

ANALISA PENGENALAN EMOSI SUARA MENGGUNAKAN METODE *LEARNING VECTOR QUANTIZATION*

SOUND EMOTION RECOGNITION ANALYSIS USING LEARNING VECTOR QUANTIZATION METHOD

Ovy Meliyandika¹, I made Budi Suksmadana², Giri Wahyu Wiriasto³

^{1,2,3}Jurusan Teknik Elektro Universitas Mataram

²imadebudisuksmadana@unram.ac.id, ³giriwahyuwiriasto@unram.ac.id,

ABSTRAK

Mengenali suara merupakan hal yang mudah, dengan cara mendengarkan dengan seksama dan manusia mempunyai kecerdasan dalam mengenali pola suara. Berbeda dengan komputer, proses pengenalan suara merupakan proses yang sulit, hal ini dikarenakan komputer memerlukan suatu mekanisme yang standar dan logis dalam mengenali pola suara. Dengan metode *Mel Frequency Cepstral Coefficient (MFCC)* memiliki peran penting dalam menentukan karakteristik dari sebuah suara. Dengan metode *Mel Frequency Cepstral Coefficient (MFCC)* memiliki peran penting dalam menentukan karakteristik dari sebuah suara. Metode ini sering digunakan untuk verifikasi suara, pengenalan suara, deteksi emosi dari suara. Penelitian ini dilakukan untuk menganalisa pengklasifikasian emosi suara menggunakan metode *Learning Vector Quantization (LVQ)*. Pada mahasiswa Universitas Mataram didapatkan data emosi suara manusia berjumlah 80 mahasiswa. Dari 80 mahasiswa didapatkan data emosi suara bahagia dan marah. Perekaman dilakukan pada masing-masing mahasiswa yang berlangsung 4 detik sampai dengan 8 detik dengan bantuan googleform sebagai media bantu untuk pengambilan data. Media bantu googleform yang di maksud adalah peneliti memberikan google form ke masing-masing mahasiswa yang didalamnya terdapat 5 kalimat bahagia dan 5 kalimat marah sebagai contoh bacaan sehingga mendapatkan emosi suara yang di inginkan peneliti. *Model Learning Vector Quantization (LVQ)* mampu menentukan jenis emosi suara yang masuk ke dalam model, tingkat akurasi emosi suara test sebesar 87% menggunakan dataset sebanyak 380 data. Metode *confusion matrix* digunakan untuk mengevaluasi model yang sudah di buat, tingkat akurasi sebesar 87% dengan perhitungan parameter lainnya seperti precision dengan nilai rata-rata 87%, recall 87% dan f-1 score 87% dengan jumlah data uji 93 data. Sehingga bisa disimpulkan model bekerja dengan baik.

Kata kunci : Suara, Emosi, *Mel Frequency Cepstral Coefficient (MFCC)*, *Learning Vector Quantization*, *confusion Matrix*

ABSTRACT

Recognizing voices is an easy thing, by listening carefully and humans have intelligence in recognizing sound patterns. In contrast to computers, speech recognition is a difficult process, this is because computers require a standard and logical mechanism for recognizing sound patterns.

The Mel Frequency Cepstral Coefficient (MFCC) method has an important role in determining the characteristics of a sound. The Mel Frequency Cepstral Coefficient (MFCC) method has an important role in determining the characteristics of a sound. This method is often used for voice verification, voice recognition, emotion detection from voice. This research was conducted to analyze the classification of sound emotions using the Learning Vector Quantization (LVQ) method. For Mataram University students, 80 students obtained human voice emotion data. From 80 students, the emotional data of happy and angry voices were obtained. Recording is done for each student which lasts 4 seconds to 8 seconds with the help of Googleform as an auxiliary medium for data retrieval. The Googleform auxiliary media that is intended is that the researcher gives a Google form to each student in which there are 5 happy sentences and 5 angry sentences as examples of reading so as to get the sound emotion that the researcher wants. The Learning Vector Quantization (LVQ) model is able to determine the type of sound emotion included in the model, the level of accuracy of the voice emotion test is 87% using a dataset of 380 data. The confusion matrix method is used to evaluate the models that have been made, with an accuracy rate of 87% with the calculation of other parameters such as precision with an average value of 87%, recall of 87% and f-1 score of 87% with a total of 93 test data. So it can be concluded that the model works well.

