhadap data test dan 80% terhadap data baru. MFCC dan KNN juga mampu mendeteksi emosi berdasarkan intonasi suara pada wanita dengan akurasi 92,5% terhadap data test namun terhadap data baru hanya mampu mencapai akurasi sebesar 50%. Model pada wanita mengalami kondisi

			overfitting yang disebabkan oleh data augmentasi pitch shifting yang tidak mampu menghasilkan sampel data yang cukup untuk secara akurat merepresentasikan semua kemungkinan nilai data input (Nawasta, 2023).
7.	Mutammimul Ula, Ria Zulhusna, Rizki Putra Fhonna, Angga Pratama	Penerapan Model Klasifikasi <i>K-Nearest Neighbor</i> Dalam Pencarian Kesesuaian Pekerjaan	Penerapan sistem informasi lowongan kerja untuk wilayah dapat memberikan informasi data lowongan pekerjaan yang ada di Aceh. Hasil penelitian ini memberikan pekerjaan yang cocok untuk para pencari pekerja. Sistem klasifikasi menggunakan model KNN dapat mengklasifikasikan data lowongan pekerjaan, berdasarkan kuesioner yang diisi oleh pekerja dan hasil rekomendasi lowongan kerja diperoleh dengan akurasi aplikasi 8,36% dengan data akuratnya 19 yang rekomendasi dan 6 data yang tidak rekomendasi. (Ula, 2022)
8.	F. Tempola, R. Rosihan, R. Adawiyah	Validasi Holdout Untuk Perbandingan Klasifikasi <i>Naïve Bayes</i> dan KNN Penerima Kartu Indonesia Pintar	Pengujian dilakukan pada data latih 1 hingga 149 dan data uji 149 hingga 1 dengan menggunakan nilai $K=3$ mendapatkan akurasi terendah sebesar 48% dan akurasi tertinggi sebesar 100% dengan rata-rata total akurasi sebesar 84,68% serta pengujian menggunakan $K=5$ mendapatkan akurasi terendah sebesar 0% dan akurasi tertinggi sebesar 100% dengan rata-rata total akurasi sebesar 84,08%, dan menggunakan $K=7$ akurasi terendah 0% dan akurasi tertinggi sebesar 100% dengan rata-rata total akurasi sebesar 84,89%. (Tempola, 2021)

Untuk itu perbedaan hasil analisa penelitian-penelitian terkait diatas dengan penelitian yang akan dilakukan penulis mencoba melakukan Klasifikasi Suara

Berdasarkan Bahasa Daerah Menggunakan Algoritma KNN, berbeda dari penelitian sebelumnya “Penerapan Algoritma *K-Nearest Neighbor* untuk Memprediksi Kelulusan Mahasiswa” yang dibuat oleh Rizki Alexander Manullang yang Hasil penelitian Klasifikasi Kematangan Buah Alpukat Mentega Menggunakan Metode *K-Nearest Neighbor* Berdasarkan Warna Kulit Buah diperoleh dengan mengkasifikasi citra terhadap tiga objek identifikasi, diantaranya Mentah, Setengah Matang, dan Matang, dengan total sample sejumlah 20 buah alpukat alpukat mentega. Terdiri dari 7 buah alpukat mentega mentah, 6 buah alpukat mentega setengah matang, dan 7 buah alpukat mentega matang. Dimana hasil pengujian sampel buah alpukat mentega mentah di atas menunjukan tingkat akurasi sebesar 85,71%, alpukat mentega setengah matang di atas menunjukan tingkat akurasi sebesar 66,66%, alpukat mentega matang sebesar 80%.

2.2. Algoritma *K-Nearest Neighbor* (KNN)

Algoritma KNN adalah algoritma pembelajaran terawasi dan banyak digunakan untuk klasifikasi terhadap objek berdasarkan data pembelajaran yang paling dekat dengan objek tersebut. Yaitu mengelompokkan suatu data baru berdasarkan jarak data ke beberapa data terdekat atau tetangga terdekat (Tempola, 2021). Yang pertama sebelum mencari jarak dari data ke tetangga adalah menentukan nilai *K* (*Neighbor*) biasanya nilai *k* yang digunakan adalah nilai ganjil untuk menghindari ikatan saat menentukan kelas mayoritas, Ketika *k* adalah genap ada kemungkinan memiliki jumlah tetangga yang sama dari kelas berbeda sehingga kelas mayoritas menjadi ambigu yang menyebabkan potensi penurunan akurasi, dengan memilih nilai ganjil untuk *k* memastikan suara mayoritas yang jelas dan meningkatkan performa algoritma *K-Nearest*, maka digunakan rumus *Euclidean* dengan persamaan 2.1 (Pratama, 2022)

Tahapan dari Algoritma KNN:

1. Tentukan nilai K atau tetangga.
2. Hitung jarak Euclidean antara data test dengan semua data training.
3. Urutkan jarak dari yang terkecil.
4. Tentukan kelas mayoritas

$$d(a, b) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \dots\dots\dots(2.1)$$

Keterangan

$d(a,b)$: jarak *Euclidean*

x : data 1

y : data 2

n : jumlah fitur.

2.3. *Python*

Python, sebuah bahasa pemrograman yang saat ini sangat populer dan banyak digunakan, dikembangkan oleh Guido van Rossum dan pertama kali dirilis pada tahun 1991. Bahasa Python yang mudah dipelajari dan fleksibel ini memiliki sejumlah fungsi yang luas dan beragam, termasuk pengembangan web (sisi server), pembuatan perangkat lunak atau aplikasi desktop, penyelesaian persamaan matematika yang kompleks, pemrosesan data besar, kecerdasan buatan, dan pembuatan skrip sistem. Selain itu, Python juga dapat digunakan untuk pemrograman mikrokontroler melalui MicroPython, yang memungkinkan pengembangan perangkat Internet of Things (IoT) dengan lebih mudah..

Beberapa aplikasi konkret dari *Python* melibatkan penggunaannya di *server* untuk menciptakan aplikasi web, integrasinya dengan perangkat lunak untuk mengatur alur kerja, koneksi ke sistem basis data, penanganan data besar serta pengerjaan matematika

kompleks. *Python* juga dapat digunakan untuk *prototyping* cepat atau pengembangan perangkat lunak yang siap diproduksi. (Ma'arif, 2020).

2.4. Suara

Suara dalam ilmu fisika ialah gelombang longitudinal yang merambat melalui zat perantara atau medium yang dapat dilalui suara berupa zat padat, cair dan gas. Hal tersebut juga berpengaruh dari seberapa jauh sifat benda tersebut untuk dapat menghantarkan suara melalui udara. Suara memiliki peran penting bagi semua makhluk hidup, contohnya manusia. suara sendiri adalah alat utama dalam berkomunikasi, menyatakan kepribadian dan kondisi pembicara.

Pengenalan suara adalah suatu proses mengidentifikasi suara berdasarkan parameter yang diambil oleh penangkap input suara. Pengenalan suara sendiri terbagi menjadi tiga bentuk pendekatan, yaitu pendekatan kecerdasan buatan (*artificial intelligent approach*), akustik forensik (*acoustic phonetic approach*), dan pendekatan pengenalan pola (*pattern recognition approach*) (Prayugo, 2023).

2.5. Mel Frequency Cepstrum Corfficients (MFCC)

Mel Frequency Cepstrum Corfficients atau MFCC adalah ekstraksi fitur yang sering digunakan dalam pengenalan suara. MFCC di perkenalkan oleh Davis dan Mermelstein pada tahun 1980 mengusulkan memakai MFCC untuk sistem pengenalan suara (Meutia Puspitasari 2023). MFCC adalah ekstraksi fitur yang mengubah sinyal suara menjadi vector parameter Adapun kelebihan dari MFCC:

1. Dapat menangkap karakter suara untuk mengidentifikasi pola suara tertentu.
2. Memberikan output berupa vector dengan ukuran data yang kecil tetapi tidak menghilangkan karakteristik suara yang diekstrak.