Keywords : Voice, Emotion, Mel Frequency Cepstral Coefficient (MFCC), Learning Vector Quantization, Confusion Matrix

I. PENDAHULUAN

Suara manusia merupakan salah satu contoh dari sinyal analog yang berisikan informasi. Suara manusia juga unik, berbeda untuk masing-masing pribadi. Karakter suara seorang manusia ada 2 macam ada yang non akustik dan nada yang akustik non akustik contohnya adalah pulsa dan waktu sedangkan untuk akustik suara manusia terdiri dari pitch, formant, bandwidth formant, energi suara, dan durasi pengucapannya. Dari ciri akustik inilah kita dapat mengidentifikasi keadaan emosi seseorang apakah dia sedang merasa senang, marah atau sedih. Emosi merupakan suatu kondisi mental seseorang yang dapat mendorongnya untuk melakukan suatu tindakan atau berekspresi yang dapat dipicu dari dalam atau luar dirinya. Dalam kehidupan sehari-hari sangat penting untuk memahami kondisi emosional seseorang dengan emosi tertentu. Emosi juga merupakan salah satu aspek penting bagi kehidupan. Emosi seseorang dapat diketahui salah satunya dari ekspresi wajah, namun terkadang ekspresi wajah seseorang tidak sesuai dengan apa yang sedang dialaminya[1].

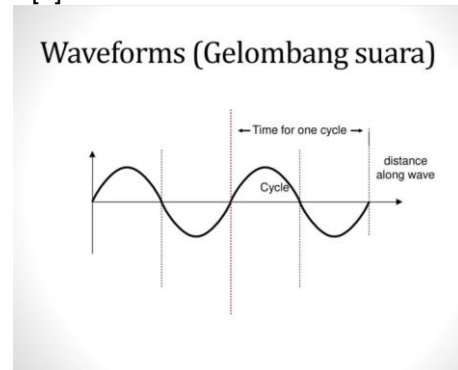
Bagi manusia mengenali suara merupakan hal yang mudah, dengan cara mendengarkan dengan seksama dan manusia mempunyai kecerdasan dalam mengenali pola suara. Berbeda dengan komputer, proses pengenalan suara merupakan proses yang sulit, hal ini dikarenakan komputer memerlukan suatu mekanisme yang standar

dan logis dalam mengenali pola suara. Dengan metode Mel Frequency Cepstral Coefficient (MFCC) memiliki peran penting dalam menentukan karakteristik dari sebuah suara. Metode ini sering digunakan untuk verifikasi suara, pengenalan suara, deteksi emosi dari suara. Untuk mentransformasi sinyal suara sebagai inputan menjadi sinyal frekuensi suara dalam format WAV diperlukan suatu metode Jaringan Syaraf Tiruan yang dapat meminimalkan masalah, yaitu menggunakan Learning Vector Quantization (LVQ) dengan bantuan software python dalam mengenali kekuatan suara (desibel) dan beberapa karakteristik suara lainnya. [2]. Pentingnya mengenali emosi secara otomatis dalam ucapan manusia telah berkembang dengan meningkatnya peran antarmuka bahasa lisan di manusia-komputer aplikasi interaksi. Dalam tulisan ini, pidato bahasa Mandarin berdasarkan emosi metode klasifikasi disajikan. Lima emosi utama manusia, termasuk kemarahan, kebosanan, kebahagiaan, netral dan kesedihan, diselidiki. Menggabungkan yang berbeda aliran fitur untuk mendapatkan hasil yang lebih akurat adalah statistik yang terkenal teknik. Untuk pengenalan emosi ucapan, kami menggabungkan 16 koefisien LPC, 12 Komponen LPCC, 16 komponen LFPC, 16 koefisien PLP, 20 MFCC komponen dan jitter sebagai fitur dasar untuk membentuk vektor fitur. Dua korpora dipekerjakan. Pengenal yang disajikan dalam makalah ini didasarkan pada tiga teknik klasifikasi: LDA, K-NN dan

HMM. Hasil menunjukkan bahwa yang dipilih fitur yang kuat dan efektif untuk pengenalan emosi dalam valensi dan dimensi gairah dari duacorpora. Menggunakan klasifikasi emosi HMM metode, akurasi rata-rata 88,7% tercapai[3]. MFCC (Mel-Frequency Cepstral Coefficient) dan DTW (Dynamic Time Warping) adalah suatu metode dalam mengolah suara, pada penelitian ini pengolahan suara yang dilakukan bertujuan untuk pengenalan jenis suara pria dan wanita. Penentuan jenis suara pria atau wanita biasanya dilakukan dalam penentuan kelompok paduan suara. MFCC merupakan metode untuk ekstraksi ciri, selain MFCC digunakan metode lain yaitu DTW yang merupakan metode pencocokan suara latih dan suara uji, teknik ini berguna untuk menghitung jarak antara dua data dengan pola-pola yang berbeda dan menghitung nilai dan jarak dari data tersebut. Pada penelitian ini dibangun suatu sistem yang dapat mengenali jenis suara pria dan wanita, Jenis suara pada pria dibagi menjadi Tenor, Bariton dan Bass, sedangkan pada wanita dibagi menjadi Sopran, Mezzosopran dan Alto. Hasil Penelitian yang didapat yaitu untuk tingkat akurasi pada wanita dengan jenis suara alto didapatkan presentase 80%, untuk tingkat akurasi jenis suara mezzosopran didapatkan 90%, untuk tingkat akurasi jenis suara sopran didapatkan 80%. Kemudian pada jenis suara pria, untuk jenis suara bass didapatkan tingkat akurasi sebesar 80%, untuk jenis suara bariton didapatkan tingkat akurasi 70%, dan untuk jenis suara tenor didapatkan tingkat akurasi sebesar 60%[4].

Emosi merupakan suatu keadaan di dalam diri seseorang, yang kelihatan dan sulit diukur. Emosi sulit deprogram, sifatnya unik, dan emosi merupakan milik kita sendiri. Manusia memiliki tempramen bawaan yang berbeda, sehingga rasa senang dan tidak senangpun berbeda. Berdasarkan hasil penelitian menunjukkan bahwa emosi sebagian besar merupakan fungsi biologis. Meskipun demikian, cara kita merespon terhadap emosi sangat dipengaruhi oleh faktor lingkungan[5]. Ada sebuah proses terciptanya sebuah suara. Hal ini diawali karena adanya sebuah sumber suara yang menghasilkan energy akustik. Energi akustik ini nanti bergerak melalui medium seperti udara, air, ataupun dinding dan akhirnya diterima oleh

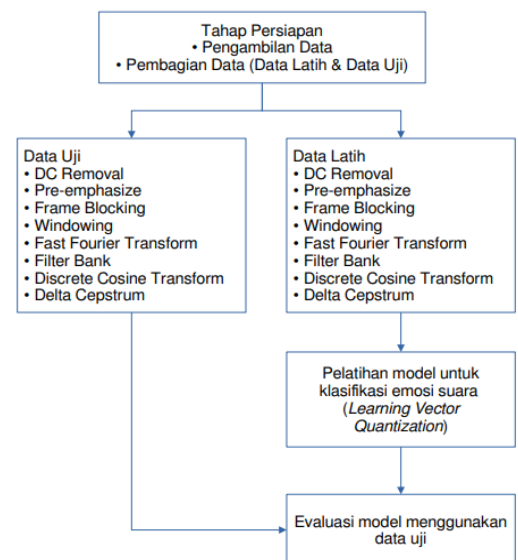
telinga hingga ke otak, lalu otak mengintrepetasikan apa yang disebut dengan suara.[6].



Gambar 1. Gelombang Suara

II. METODOLOGI

Pada penelitian ini, digunakan data emosi suara subjek hidup (manusia) secara langsung/real dan data emosi suara dari internet. Pelabelan data dilakukan oleh penulis sesuai dengan kelas yang telah didefinisikan. Kemudian menggunakan data tersebut, model akan dilatih untuk melakukan klasifikasi terhadap emosi suara menjadi beberapa kelas yaitu, emosi suara marah dan bahagia. Penelitian ini dilakukan untuk menganalisa pengklasifikasian emosi suara menggunakan metode Learning Vector Quantization (LVQ).



Gambar 2. Diagram Blok Penelitian

Pengambilan data pada penelitian ini merupakan proses perekaman emosi suara pada subjek hidup (manusia) secara langsung/real. Pengambilan data ini dilakukan di mahasiswa Teknik Elektro Fakultas Teknik Universitas Mataram.

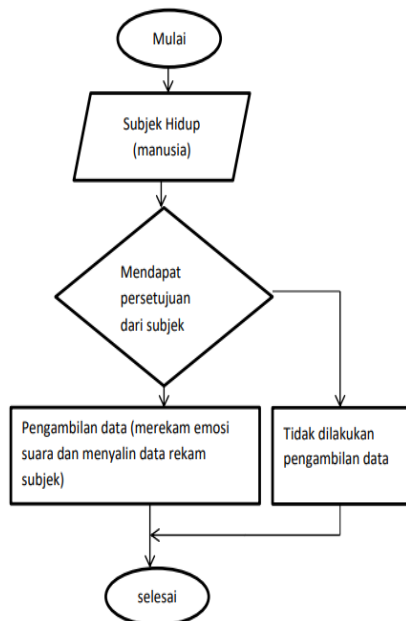


Gambar 3. Perangkat pengambilan data secara langsung dari internet

a. **Kriteria subjek**

Pengambilan data pada penelitian ini dilakukan pada subjek dengan tidak ada batasan usia. Pengambilan data didapatkan pada subjek hidup (manusia) laki-laki dan perempuan. Pada tahap ini dilakukan dengan merekam emosi suara dengan durasi 30 detik hingga 2 menit. Setelah merekam emosi suara dari subjek, peneliti akan menyalin beberapa informasi dari data subjek sebagai bahan kajian lebih lanjut.