3. Cara kerja MFCC mirip cara kerja pendengaran manusia dalam memberikan persepsi suara yang di dengarkan.

2.6. Klasifikasi

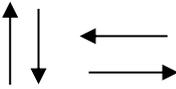
Klasifikasi merupakan proses menemukan model atau fungsi yang mendeskripsikan dan membedakan data kedalam kelas- kelas. Memeriksa karakteristik dan objek dan memasukkan objek dalam salah satu kelas yang sudah didefinisikan sebelumnya (Andreas, 2019).

2.7. Flowchart

Flowchart merupakan representasi grafis dari langkah-langkah dan urutan prosedur suatu program. Umumnya, *flowchart* memainkan peran penting dalam penyelesaian masalah yang memerlukan kajian dan evaluasi lebih lanjut. *Flowchart* dapat digunakan untuk menggambarkan berbagai jenis kegiatan, termasuk kegiatan manual, pemrosesan data, atau keduanya. *Flowchart* sendiri merupakan serangkaian simbol-simbol yang digunakan untuk membangun dan menggambarkan proses atau alur suatu program (Budiman Ilham, 2021). Berikut adalah tabel 2.2 *Flowchart*

Tabel 2. 2 *Flowchart*

Simbol	Nama	Fungsi
	Terminal	Digunakan untuk memulai atau mengakhiri program.
	<i>Input/Output</i>	Digunakan untuk menyatakan input atau output tanpa melihat jenisnya.
	<i>Manual Operation</i>	Digunakan untuk menunjukkan pengolahan yang tidak dilakukan oleh komputer.
	<i>Decision</i>	Digunakan untuk memilih proses yang akan dilakukan berdasarkan kondisi tertentu.

	<i>Processing</i>	Digunakan untuk menunjukkan pengolahan data yang dilakukan oleh komputer.
	<i>Disk Storage</i>	Digunakan untuk menyatakan masukan dan keluaran yang berasal dari disk.
	<i>Flow Direction Symbol / Connecting Line</i>	Berfungsi untuk menghubungkan simbol yang satu dengan yang lainnya, menyatakan arus suatu proses.

2.8. *White Box Testing*

Pengujian kotak putih, atau *white box testing*, adalah metode pengujian perangkat lunak yang dilakukan dengan menganalisis dan meneliti struktur internal serta kode dari perangkat lunak tersebut. Berbeda dengan pengujian kotak hitam (*black box testing*) yang hanya mengevaluasi *input* dan *output*, *white box testing* berfokus pada aliran *input* dan *output* di dalam perangkat lunak. Tujuan utama dari *white box testing* adalah memastikan setiap komponen program berfungsi sebagaimana mestinya dengan memeriksa kode program secara mendetail (Salsabilla, 2023).

Cycloimatic Complexity, adalah besaran perangkat lunak yang menyediakan acuan kuantitatif kompleksitas suatu logika dalam program. Rumus menghitung *Cyclomatic Complexity* dapat dilihat dengan persamaan 2.2

$$CC = E - N + 2 \dots \dots \dots (2.2)$$

Keterangan 2.2:

E = jumlah *edges* pada *flowgraph*

N = jumlah *nodes* pada *flowgraph*

P = jumlah *predicates* nodes pada *flowgraph*

2.9. Bahasa Daerah

Bahasa adalah sifat manusiawi yang membedakan manusia dari makhluk lain.

Bahasa, yang digunakan untuk berkomunikasi, merupakan komponen paling penting dari kebudayaan, dan merupakan komponen yang menggambarkan setiap aspek kebudayaan. Menurut Devianty (2017: hlm. 226), bahasa memiliki banyak fungsi untuk kebudayaan, termasuk sebagai alat untuk pengembangan, pembinaan, jalur pembinaan, dan inventarisasi. Oleh karena itu, bahasa tidak dapat dipisahkan dari kehidupan budaya manusia karena keduanya berhubungan satu sama lain secara kausalitas atau timbal balik. Bahasa adalah bagian dari budaya manusia, dan bahasa banyak mempengaruhinya. Lebih penting lagi, bahasa adalah komponen yang memungkinkan kebudayaan terbentuk, jadi bahasa tidak dapat terjadi tanpanya. Bahasa merupakan cerminan kebudayaan masyarakat (Barus, 2019).

Bahasa Tidore, sebagai salah satu bahasa daerah diprovinsi Maluku Utara, merupakan bagian integral dari kekayaan budaya setempat yang perlu dilestarikan. Kajian ini menggaris bawahi bahwa bahasa daerah tidak hanya mencerminkan identitas budaya suatu daerah tetapi juga memperkuat identitas budaya nasional, dengan Bahasa Indonesia berperan sebagai bahasa persatuan. Bahasa daerah tetap menjadi sarana yang penting untuk merepresentasikan esensi kebudayaan setempat. Bahasa lokal menunjukkan budaya nasional dan menerima bahasa Indonesia sebagai bahasa persatuan. Bahasa lokal masih digunakan untuk menunjukkan pusat budaya lokal (Djumadil, 2021).

Penelitian ini juga mengungkapkan fakta bahwa beberapa bahasa daerah, termasuk bahasa di daerah Halmahera, Maluku Utara, menghadapi tantangan serius karena jumlah penuturnya yang terbatas, bahkan hanya tinggal puluhan penutur. Bahasa Tidore, contohnya, merupakan salah satu bahasa non-Austronesia sejajar dengan bahasa non-Austronesia lainnya seperti bahasa Ternate dan bahasa-bahasa di daratan

Halmahera Utara. Penting untuk mengapresiasi dan menjaga bahasa-bahasa ini sebagai bagian tak terpisahkan dari keanekaragaman bahasa dan budaya di Indonesia

2.10. Bahasa Indonesia

Sejarah perkembangan bahasa Indonesia sebagai identitas bangsa terbagi dalam tiga tahapan utama. Pertama, perkembangan fungsi bahasa Indonesia dimulai sejak sebelum kemerdekaan. Pada tahap ini, bahasa Indonesia berperan sebagai bahasa pemersatu berbagai suku bangsa. Saat itu, bahasa Indonesia mampu menyatukan bangsa. Oleh karena itu, banyak yang berpendapat bahwa bahasa Indonesia berkontribusi besar terhadap persatuan bangsa hingga kini. Kedua, pada masa awal kemerdekaan, fungsi bahasa Indonesia mengalami peningkatan menjadi bahasa negara. Selain menjadi simbol persatuan, bahasa Indonesia juga menjadi bagian integral dari kehidupan bernegara. Ketiga, fungsi bahasa Indonesia terus berkembang seiring dengan peningkatan kekuatan negara. Jumlah ahli bahasa dan sastra semakin banyak, dan banyak karya ilmiah berkualitas telah ditulis dalam bahasa Indonesia (Nurrahmah, 2023).

2.11. Bahasa Tidore

Bahasa memiliki peran yang sangat penting di dunia ini karena berfungsi sebagai alat komunikasi, baik secara lisan maupun tulisan, bagi para penuturnya. Selain itu, bahasa juga menjadi identitas bagi masyarakat yang menggunakannya di wilayah atau daerah tertentu di Indonesia, di mana bahasa tersebut dikategorikan sebagai bahasa daerah. Salah satu contohnya adalah Bahasa Tidore, yang merupakan bahasa daerah yang bahasa daerah yang digunakan di Tidore (Husnia, 2020).

Berikut merupakan Kosakata bahasa daerah Tidore yang terdapat pada kamus bahasa Tidore dapat dilihat pada tabel 2.3.