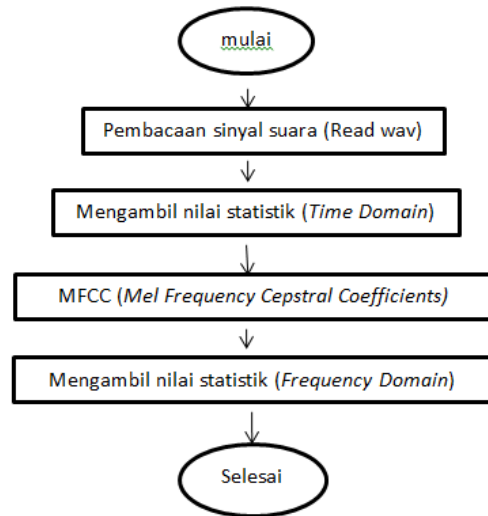
b. **Proses pengambilan data**



Gambar 4. Diagram Alir Prosedur Pengambilan Data

c. **Ekstraksi ciri**

Ekstraksi ciri data suara merupakan tahap lanjutan untuk mengolah emosi suara yang telah didapat. Ekstraksi ciri data suara dilakukan dengan menggunakan perangkat lunak python. Ekstraksi ciri data suara memiliki 4 tahapan utama yakni (1) membaca seluruh file wav dalam satu folder, (2) mengambil nilai statistik data suara pada domain waktu, (3) mentrasformasi file wav ke dalam bentuk domain frekuensi dengan menggunakan metode Mel Frequency Cepstral Coefficients, (4) mengambil nilai data statistik data suara yang dibutuhkan dalam domain frekuensi.



Gambar 5. Alur ekstraksi ciri data suara

d. **Learning Vector Quantization**

Motivasi untuk sebuah algoritma yang diterapkan pada jaringan syaraf LVQ adalah untuk menentukan unit output yang terdekat dengan vector input. Hal tersebut akan berakhir, jika x dan w berada dalam kelas yang sama, maka bobot dipindahkan ke vector input yang baru dan jika x dan w berada pada kelas yang berbeda, maka bobot akan dipindahkan dari vector input.

Algoritma dari pembelajaran LVQ secara umum adalah sebagai berikut:

```

Step 0: Inisialisasi vektor reverensi
      Inisialisasi alpha
      Inisialisasi decalpha
Step 1: Bila kondisi STOP belum terpenuhi,
      kerjakan Step 2-6
Step 2: Untuk setiap vektor training x, kerjakan
      Step 3-4
Step 3: Dapatkan j sedemikian hingga
      ||x-wj|| minimum
Step 4: Update Wj sebagai berikut:
      Jika T=Cj maka
      Wj (baru) = Wj (lama)+alpha(x-Wj (lama))
      Jika T != Cj maka
      Wj (baru) = Wj (lama) -alpha(x-Wj (lama))
Step 5: Reduksi learning rate (alpha)
      alpha = alpha - alpha*decalpha
Step 6: Tes kondisi STOP
      - dengan membatasi jumlah iterasi
      - setelah alpha mencapai nilai toleransi

```

Dengan :

- x : vector pelatihan
- x_1, x_2, \dots, x_n
- T : kategori atau kelas yang benar untuk vector pelatihan
- W_j : vector bobot untuk unit output j ($x_{1j}, x_{2j}, \dots, x_{nj}$)
- $\|x-W_j\|$: jarak Euclidean antara vector input (vector bobot) dan unit output ke j
- Alpha : Learning rate
- Decalpha : pengurangan learning rate

e. Confusion matrix

Confusion matrix adalah salah satu metode tools analitik prediktif yang menampilkan dan membandingkan nilai actual atau nilai sebenarnya dengan nilai hasil prediksi model yang dapat di gunakan untuk menghasilkan metrik evaluasi suatu model seperti accuracy, precision, recall, f-1 score.

Ada empat nilai yang dihasilkan di dalam confusion matrix, diantaranya true positive (TP), false positive (FP), false negative (FN) dan true negative (TN). Ilustrasi tabel bisa dilihat pada tabel berikut ini.

Tabel 1. Nilai confusion matrix

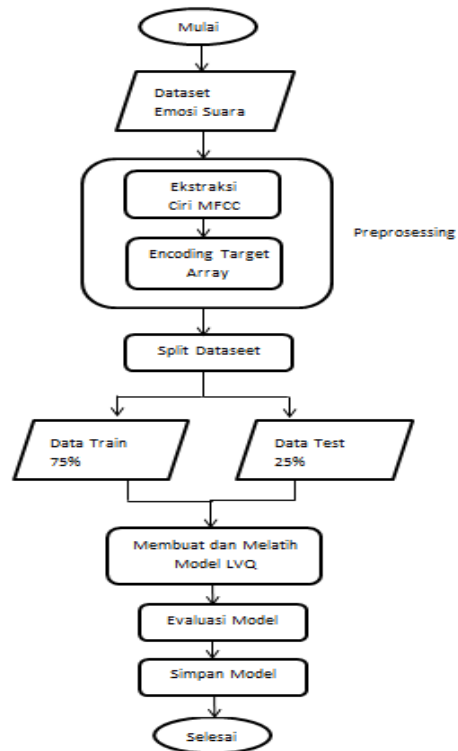
		Positive	Negative
Nilai Prediksi	Positive	TP	FP
	Negative	FN	TN

1. Instrumen penelitian

perangkat keras yang digunakan dalam penelitian ini adalah laptop Asus X441M, CPU intel Celeron tipe N4000 dan N4020 1.1 Ghz, Ram 4GB, dan sistem operasi Windows 10. Sedangkan untuk perangkat lunak yang digunakan adalah Google Colab dan bahasa pemograman phyton untuk membuat dan menguji model penelitian.

a. Diagram alir implementasi

Berikut adalah diagram alir dari implementasi penelitian tersebut.