Tabel 2. 3 Kosakata Tidore

Kalimat	Arti
<i>ake</i>	air
<i>ahu</i>	hidup
<i>aru</i>	gayung
<i>amo</i>	sukun
<i>Ake malao</i>	Muara sungai
<i>baja</i>	bujuk
<i>balaco</i>	hujau
<i>bawa</i>	bawang
<i>balutu</i>	Nyenyak
<i>barisi</i>	Rumput berduru
<i>beke</i>	Ikan kering
<i>cala</i>	ribu
<i>cou</i>	Kerja bakti
<i>cam</i>	kunyah
<i>dahe</i>	terima
<i>dafu</i>	Besan perempuan
<i>diyali</i>	pengganti
<i>eno</i>	Sampah
<i>e.la</i>	melangkah
<i>foli</i>	belanja

BAB III

METODE PENELITIAN

3.1. Objek dan Waktu Penelitian

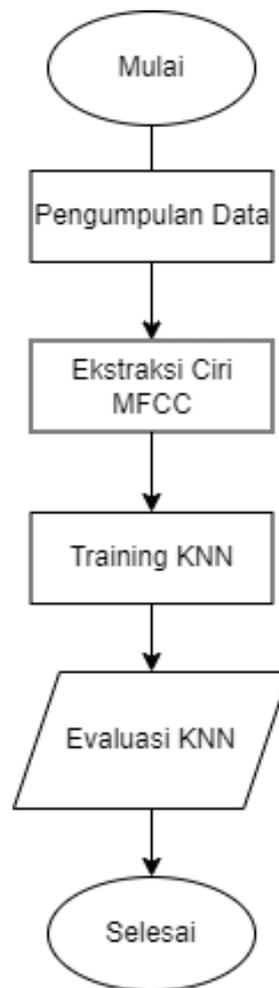
Objek yang diteliti pada penelitian ini yaitu Bahasa Tidore yang dimana data yang akan diteliti adalah suara manusia yang kemudian diklasifikasi menggunakan Algoritma *K-Nearest Neighbor*. Waktu penelitian yang dilakukan pada semester ganjil tahun ajaran 2023/2024.

3.2. Metode Pengumpulan Data

Pada tahapan ini Data suara yang berupa kosa kata dalam bahasa Daerah dikumpulkan secara langsung. Data suara di dapatkan dengan mengambil sampel suara dengan cara merekam pengucapan kalimat dalam bahasa daerah Tidore dan disimpan dalam format WAV. Data suara dibagi menjadi dua, yakni data training sebesar 80% dan data testing sebesar 20%.

3.3. Alur Penelitian

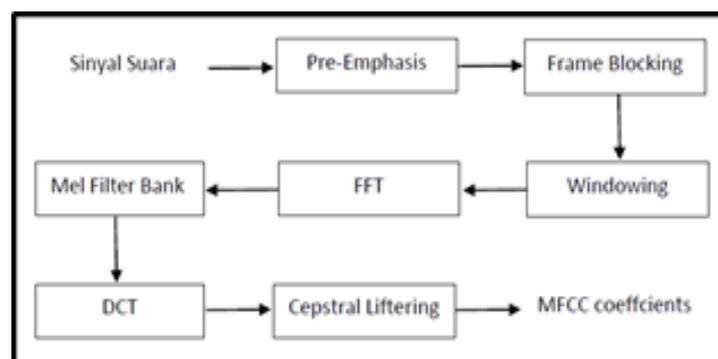
Dalam melakukan penelitian ini diperlukan beberapa tahapan untuk mencapai tujuan yang ditetapkan. Pada tahapan penelitian ini yang pertama yaitu mengumpulkan data yang berupa sampel suara. Selanjutnya data tersebut diekstraksi ciri menggunakan MFCC dan dilakukan pembagian data yang dibagi menjadi data *training* dan data *testing*, setelah itu dilakukan proses klasifikasi menggunakan *K-Nearest Neighbord* (KNN), dan didapatkan hasil dari klasifikasi *K-Nearest Neighbor* dan. Tahapan pada penelitian ini dapat dilihat pada gambar 3.1.



Gambar 3. 1 Diagram alir Penelitian

3.4. Ekstraksi Ciri *Mel Frequency Cepstrals Coefficients*

Setelah Data telah terkumpul maka dilakukan Ekstraksi Ciri menggunakan MFCC proses ekstraksi dapat dilihat pada gambar 3.2.



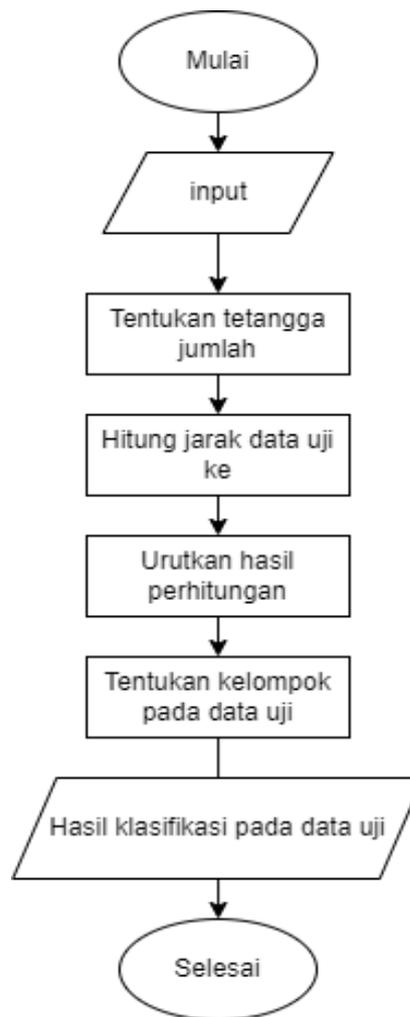
Gambar 3. 2 Tahapan Ekstraksi Ciri MFCC

1. *Pre-emphasis* merupakan tahap pertama pada ekstraksi ciri MFCC. Proses ini akan mempertahankan frekuensi-frekuensi tinggi pada spektrum yang biasanya hilang atau tereliminasi dalam proses produksi suara.
2. *Frame blocking* adalah tahap dimana sinyal suara disegmentasi menjadi beberapa bingkai. Secara umum pada proses *frame blocking*, dalam setiap bingkainya berukuran 20 - 25 milidetik dengan ukuran bagian yang tumpang tindih (M) antara satu bingkai dengan bingkai lainnya sebesar 30-50% dari panjang frame. Panjang frame (N) yang digunakan dalam proses ini sangat berpengaruh terhadap keberhasilan dalam analisis spektral. Semakin panjang ukuran frame yang digunakan maka akan semakin baik pula resolusi frekuensi yang ditunjukkan akan tetapi akan berpengaruh terhadap resolusi waktu yang dihasilkan
3. *Windowing* merupakan tahap pembobotan pada setiap frame yang telah terbentuk pada proses sebelumnya dengan menggunakan fungsi *window*. *Windowing* ini dilakukan untuk mengurangi kesenjangan atau discontinuitas signal setelah proses *frame blocking*.
4. *Fast Fourier Transform* (FFT) merupakan pengembangan dari *Discrete Fourier Transform* (DFT) yang bertujuan untuk mengkonversi sinyal digital dari domain waktu ke domain frekuensi. Dalam proses FFT ini, setiap frame N sampel akan diubah dari domain waktu ke domain frekuensi.
5. *Filter bank* adalah salah satu filter yang digunakan untuk mengetahui ukuran energy dari *frequency band* tertentu dalam sinyal suara. Untuk mengetahui ukuran energi yang tersedia pada setiap titik, nilai frekuensi yang terdapat pada spektrum FFT harus dipetakan ke dalam Mel scale dengan bantuan triangular filter bank.

6. *Discrete Cosine Transform (DCT)* merupakan langkah terakhir ekstraksi ciri dengan MFCC. Konsep dasar DCT adalah mendekorelasikan mel spectrum untuk menghasilkan representasi yang baik dari property spektral lokal. Pada tahap ini, nilai *Mel spectrum dalam frequency domain* akan dikonversi ke dalam time domain dengan tujuan untuk mendapatkan nilai koefisiennya.
7. Proses *cepstral liftering* ini adalah proses akhir dalam mengekstraksi fitur dengan MFCC. *Cepstral liftering* merupakan sebuah teknik yang digunakan untuk meminimalisasi sensitivitas dari koefisien cepstral yang dihasilkan dari tahapan utama dalam ekstraksi ciri menggunakan MFCC. Proses ini berfungsi untuk meningkatkan kualitas pengenalan suara.

3.5. Algoritma K-Nearest Neighbor

Algoritma yang digunakan pada proses klasifikasi ini adalah *K-Nearest Neighbor*, Algoritma ini di harapkan dapat mengklasifikasikan Bahasa daerah tidore. Berikut merupakan proses Algoritma KNN. Pertama menentukan parameter K (jumlah tetangga paling dekat), parameter K pada *testing* ditentukan berdasarkan nilai K pada saat *training*, setelah itu menghitung kuadrat jarak *Euclidean Distance* masing masing objek terhadap data sampel yang diberikan, setelah itu mengurutkan objek objek tersebut kedalam kelompok yang mempunyai jarak *Euclidean* terkecil, menentukan kelompok (klasifikasi KNN), didapatkan hasil klasifikasi



Gambar 3. 3 Diagram Proses Algoritma KNN

3.6. Contoh Perhitungan K-Nearest Neighbor

Berikut adalah contoh perhitungan KNN:

Tabel 3. 1 Contoh Kelas

Data	Fitur 1	Fitur 2	Kelas
A	1	2	Tidore
B	2	3	Indonesia
C	3	1	Tidore
D	4	4	Indonesia

Misalkan kita memiliki data baru $X = (2, 2)$ dan dua kelas tidore dan Indonesia

1. Tentukan nilai k (*neighbor*)

Misalkan kita gunakan $k = 3$, dengan Rumus $d(a, b) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$

2. Hitung jarak Euclidean antara data baru $X = (2, 2)$ dengan setiap data pelatihan

a. Jarak ke titik (1, 2)

$$\begin{aligned} &= \sqrt{(x_1 - y_1)^2 + (x_5 - y_5)^2} \\ &= \sqrt{(1 - 2)^2 + (2 - 2)^2} = \sqrt{(-1)^2 + (-0)^2} = \sqrt{1 + 0} = 1 \end{aligned}$$

b. Jarak ke titik (2, 3)

$$\begin{aligned} &= \sqrt{(x_2 - y_2)^2 + (x_5 - y_5)^2} \\ &= \sqrt{(2 - 2)^2 + (3 - 2)^2} = \sqrt{(0)^2 + (1)^2} = \sqrt{0 + 1} = 1 \end{aligned}$$

c. Jarak ke titik (3, 1)

$$\begin{aligned} &= \sqrt{(x_3 - y_3)^2 + (x_5 - y_5)^2} \\ &= \sqrt{(3 - 2)^2 + (1 - 2)^2} = \sqrt{(1)^2 + (-1)^2} = \sqrt{1 + 1} = 2 \end{aligned}$$

d. Jarak ke titik (4, 4)

$$\begin{aligned} &= \sqrt{(x_4 - y_4)^2 + (x_5 - y_5)^2} \\ &= \sqrt{(4 - 2)^2 + (4 - 2)^2} = \sqrt{(2)^2 + (2)^2} = \sqrt{4 + 4} = 8 \end{aligned}$$

3. Urutkan jarak dari yang terkecil:

A = 1

B = 1

C = 2

D = 8

4. Ambil K Tetangga terdekat yang sudah ditentukan yaitu $K = 3$

A (Tidore)

B (Indonesia)

C (Tidore)

5. Tentukan kelas mayoritas dari k tetangga

Kelas Tidore 2 dan Kelas Indonesia 1 jadi data baru X diklasifikasikan sebagai kelas

Tidore

3.7. Perencanaan Pengujian Sistem

Pada penelitian ini metode pengujian yang digunakan adalah *White box*. Bagian yang akan diuji proses ekstraksi dan penyimpanan fitur. Berikut merupakan Langkah-langkah pengujian sebagai berikut:

1. Membuat *flowchart* ekstraksi
2. Membuat *Flowgraph* ekstraksi
3. Mencari *Cyclomatic Complexity*
4. Tentukan *independent Part*

3.8. Alat Penelitian

Adapun alat dan bahan pada penelitian ini, terdiri dari perangkat keras (*hardware*) dan perangkat lunak (*software*) sebagai penunjang selama penelitian.

3.7.1 Perangkat Keras (*Hardware*)

Adapun spesifikasi perangkat keras (*hardware*) yang digunakan pada penelitian ini sebagai penunjang proses perancangan dan pembuatan program selama penelitian berlangsung. Spesifikasi hardware dapat dilihat pada tabel 3.1.

Tabel 3. 1 Spesifikasi Perangkat Keras

Jenis	Spesifikasi
<i>Processor</i>	Intel Core i3-1115G4
RAM	8 GB DDR4
SSD	512 GB
<i>System Type</i>	64-bit <i>operating system</i>

3.7.2 Perangkat Lunak (*Software*)

Selain kebutuhan *hardware*, dibutuhkan juga *software* untuk melakukan perancangan dan pembuatan program. Spesifikasi perangkat lunak (*software*) dapat dilihat pada tabel 3.2.

Tabel 3. 2 Spesifikasi Perangkat Lunak

Nama	Keterangan
<i>Windows 11 Home</i>	Sistem operasi yang digunakan selama penelitian
<i>Visual Studio Code</i>	<i>Software</i> yang digunakan untuk menulis <i>script code</i> .

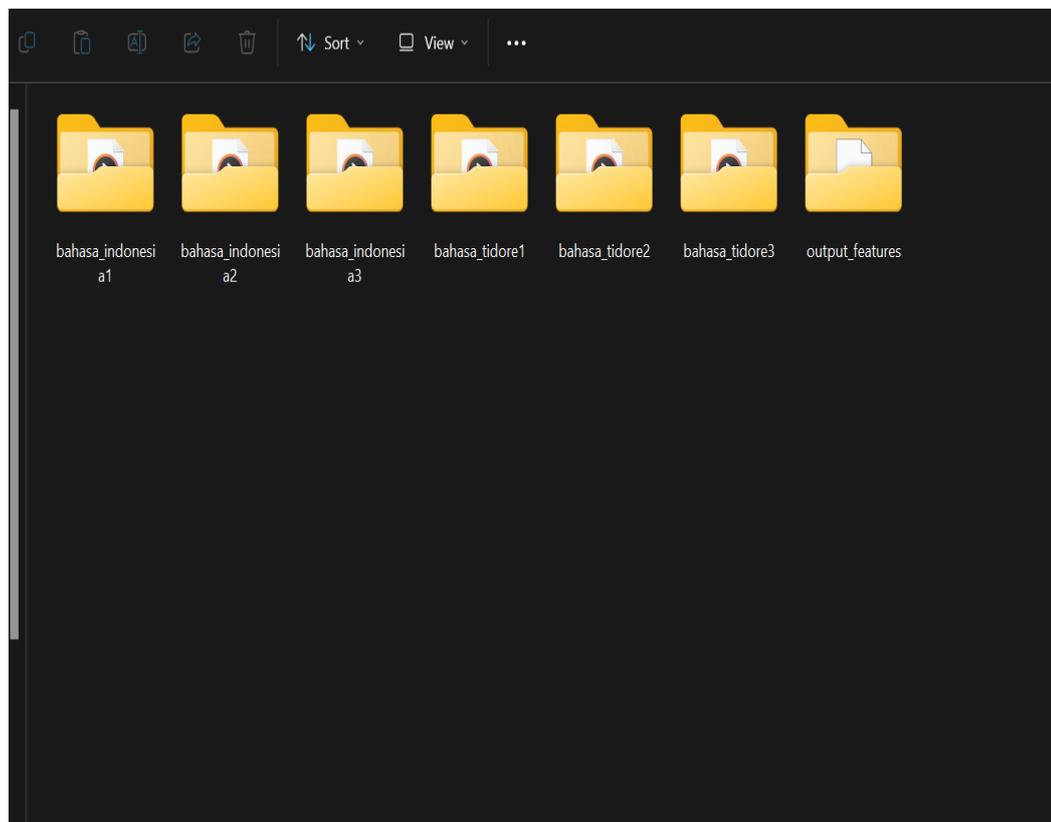
BAB IV

HASIL DAN PEMBAHASAN

Pada bab ini menjabarkan tentang hasil dan pembahasan tentang proses awal hingga akhir dalam membuat Penerapan Algoritma *K-Nearest Neighbor* (KNN) Dalam Klasifikasi Suara Berdasarkan Bahasa Daerah Tidore, data yang akan digunakan yaitu berasal dari rekaman suara yang dikumpulkan oleh penulis.