Gambar 6. Diagram alir pengujian LVQ

III. HASIL DAN PEMBAHASAN

Pada proses pengambilan data emosi suara ini dilakukan pada mahasiswa Universitas Mataram didapatkan data emosi suara manusia berjumlah 80 mahasiswa. Dari 80 mahasiswa didapatkan data emosi suara bahagia dan marah. Perekaman dilakukan pada masing-masing mahasiswa yang berlangsung 4 detik sampai dengan 8 detik dengan bantuan googleform sebagai media bantu untuk pengambilan data. Media bantu googleform yang di maksud addalah peneliti memberikan google form ke masing-masing mahasiwa yang didalmnya terdapat 5 kalimat

bahagia dan 5 kalimat marah sebagai contoh bacaan sehingga mendapatkan emosi suara yang di inginkan peneliti.



Gambar 7. Proses perekaman emosi suara

A. Implementasi penelitian

1. Implementasi

Pada tahap ini implementasi yang digunakan berupa Google Colab dengan menggunakan bahasa pemrograman python pada browser google chrome. Ada tiga proses utama dari model yang sudah dibuat diantaranya adalah proses ekstraksi ciri yang berfungsi sebagai filter suara untuk mendapatkan nilai statistik (frequency domain). Proses selanjutnya adalah klasifikasi yang merupakan bagian inti pada penelitian ini dengan menggunakan metode LVQ yang berfungsi sebagai proses pengenalan terhadap kelas yang di tentukan dengan menggunakan data latih dan data uji. Proses terakhir merupakan proses evaluasi model untuk mengetahui seberapa baik model yang sudah dibuat dengan perhitungan persentase parameter kelas dengan membandingkan nilai actual dan prediksi model yang dapat digunakan untuk menghasilkan metrik evaluasi suatu model

a. Import dataset

Pada tahap pertama proses yang dilakukan adalah meng-upload Dataset suara marah dan suara bahagia ke Google drive yang sudah di siapkan yang nantinya akan digunakan pada Gogogle colab.



Gambar 8. Dataset suara Google drive

b. Import libraries

Tahap selanjutnya merupakan pembuatan librari digunakan untuk melakukan pemodelan dan pengujian di Google Colab.

```
!pip install pydub
!pip install sklvq
[ ] import numpy as np
import os
import glob
import pickle
import soundfile , librosa
from
sklearn.model_selection import
train_test_split
# from
sklearn.neural_network import
MLPClassifier
From sklearn.metrics import
accuracy_report
from pydub import
AudioSegment
from sklvq import GLVQ,
GMLVQ, LGMLVQ
```

c. Ekstraksi fitur MFCC

Ekstraksi ciri MFCC bertujuan untuk mendapatkan fitur-fitur berupa parameter yang memiliki tujuh tahapan. Yang pertama adalah pre-emphasis, kedua frame blocking, ketiga windowing, keempat Fast Fourier Transform (FFT), kelima Mel Frequency Wharping (MFW), keenam Discrete Cosine Transform (DCT), ketujuh Cepstral Liftering. Ketujuh proses ini memiliki kegunaan yang berbeda-beda

```
def feature_extractor(file, mfcc,
chroma, mel):
    with soundfile.SoundFile(file)
as sound_file:
        file_array =
sound_file.read(dtype="float32")

sample_rate=sound_file.samplerate
result=np.array([])
if mfcc:

mfccs=np.mean(librosa.feature.mfcc(y=fi
le_array, sr=sample_rate, n_mfcc=40).T,
axis=0)

result=np.hstack((result, mfccs))
if chroma:

stft=np.abs(librosa.stft(file_array))
```

```

chroma=np.mean(librosa.feature.chroma_s
tft(S=stft, sr=sample_rate).T,axis=0)

result=np.hstack((result, chroma))
    if mel:

mel=np.mean(librosa.feature.melspectrogr
am(file_array,
sr=sample_rate).T,axis=0)

result=np.hstack((result, mel))
    return result

```

d. Output proses ekstraksi ciri merupakan beberapa data matrix 1D pada proses ekstraksi ciri yang memiliki jumlah 93 data dari data uji dan 277 data latih pada proses pelatihan dan pengujian model.

0	03-01-03-145-01-01-	angry
1	03-01-03-142-01-01-	happy
2	03-01-03-113-01-01-	happy
3	03-01-03-103-01-01-	happy
4	03-01-05-153-01-01-	happy
5	03-01-03-131-01-01-	happy
6	03-01-03-164-01-01-	happy
7	03-01-05-50-01-01-0	angry
8	03-01-05-177-01-01-	happy
9	03-01-03-93-01-01-0	angry
10	03-01-05-24-01-01-0	happy

Gambar 9. Output proses ekstraksi ciri

e. Split data loader

Pembagian data menjadi data training dan data testing dengan rasio pembagian 75:25, dimana 75% untuk data training dan 25% untuk data testing.

```

x_train,x_test,y_train,y_test=d
ata_loader(test_size=0.25)

```

f. Konfigurasi model

Langkah selanjutnya adalah pengaturan model yang berfungsi untuk melatih model LVQ dengan menggunakan data training dan data testing.

```

classifier = LGMLVQ(

relevance_localization="class",
# Can either be "class" or
"prototypes"

```

```

distance_type="local-
adaptive-squared-euclidean",
activation_type="swish",
activation_params={"beta":
2},
solver_type="lbfgs",
)

```

g. Menghitung akurasi training

Sebelum menghitung akurasi testing dan training, langkah selanjutnya adalah menghitung akurasi data audio training. Yang dimana akurasi pada tahap training dengan rasio 75% mendapatkan nilai akurasi sebanyak 100%.

```

classifier.fit(x_train,y_train)
classifier.score(x_train,
y_train)

y_train_pred =
classifier.predict(x_train)
train_accuracy =
accuracy_score(y_train,
y_train_pred)
print(f"akurasi
train:{train_accuracy * 100}%")

```

h. Menghitung akurasi testing

Pada tahap ini dilakukan klasifikasi data training dan data testing yang bertujuan untuk mendapatkan hasil atau akurasi model klasifikasi.

```

y_pred=classifier.predict(x_tes
t)

accuracy=accuracy_score(y_true=
y_test, y_pred=y_pred)
print(f"akurasi: {accuracy *
100}%")

```

i. Hasil classification report dari confusion matrix

$$\begin{aligned}
 &= \frac{81}{93} \\
 &= 0.87 * 100\% \\
 &= 87\%
 \end{aligned}$$

Berfungsi untuk menampilkan hasil dari classification report dari Confusion Matrix.

```
print(classification_report(y_test, y_pred))
```

Baris kode ini menggunakan fungsi classification_report dari modul sklearn.metrics untuk mengevaluasi kinerja klasifikasi LGMLVQ pada data pengujian. Ringkasan tersebut termasuk precision, recall, dan skor F1 untuk setiap kelas, serta akurasi secara keseluruhan dan skor F1 rata-rata makro.

- Data uji marah dan bahagia

$$\begin{aligned}
 - \text{Precision} &= \frac{TP}{TP+FP} \\
 &= \frac{39}{39+5} = 0.88 \\
 &= 0.88 * 100\% \\
 &= 88\%
 \end{aligned}$$

$$\begin{aligned}
 - \text{Recall} &= \frac{TP}{TP+FN} \\
 &= \frac{39}{39+7} = 0.84 \\
 &= 0.84 * 100\% \\
 &= 84\%
 \end{aligned}$$

$$\begin{aligned}
 - \text{F1 - score} &= \frac{2*(recall*precision)}{recall+precision} \\
 &= \frac{2 * (0,84 * 0,88)}{0,84 + 0,88} \\
 &= \frac{1,4784}{1,72} \\
 &= 0.85 \\
 &* 100\% \\
 &= 85\%
 \end{aligned}$$

$$\begin{aligned}
 - \text{Accuracy} &= \frac{TP+TN}{TP+FP+FN+TN} \\
 &= \frac{39+42}{39+5+7+42}
 \end{aligned}$$

Tabel 2. Nilai precision, recall, f-1 score, support dan accuracy

	Precision	Recall	f1 score	Support
Marah	0.85	0.89	0.87	44
Bahagia	0.89	0.86	0.88	49
Accuracy			0.87	93
Makro Rata-rata	0.87	0.87	0.87	93
Perhitungan Rata-rata Nilai	0.87	0.87	0.87	93

j. Confusion matrix

Untuk melihat hasil akurasi model yang telah dibuat sehingga menampilkan dan membandingkan nilai actual atau nilai sebenarnya dengan nilai hasil prediksi model dengan nilai yang dihasilkan di dalam tabel matrix, diantaranya True Positive (TP), False Positif ((FP), False Negatif (FN) dan True Negatif (TN).

```

From sklearn.metrics import
confusion_matrix
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
cm2=confusion_matrix(y_test,y_pred
)
cm_df=pd.DataFrame(cm2,
index=['Angry','Happy'],
columns= ['Angry','Happy'])
plt.figure(figsize=(7,5))
sns.heatmap(cm_df,annot=True,
cmap='Blues', fmt='g')
plt.title('Confusion Matrix')
plt.ylabel('Label Sebenarnya')
plt.xlabel('Label Prediksi')
plt.show()