4.1. Pengumpulan Data

Data yang digunakan pada penelitian ini bersumber dari narasumber yang direkam suara menggunakan smartphone dan simpan dalam format WAV, setelah selesai merekam maka semua file disimpan dalam folder Dataset suara, dapat dilihat pada gambar 4.1.



Gambar 4. 1 Folder Penyimpanan Dataset Suara

4.2. Ekstraksi Ciri MFCC

Pada tahap ini kita akan mengekstrak fitur dari dataset yang telah dikumpulkan, dengan menggunakan *library python*, berikut gambar 4.2 *library* yang digunakan untuk ekstraksi fitur.

```

import os
import numpy as np
import librosa

```

Gambar 4. 2 Library Yang Digunakan

Setelah import library diatas, maka selanjutnya kita akan melakukan mendefinisikan direktori file yang berisi dataset suara, dengan menggunakan kode seperti pada gambar 4.3.

```

# Define directories
audio_dirs = [
    r"dataset_suara\bahasa_tidore1",
    r"dataset_suara\bahasa_indonesia1",
    r"dataset_suara\bahasa_tidore2",
    r"dataset_suara\bahasa_indonesia2",
    r"dataset_suara\bahasa_tidore3",
    r"dataset_suara\bahasa_indonesia3"
]

```

Gambar 4. 3 Mendefinisikan Direktori

Setelah mendefinisikan direktori, maka selanjutnya kita akan mengestrak dan menyimpan fitur MFCC, dengan menggunakan kode seperti pada gambar 4.4.

```

# Function to extract and save MFCC features
def save_mfcc_features(audio_dir, label_prefix):
    for file in os.listdir(audio_dir):
        if file.endswith(".wav") or file.endswith(".mp3"):
            file_path = os.path.join(audio_dir, file)
            y, sr = librosa.load(file_path)
            mfcc = librosa.feature.mfcc(y=y, sr=sr)
            output_file = f"{label_prefix}_{os.path.splitext(os.path.basename(file_path))[0]}.npy"
            output_path = os.path.join(output_dir, output_file)
            np.save(output_path, mfcc)
            print(f"Fitur MFCC untuk {os.path.basename(file_path)} telah disimpan ke {output_path}")

```

Gambar 4. 4 Ekstrak dan Menyimpan Fitur

Kode diatas digunakan untuk mengiterasi melalui setiap file dalam direktori `audio_dir`

dan memeriksa apakah file tersebut memiliki ekstensi ".wav" atau ".mp3". Jika demikian, file tersebut akan dimuat menggunakan *librosa.load* untuk mendapatkan data *audio* (*y*) dan *sampling rate* (*sr*). Kemudian, fitur MFCC diekstrak dari data audio tersebut menggunakan *librosa.feature.mfcc.*, setelah ekstraksi maka file disimpan di direktori tujuan dan di simpan dengan format .npy. berikut adalah gambar hasil running dari program dapat dilihat pada gambar 4.5.

```
Fitur MFCC untuk Jaya 11.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 11.npy
Fitur MFCC untuk Jaya 13.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 13.npy
Fitur MFCC untuk Jaya 15.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 15.npy
Fitur MFCC untuk Jaya 17.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 17.npy
Fitur MFCC untuk Jaya 19.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 19.npy
Fitur MFCC untuk Jaya 21.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 21.npy
Fitur MFCC untuk Jaya 23.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 23.npy
Fitur MFCC untuk Jaya 25.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 25.npy
Fitur MFCC untuk Jaya 27.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 27.npy
Fitur MFCC untuk Jaya 29.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 29.npy
Fitur MFCC untuk Jaya 3.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 3.npy
Fitur MFCC untuk Jaya 31.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 31.npy
Fitur MFCC untuk Jaya 33.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 33.npy
Fitur MFCC untuk Jaya 35.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 35.npy
Fitur MFCC untuk Jaya 37.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 37.npy
Fitur MFCC untuk Jaya 39.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 39.npy
Fitur MFCC untuk Jaya 41.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 41.npy
Fitur MFCC untuk Jaya 43.wav telah disimpan ke dataset_suara\output_features\bahasa_tidore1_Jaya 43.npy
```

Gambar 4. 5 Hasil Running

4.3. Membuat Dan Melatih Model KNN

Pada Tahapan ini kita akan melakukan training dari model knn yang telah dibuat, sebelum itu kita akan membagi Data menjadi set pelatihan dan Pengujian. Dapat dilihat pada gambar 4.6.

```
x_train, x_test, y_train, y_test = train_test_split(data_padded, labels, test_size=0.2, random_state=42)
```

Gambar 4. 6 Membagi Dataset

Kode diatas digunakan untuk membagi data menjadi set pelatihan ('x_train','y_train') dan set pengujian ('x_test','y_test'). Parameter 'test_size=0.2' menunjukkan bahwa 20% dari data akan digunakan untuk pengujian, sementara 80% sisanya digunakan untuk pelatihan dan parameter 'random_state=42' untuk memastikan pembagian data yang konsisten setiap

kali kode dijalankan, setelah itu maka kita akan mengubah bentuk data pelatihan dapat dilihat pada gambar 4.7.

```
X_train = np.concatenate([x.reshape(1, -1) for x in X_train], axis=0)
```

Gambar 4. 7 Mengubah Bentuk Data Pelatihan

Kode diatas berfungsi untuk mengubah array yang misalnya bukan dua dimensi menjadi array dua dimensi, dimana setiap contoh data diubah menjadi satu baris. Ini dilakukan dengan mengubah bentuk setiap contoh mfcc dari bentuk dari 3d menjadi 2d, Setelah itu maka kita membuat dan melatih model knn, dapat dilihat pada gambar 4.8.

```
knn = KNeighborsClassifier(n_neighbors=5)  
knn.fit(X_train, y_train)
```

Gambar 4. 8 Membuat Model

Kode *KNeighbors Classifier(n_neighbors=5)* adalah sebuah instance dari model knn yang berarti model akan menggunakan 5 tetangga terdekat untuk membuat prediksi. Kemudian, model tersebut dilatih menggunakan data pelatihan dengan menggunakan kode `knn.fit(X_train, y_train)`. Pada Langkah ini model belajar pola dalam data pelatihan yang dapat membedakan antara label yang berbeda, selanjutnya kita akan mencoba 3 tetangga terdekat. Dapat dilihat pada gambar 4.9.

```
knn = KNeighborsClassifier(n_neighbors=3)  
knn.fit(X_train, y_train)
```

Gambar 4. 9 Menggunakan 3 Tetangga Terdekat

4.4. Evaluasi Model K-Nearest Neighbor

Pada Tahap ini kita akan mencoba melakukan pengujian dan evaluasi performa model yang telah dibuat bisa dilihat pada gambar 4.9.

```
predictions = knn.predict(X_test_reshaped)
accuracy = np.mean(predictions == y_test)
print(f"Akurasi model KNN: {accuracy * 100:.2f}%")
```

Gambar 4. 10 Memprediksi dan Evaluasi Data

Kode diatas berfungsi membuat prediksi pada data pengujian yang telah diubah bentuknya (`X_test_reshaped`).setelah itu hasil prediksi dimpan dalam variable `predictions` setelah itu akurasi model dihitung dengan membandingkan prediksi yang dibuat dengan label sebenarnya (`y_test`). `np.mean(predictions == y_test)`, menghitung proporsi prediksi yang benar. Akurasi ini Kemudian dikalikan dengan 100 untuk mendapatkan persentase dan dicetak dengan format dua angka decimal. Setelah itu maka kita akan membuat laporan klasifikasi dapat dilihat pada gambar 4.11.

```
report = classification_report(y_test, predictions)
print("\nLaporan Klasifikasi:")
print(report)
```