```

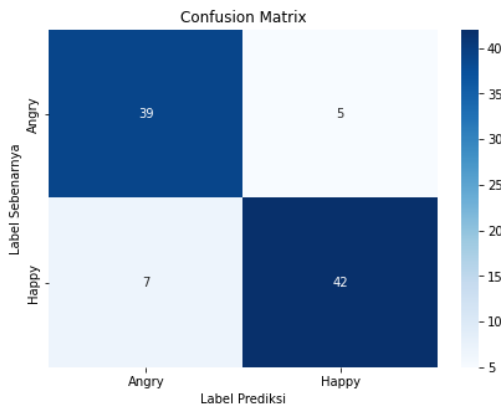

Evaluasi model dilakukan untuk melihat seberapa baik performa model yang dibuat terhadap parameter uji atau kelas yang telah ditentukan. Pada tahap ini proses evaluasi model dibantu dengan menggunakan metode confusion matrix untuk menjelaskan berapa jumlah data yang benar dan data yang salah pada saat proses pengklasifikasian oleh model.

Tabel 3. Keterangan confusion matrix

N = 93	Aktual: Positif (1)	Aktual: Negatif (0)
Marah Prediksi: Positif (1)	True Positive = 39	False Positive = 5
Bahagia prediksi: Negaif (0)	False Negative = 7	True Negative = 42
	46	47

k. Menampilkan hasil confusion matrix

Pada gambar di bawah merupakan visualisasi dari output confusion matrix yang menampilkan 2 parameter uji yaitu suara bahagia dan suara marah dengan masing-masing nilai yang didapatkan dalam proses evaluasi model.



Gambar 10. Confusion matrix hasil prediksi pada data uji

2. Hasil Uji Coba Model

Uji coba dilakukan menggunakan 25% data dari 370 keseluruhan data suara, dan digunakan 75% data latih yang telah digunakan sebelumnya untuk melatih model yang sudah dibuat. Model yang sudah dibuat dilatih dan diuji bersifat Unlimited tanpa batas perulangan sampai menemukan hasil tertinggi dari model klasifikasi.

a. Hasil Penelitian

Dari hasil training dan testing pada model LVQ, didapatkan akurasi tertinggi data train sebesar 100% dan akurasi model data test sebesar 87,24%

```

y_pred=classifier.predict(x_test)
accuracy=accuracy_score(y_true=y_
test, y_pred=y_pred)
print(f"akurasi: {accuracy *
100}%")
akurasi train:100.0%

accuracy=accuracy_score(y_true=y_
test, y_pred=y_pred)
print(f"akurasi: {accuracy *
100}%")
akurasi test: 87.24731182795699%

```

Dengan menggunakan data audio keseluruhan sebanyak 370 data digunakan 277 data sebagai data pelatihan model. Dalam proses pelatihan model bekerja dengan baik sehingga mendapatkan nilai akurasi data latih sebanyak 100% dengan jumlah data sebanyak 277 data diantaranya 138,5 data suara marah dan 138,5 data suara bahagia.

Data yang digunakan untuk proses pengujian model sebanyak 93 data diantaranya 46.5 data suara marah dan 46.5 data suara bahagia dan mendapatkan akurasi data uji sebanyak 89%. Dengan hasil akurasi data uji yang didapatkan maka hasil klasifikasi dikatakan baik.

3. Tabel Hasil Pengujian Data

Pengujian ini dilakukan dengan menggunakan 93 data suara dengan 46 data suara bahagia dan 46 data suara marah menggunakan bahasa Indonesia. Berikut adalah hasil uji pengenalan emosi suara bahagia disajikan pada tabel berikut ini.

Tabel 4. Hasil Pengujian Data Suara bahagia

Kalimat Yang Diuji	Data Suara Yang Diuji	Hasil Uji
Setelah sekian lama berjuang menyelesaikan skripsi, akhirnya aku bisa lulus juga dari kampus ini.	Bahagia1	Bahagia
	Bahagia1	Bahagia
	Bahagia1	Bahagia
	Bahagia1	Bahagia
	Bahagia1	Bahagia

Tingkat keberhasilan dalam (%) untuk deteksi emosi bahagia dapat dihitung menggunakan persamaan:

$$validitas = \frac{\text{hasil yang sesuai}}{\text{banyaknya hasil uji}} \times 100\%$$

Sehingga dapat diketahui nilai keberhasilan dari deteksi bahagia1 adalah:

$$\frac{5}{5} \times 100\% = 100\%$$

Tabel 5. Hasil Pengujian Data Suara Marah

Kalimat Yang Diuji	Data Suara Yang Diuji	Hasil Uji
Kenapa kamu sibuk mengurus hidupku, padahal kan aku yang menjalani hidupku sendiri.	Marah1	Marah
	Marah1	Marah
	Marah1	Marah
	Marah1	Marah
	Marah1	Marah
	Marah1	Marah