Gambar 4. 11 Laporan Klasifikasi

Kode diatas digunakan untuk menampilkan laporan klasifikasi dari model, laporan ini mencakup metrik evaluasi seperti presisi, *recall*, *f1-score*, dan *support* untuk setiap kelas. Selanjutnya kita akan mencoba membuat laporan klasifikasi menggunakan 5 tetangga terdekat Hasilnya dicetak ke konsol dapat dilihat pada gambar 4.12.

```

Laporan Klasifikasi:
      precision    recall  f1-score   support

 indonesia      0.66      0.62      0.64         96
   tidore       0.67      0.70      0.69        105

 accuracy                    0.67        201
 macro avg      0.67      0.66      0.67        201
 weighted avg   0.67      0.67      0.67        201

```

Gambar 4. 12 Hasil Laporan Klasifikasi

Gambar diatas merupakan hasil laporan klasifikasi yang berisi hasil evaluasi menggunakan 5 tetangga terdekat yang menunjukkan seperti *precision*, *recall*, *f1-score*, *support* serta akurasi dari model. Selanjutnya akan menampilkan menggunakan 3 tetangga terdekat, dapat dilihat pada gambar 4.13.

```

Laporan Klasifikasi:
      precision    recall  f1-score   support

 indonesia      0.78      0.72      0.75         96
   tidore       0.76      0.81      0.78        105

 accuracy                    0.77        201
 macro avg      0.77      0.76      0.76        201
 weighted avg   0.77      0.77      0.77        201

```

Gambar 4. 13 Hasil Laporan Klasifikasi

Gambar diatas merupakan hasil laporan klasifikasi yang berisi hasil evaluasi menggunakan 3 tetangga terdekat yang menunjukkan seperti *precision*, *recall*, *f1-score*, *support* serta akurasi dari model. Setelah kita mendapatkan hasil laporan klasifikasi dari 3

model yang digunakan maka selanjutnya kita akan membuat tabel dari laporan klasifikasi dapat dilihat pada tabel 4.1

Tabel 4. 1 Laporan Klasifikasi

Jumlah Tetangga	Akurasi	Precision	Recall	F1-Score
1	97%	98%	96%	97%
3	77%	76%	81%	78%
5	67%	67%	70%	69%

4.5. White Box

Pada bagian ini dilakukan pengujian menggunakan metode *white box*. Bagian yang akan diuji adalah ekstraksi dan penyimpanan fitur. Berikut merupakan Langkah-langkah pengujian ekstraksi dan penyimpanan fitur.

1. Menentukan flowchart pengujian, sebelum itu kita akan menunjukkan listing program yang akan diuji menggunakan metode *white box* untuk lebih jelasnya dapat dilihat pada gambar 4.14.

```
import os
import numpy as np
import librosa
from sklearn.neighbors import KNeighborsClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report, confusion_matrix
from joblib import dump, load
import matplotlib.pyplot as plt
import seaborn as sns

# Define directories
audio_dirs = [
    r"dataset_suara\bahasa_tidore1",
    r"dataset_suara\bahasa_indonesia1",
    r"dataset_suara\bahasa_tidore2",
    r"dataset_suara\bahasa_indonesia2",
    r"dataset_suara\bahasa_tidore3",
    r"dataset_suara\bahasa_indonesia3"
]

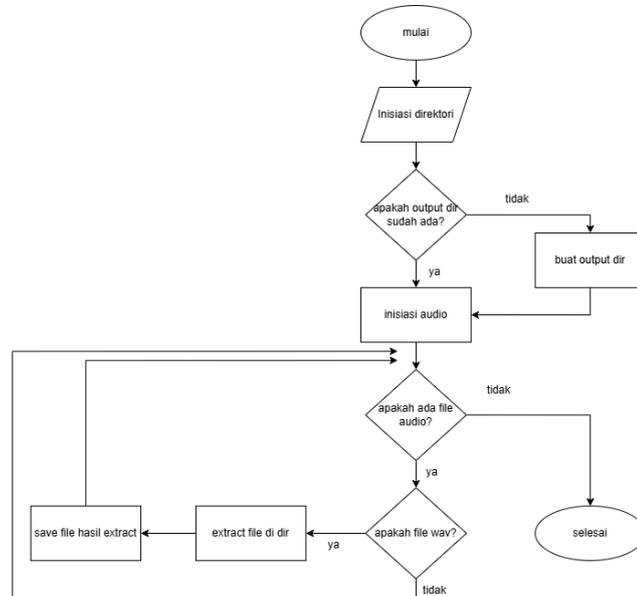
output_dir = r"dataset_suara\output_features"
if not os.path.exists(output_dir):
    os.makedirs(output_dir)

# Function to extract and save MFCC features
def save_mfcc_features(audio_dir, label_prefix):
    for file in os.listdir(audio_dir):
        if file.endswith(".wav") or file.endswith(".mp3"):
            file_path = os.path.join(audio_dir, file)
            y, sr = librosa.load(file_path)
            mfcc = librosa.feature.mfcc(y=y, sr=sr)
            output_file = f"{label_prefix}_{os.path.splitext(os.path.basename(file_path))[0]}.npy"
            output_path = os.path.join(output_dir, output_file)
            np.save(output_path, mfcc)
            print(f"Fitur MFCC untuk {os.path.basename(file_path)} telah disimpan ke {output_path}")

# Iterate through directories and save MFCC features
for i, audio_dir in enumerate(audio_dirs):
    label_prefix = f"bahasa_tidore{[i+1] if i % 2 == 0 else f"bahasa_indonesia{[i+1]"
    save_mfcc_features(audio_dir, label_prefix)
```

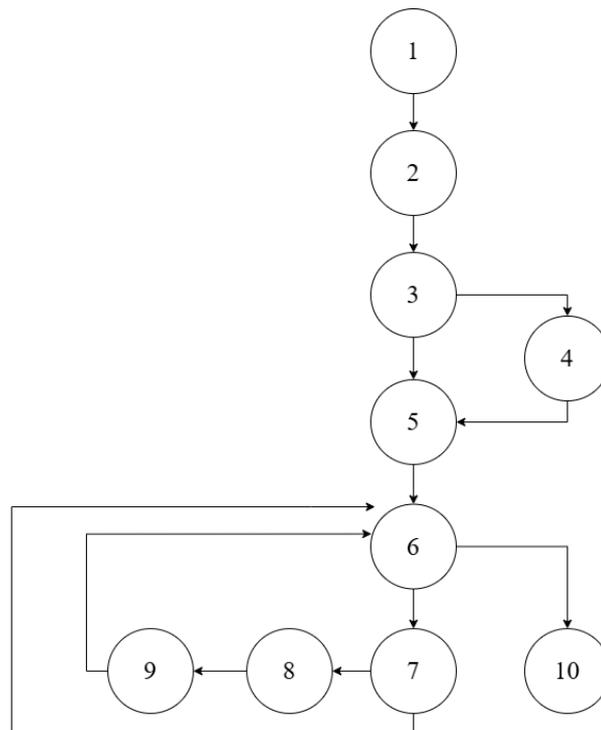
Gambar 4. 14 Listing Program Ekstraksi dan Penyimpanan Fitur

Kode diatas digunakan untuk ekstraksi dan penyimpanan fitur, setelah itu akan kita uji menggunakan *white box*, lebih jelasnya dapat dilihat pada gambar flowchart 4.15.