Tingkat keberhasilan dalam (%) untuk deteksi emosi bahagia dapat dihitung menggunakan persamaan:

$$validitas = \frac{\text{hasil yang sesuai}}{\text{banyaknya hasil uji}} \times 100\%$$

Sehingga dapat diketahui nilai keberhasilan dari deteksi bahagia1 adalah:

$$\frac{6}{6} \times 100\% = 100\%$$

Dari tabel data uji suara bahagia dan suara marah diatas dapat dilihat nilai persentase pada tabel bahagia 1 sampai dengan 5 mendapatkan nilai yang terbilang sangat baik. Namun pada tabel bahagia 2 dan 3 terdapat masing-masing 1 data yang salah di prediksi oleh sistem dan akurasi yang didapatkan

pada tabel bahagia 2 mendapatkan nilai sebesar 92% dan pada tabel bahagia3 mendapat nilai sebesar 88%. Sebaliknya pada tabel suara marah 1 sampai 5 mendapatkan nilai persentase sempurna yaitu 100%. Sehingga dapat disimpulkan bahwa kecocokan pada database emosi yang lebih baik adalah suara marah dengan akurasi dalam setiap kalimat tabel marah 1 sampai marah 5 sebesar 100%.

IV. PENUTUP

A. KESIMPULAN

Berdasarkan penelitian yang dilakukan peneliti untuk menguji performa Learning Vector Kuantization (LVQ) dalam mengklasifikasikan data suara emosi bahagia dan marah dapat disimpulkan sebagai berikut:

1. Model Learning Vector Quantization (LVQ) mampu menentukan jenis emosi suara yang masuk ke dalam model. Diperoleh tingkat akurasi emosi suara test sebesar 87% ketika dilakukan pengujian model menggunakan dataset sebanyak 380 data.
2. Metode confusion matrix digunakan untuk mengevaluasi model yang sudah di buat. Diperoleh tingkat akurasi dari evaluasi model sebesar 87% dengan perhitungan parameter lainnya seperti precision dengan nilai rata-rata 87%, recall 87% dan f-1 score 87% dengan jumlah data uji 93 data. Sehingga bisa disimpulkan bahwa model bekerja dengan baik.

B. SARAN

Dari hasil penelitian hal ini dapat menjadi pertimbangan untuk kekurangan dalam penelitian dan potensi yang dapat dikembangkan dari penelitian ini, saran dari penulis antara lain:

1. Pengambilan data emosi suara pada subject hidup sebaiknya memperhatikan intonasi pada saat pengambilan data untuk meminimalisir noise dan tidak adanya kesamaan dalam intonasi suara yang direkam.
2. Untuk penelitian selanjutnya diusahakan membuat suatu sistem atau aplikasi agar pada saat pengambilan data digunakan satu aplikasi saja untuk mendapatkan kualitas suara yang diinginkan.

3. Menambah data latih yang digunakan untuk melatih model sehingga pada saat pengklasifikasian suara emosi semakin banyak data semakin baik model dan akurasi semakin meningkat.

DAFTAR RUJUKAN

- [1] Mahardhika, P - (2019) APLIKASI PENGENALAN SUARA MANUSIA DENGAN METODE JARINGAN SYARAF TIRUAN LEARNING VECTOR QUANTIZATION. Skripsi thesis, Universitas Pembangunan Nasional Veteran Jakarta.
- [2] Nwe, T. L., Foo, S. W., & De Silva, L. C. (2003). Speech emotion recognition using hidden Markov models. *Speech Communication*, 41(4), 603–623. [https://doi.org/10.1016/S0167-6393\(03\)00099-2](https://doi.org/10.1016/S0167-6393(03)00099-2)
- [3] Pao, T. L., Chen, Y. Te, Yeh, J. H., & Lu, J. J. (2020). Detecting emotions in Mandarin speech. *Proceedings of the 16th Conference on Computational Linguistics and Speech Processing, ROCLING 2004*, 10(3), 347–362
- [4] Safriadi, S., & Rahmadani, R. (2020). Klasifikasi Gender Berdasarkan Suara dengan Naive Bayes dan Mel Frequency Cepstral Coefficient. *VOCATECH: Vocational Education and Technology Journal*, 2(1), 19-26.
- [5] Abushariah, A. A. M., Gunawan, T. S., Chebil, J., & Abushariah, M. A. M. (2012). Voice based automatic person identification system using vector quantization. *2012 International Conference on Computer and Communication Engineering, ICCCE 2012*, July, 549–554. <https://doi.org/10.1109/ICCCE.2012.6271247>
- [6] Indrawati, N. Y., Musrini B. M., & Hafiz D. A. (2016). Aplikasi Pendeteksi Emosi Manusia Menggunakan Metode Mfcc Dan Dtw. *Seminar Nasional Teknologi Informasi 2016*, 29–35.
- [7] Arief, R. (2021). Klasifikasi Audio Ucapan Emosional Menggunakan Model LSTM. *Konferensi Nasional Ilmu Komputer*, 524–529