Gambar 4. 15 Ekstraksi Fitur

2. Membuat *Flowgraph* pengujian, untuk lebih jelasnya dapat dilihat pada gambar 4.15.



Gambar 4. 16 *Flowgraph* Pengujian

3. Setelah menentukan *flowgraph* maka selanjutnya mencari *Cyclomatic Complexity*.

$$CC = (E - N) + 2$$

$$= (12 - 10) + 2$$

$$= 2 + 2 = 4$$

4. Menentukan *Independent part*

Part 1 = 1-2-3-5-6-7-8-9-10

Part 2 = 1-2-3-4-5-6-7-8-9-10

Part 3 = 1-2-3-5-6-7-10

Part 4 = 1-2-3-5-6-7

4.6. Pembahasan

Pada tahap ini membandingkan penelitian yang terdahulu dengan penelitian yang dilakukan, penelitian yang dilakukan oleh Yovi Pratama yang berjudul *Klasifikasi Penyakit Gagal Jantung Menggunakan Algoritma K-Nearest Neighbor*, menghasilkan akurasi tertinggi sebesar 70,65%, dan dari penelitian yang dilakukan untuk klasifikasi suara berdasarkan Bahasa daerah Tidore menghasilkan nilai akurasi 76,62%. Sehingga algoritma KNN pada klasifikasi suara terbilang cukup baik dalam mengklasifikasikan Suara

BAB V

PENUTUP

5.1. Kesimpulan

Dari penelitian yang dilakukan, terdapat beberapa kesimpulan yang dapat diambil dari penelitian ini:

1. Algoritma K-Nearest Neighbor dapat digunakan untuk mengklasifikasikan suara kedalam dua kelas (Bahasa tidore dan Bahasa Indonesia) dengan cukup baik.
2. Akurasi model tergantung pada jumlah tetangga (neighbors) yang digunakan dalam KNN contohnya jika hanya menggunakan 1 neighbors maka akurasi yang dihasilkan sebesar 97,01%, dan jika menggunakan 3 neighbors menghasilkan akurasi sebesar 76,62%. Serta pada 5 neighbors 66,67%.
3. Penggunaan laporan klasifikasi pada model knn menunjukkan pada Bahasa tidore menggunakan pada 5 *Neighbor precssion* sebesar 67%, *recall* sebesar 62% serta *f1-score* 69%, 3 *Neighbor precssion* sebesar 76%, *recall* sebesar 81% serta *f1-score* 78%, dan 1 *Neighbor precision* sebesar 98%, *recall* sebesar 96% serta *f1-score* 97% oleh karena itu dapat disimpulkan bahwa performa terbesar pada model knn menggunakan 1 neighbors dan terkecil pada 5 neighbors, tetapi menggunakan 3 neighbors memberikan keseimbangan yang lebih antara bias dan varians sehingga menghasilkan kinerja yang lebih konsisten.

5.2. Saran

Adapun saran yang dapat diberikan untuk mengembangkan penelitian ini adalah:

1. Penelitian selanjutnya dapat menggunakan lebih banyak kelas Bahasa lainnya untuk meningkatkan akurasi.

2. Untuk penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih banyak dan berbeda, hal ini tidak hanya dapat meningkatkan akurasi tetapi juga membuat model dapat mengklasifikasikan kelas Bahasa menjadi lebih baik.
3. Menggunakan algoritma yang berbeda dengan penelitian ini agar dapat mengetahui hasil dan kualitas dari masing-masing metode

DAFTAR PUSTAKA

- Andreas, (2019). Klasifikasi Data Nasabah Yang Berpotensi Membuka Simpanan Deposito Menggunakan Algoritma *Backpropagation*. 13–15.
- Barus, J. I., & Sastromiharjo, A. 2019. Intervensi Bahasa Daerah Terhadap Bahasa Indonesia (Studi Campur Kode Bunyi Bahasa Karo Dalam Percakapan Bahasa Indonesia). Seminar Internasional, ISSN:2655-1780, Vol. 9 Issue.8 Agustus 2019
- Budiman Ilham, Sopyan Saori, Ramdan Nurul Anwar, Fitriani & Muhammad Yuga Pangestu 2021. Analisis Pengendalian Mutu Di Bidang Industri Makanan. Universitas Muhammadiyah Sukabumi, ISSN:2722-9475, Vol. 1 Issue.10 Maret 2022
- Djumadil, S. M. S. 2021. Frasa Verba Bahasa Tidore Dalam Tata Bahasa Generatif. Universitas Khairun, Vol. 9 Issue. 2 Juli 2021
- Duwila, E. 2022. Bahasa Ternate dan Bahasa Tidore (Kajian Linguistik Historis Komparatif dan Dialektologi). , ISSN:1693-1041, Vol. 19 Issue. 2 Juni 2021
- Ma'arif, A. (2020). Buku Ajar Pemrograman Lanjut Bahasa Pemrograman Python. Universitas Ahmad Dahlan, 7-8.
- Manullang, R. A., Sianturi, F. A., Penerapan, [, *Neighbor*, A. K.-N., Memprediksi, U., & Mahasiswa, K. (2021). JIKOMSI [Jurnal Ilmu Komputer dan Sistem Informasi] Penerapan Algoritma *K-Nearest Neighbour* Untuk Memprediksi Kelulusan Mahasiswa. Jurnal Ilmu Komputer Dan Sistem Informasi, 4(2), 42–50.
- Momole, G. M. 2022. Perbandingan Naïve Bayes dan Random Forest Dalam Klasifikasi Bahasa Daerah. JATISI (Jurnal Teknik Informatika Dan Sistem Informasi), ISSN:2407-4322, Vol. 9 Issue.2 Juni 2022
- Nawasta, R. A., Cahyana, N. H., Heriyanto 2023. Ekstraksi *Ciri Mel-Frequency Cepstral Coefficient* Menggunakan *K-Nearest Neighbor* Untuk Deteksi Emosi Berdasarkan Intonasi, ISSN:1829-667X, Vol. 20 Issue.1 Februari 2023
- Nurrahmah, F. 2023. Penguatan Bahasa Indonesia Sebagai Identitas Nasional Melalui Poster Interaktif Pada Instagram, ISSN:2986-5808, Vol. 1 Issue.1 February 2023
- Pratama, Y., Prayitno, A., Azrian, D., Aini, N., Rizki, Y., & Rasywir, E. 2022. Klasifikasi Penyakit Gagal Jantung Menggunakan Algoritma *K-Nearest Neighbor*. *Bulletin of Computer Science Research*, ISSN:2774-3659, Vol. Issue.1 Desember 2022

- Prayugo, M. B. D. (2023). *Implementasi metode jaringan saraf tiruan backpropagation pada pengenalan suara manusia*. 29-30.
- Salsabilla, A. P., Pujianti, I. P., & Permatasari, H. 2023. Pengujian Sistem Informasi Pada Aplikasi Hugo Car Wash, ISSN:2962-1968, Vol. 8 *Issue*.25 Juli 2023
- Saputra, J., Sa, Y., Yoga Pudya Ardhana, V., & Afriansyah, M. (2023). RESOLUSI : Rekayasa Teknik Informatika dan Informasi Klasifikasi Kematangan Buah Alpukat Mentega Menggunakan Metode K-Nearest Neighbor Berdasarkan Warna Kulit Buah, ISSN:2745-7966, Vol. 3 Issue 5 Mei 2023
- Syam, M., Raharjo, I. J., & Patmasari, R. (2019). Identifikasi Asal Daerah Berdasarkan Suara Manusia Dengan Metode Linier Predictive Coding (Lpc), ISSN:2355-9365, Vol. 6 Issue 3 Desember 2019
- Tempola, F., Rosihan, R., & Adawiyah, R. 2021. *Holdout Validation for Comparison Classification Naïve Bayes and KNN of Recipient Kartu Indonesia Pintar*. IOP Conference Series: Materials Science and Engineering, ISSN:1757-8981, Vol. 1125 Issue. 2021
- Ula, M., Zulhusna, R., Putra Fhonna, R., & Pratama, A. 2022. Penerapan Model Klasifikasi K-Nearest Neighbor Dalam Pencarian Kesesuaian Pekerjaan, ISSN:2442-9562, Vol. 6 Issue 1 Mei 2022



KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS KHAIRUN
FAKULTAS TEKNIK PROGRAM STUDI INFORMATIKA
Kampus III Universitas Khairun, Kelurahan Jati Kota Ternate Selatan
<http://if.unkhair.ac.id>, <http://unkhair.ac.id> Group FB: if.unkhair

KARTU BIMBINGAN HASIL

Nama Mahasiswa : Muh. Irfan Mansyur
NIM : 07352011055
Dosen Pembimbing II : Muhammad Fhadli, S.Kom., M.Sc.
Judul : Penerapan Algoritma *K-Nearest Neighbor* (KNN) Dalam Klasifikasi Suara Berdasarkan Bahasa Daerah Tidore

NO	Tanggal	Uraian	Paraf
		- Perbaiki format	
		- Perbaiki bab 3	
		- Perbaiki bab 4	
		- belajar coding	
		Acc	



UNIVERSITAS KHAIRUN
FAKULTAS TEKNIK
PROGRAM STUDI INFORMATIKA

DAFTAR PERBAIKAN SEMINAR HASIL SKRIPSI

Dengan ini dinyatakan bahwa pada

Hari / tanggal : RABU, 03 JULI 2024
Pukul : 11:00 - 13:00
Tempat : RUANG SIDANG

telah berlangsung Seminar Hasil Skripsi dengan Peserta:

Nama Mahasiswa : MUH. IRFAN MANSYUR
NPM : 07352011055
Judul : PENERAPAN ALGORITMA K-NEAREST NEIGHBOR (KNN) DALAM
KLASIFIKASI SUARA BERDASARKAN BAHASA DAERAH TIDORE

dinyatakan HARUS menyelesaikan perbaikan, yaitu:

- Ikuti format penulisan
- Abstrak harus ada di laporan hasil
- Tambahkan teori ttg:
1. Bhs Indonesia; 2. Bahasa Daerah Tidore
(ambil sumber dari Kantor Bahasa Maluku)
- Jumlah K ada 3 (tambahkan Bahasa Lainnya)
- Jelaskan proses konversi dari data suara analog ke data suara digital
- Buat program euclidian distance dengan C++
- Buat program biodata dinamis dengan memanfaatkan pengulangan dan dibuat dengan bhs pemrograman Phyton
- Input data suara bisa secara realtime

Dosen Penguji I,

KUSIHAN, S.T., M.Cs.
NIP. 197607192010121001



UNIVERSITAS KHAIRUN
FAKULTAS TEKNIK
PROGRAM STUDI INFORMATIKA

DAFTAR PERBAIKAN SEMINAR HASIL SKRIPSI

Dengan ini dinyatakan bahwa pada

Hari / tanggal : RABU, 03 JULI 2024
Pukul : 11:00 - 13:00
Tempat : RUANG SIDANG

telah berlangsung Seminar Hasil Skripsi dengan Peserta:

Nama Mahasiswa : MUH. IRFAN MANSYUR
NPM : 07352011055
Judul : PENERAPAN ALGORITMA K-NEAREST NEIGHBOR (KNN) DALAM
KLASIFIKASI SUARA BERDASARKAN BAHASA DAERAH TIDORE

dinyatakan HARUS menyelesaikan perbaikan, yaitu:

- | | |
|--|--|
| 1. bab 3 tidak ada contoh perhitungan. buatlah sesuai | |
| langkah2 metodenya | |
| 2. di teori metodenya hanya sedikit tidak ada langkah2 | |
| metodenya | |
| 3. hasil belum jelas, belum jelas cara merubah suara | |
| dan masuk dalam metodenya | |
| 4. data juga blum jelas karena ini terkait suara | |
| | |
| | |

AGG, 11/07/2024

Dosen Penguji III,


SYARIFUDDIN N. KAPITA, S.Pd., M.Si.
NIP. 199103122024211001



**UNIVERSITAS KHAIRUN
FAKULTAS TEKNIK
PROGRAM STUDI INFORMATIKA**

DAFTAR PERBAIKAN UJIAN SKRIPSI/TUTUP

Dengan ini dinyatakan bahwa pada

Hari / tanggal : SENIN, 29 JULI 2024
Pukul : 10:30 - 12:00
Tempat : RUANG PRODI

telah berlangsung Ujian Skripsi/Tutup dengan Peserta:

Nama Mahasiswa : MUH. IRFAN MANSYUR
NPM : 07352011055
Judul : PENERAPAN ALGORITMA K-NEAREST NEIGHBOR (KNN) DALAM
KLASIFIKASI SUARA BERDASARKAN BAHASA DAERAH TIDORE

dinyatakan HARUS menyelesaikan perbaikan, yaitu:

Selesaikan secepatnya semua masukan dan koreksi dari penguji

*Acc to
Clap*

Dosen Pembimbing I,

Clap

MUHAMMAD SABRI AHMAD, S.Kom., M.Kom.
NIP. 198905012019031013



UNIVERSITAS KHAIRUN
FAKULTAS TEKNIK
PROGRAM STUDI INFORMATIKA

DAFTAR PERBAIKAN UJIAN SKRIPSI/TUTUP

Dengan ini dinyatakan bahwa pada

Hari / tanggal : SENIN, 29 JULI 2024

Pukul : 10:30 - 12:00

Tempat : RUANG PRODI

telah berlangsung Ujian Skripsi/Tutup dengan Peserta:

Nama Mahasiswa : MUH. IRFAN MANSYUR

NPM : 07352011055

Judul : PENERAPAN ALGORITMA K-NEAREST NEIGHBOR (KNN) DALAM
KLASIFIKASI SUARA BERDASARKAN BAHASA DAERAH TIDORE

dinyatakan HARUS menyelesaikan perbaikan, yaitu:

- Ikuti format penulisan
- Pelajari dengan lengkap (akan ditanya saat asistensi):
 1. Struktur dasar algoritma
 2. Sequential
 3. Pilihan/Selection
 4. Perulangan/Looping

Aku
sudah
selesai

Dosen Penguji I,

ROSIHAN, S.T., M.Cs.
NIP. 197607192010121001



UNIVERSITAS KHAIRUN
FAKULTAS TEKNIK
PROGRAM STUDI INFORMATIKA

DAFTAR PERBAIKAN UJIAN SKRIPSI/TUTUP

Dengan ini dinyatakan bahwa pada

Hari / tanggal : SENIN, 29 JULI 2024
Pukul : 10:30 - 12:00
Tempat : RUANG PRODI

telah berlangsung Ujian Skripsi/Tutup dengan Peserta:

Nama Mahasiswa : MUH. IRFAN MANSYUR
NPM : 07352011055
Judul : PENERAPAN ALGORITMA K-NEAREST NEIGHBOR (KNN) DALAM
KLASIFIKASI SUARA BERDASARKAN BAHASA DAERAH TIDORE

dinyatakan HARUS menyelesaikan perbaikan, yaitu:

Pelajari Dasar Informatika

ACC
Ridha

31-07-24

Dosen Penguji II,

Ridha

Dr. MUHAMMAD RIDHA ALBAAR, S.Kom., M.Kom.
NIP. 198504232008031001



**UNIVERSITAS KHAIRUN
FAKULTAS TEKNIK
PROGRAM STUDI INFORMATIKA**

DAFTAR PERBAIKAN UJIAN SKRIPSI/TUTUP

Dengan ini dinyatakan bahwa pada

Hari / tanggal : SENIN, 29 JULI 2024
Pukul : 10:30 - 12:00
Tempat : RUANG PRODI

telah berlangsung Ujian Skripsi/Tutup dengan Peserta:

Nama Mahasiswa : MUH. IRFAN MANSYUR
NPM : 07352011055
Judul : PENERAPAN ALGORITMA K-NEAREST NEIGHBOR (KNN) DALAM
KLASIFIKASI SUARA BERDASARKAN BAHASA DAERAH TIDORE

dinyatakan HARUS menyelesaikan perbaikan, yaitu:

pelajari kembali penjabaran rumus metode yang diangkat sampai paham dan rumus2 serupa

Ace, 12/08/2024

Dosen Penguji III,

SYARIFUDDIN N. KAPITA, S.Pd., M.Si.
NIP. 199103122024211001