

## KLASIFIKASI KEMIRIPAN SUARA REKAMAN MENGGUNAKAN METODE MEL-FREQUENCY CEPSTRA COEFFICIENT DAN MINKOWSKI

### CLASSIFICATION OF RECORDED VOICE SIMILARITY USING MEL-FREQUENCY CEPSTRAL COEFFICIENT AND MINKOWSKI METHOD

Aidi Putrasyah<sup>1</sup>, Ermatita<sup>2</sup>, Abdiansah<sup>3</sup>

<sup>1,2,3</sup>Jurusan Magister Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya  
aidiputrasyah@gmail.com

#### ABSTRACT

Current technology has dominated various fields ranging from education, offices, commercial, industry and even law. With advances in information technology, human needs will be available easily, practically and without limits. On the other hand, the development of information technology raises concerns about the development of crime-related crimes. In some cases there is usually evidence left behind, such as voice recordings resulting from telephone conversations. This recorded voice can be used as a support in the investigation during the trial to identify the perpetrators of the crime. Research to identify sound still has obstacles, namely noise that occurs when recording sound. The MFCC method is a method that can be used to extract features and reduce noise. Research using algorithms in analyzing recorded voices is circulating on the internet, one of which is the KNN algorithm which can be used for classification, identification, and prediction. Identification of recorded voices using the KNN algorithm with the Minkowski method is carried out for voice recognition in determining whether or not the voice recorded by the perpetrator of the crime is identical with the voice sample of the suspect. This study involved two respondents as suspects and three respondents as perpetrators. Each respondent will make a voice recording, where this recording will be cut into even parts consisting of training data and test data. The results are expected to be able to identify the voice recording of the perpetrator against the voice of the suspect by getting the smallest distance that is close to similarity so that the sound recording evidence can be accounted for in the trial.

**Keywords:** Technology, Identification, Extraction, KNN, Minkowski

#### ABSTRAK

Teknologi saat ini telah mendominasi berbagai bidang mulai dari pendidikan, perkantoran, komersial, industri bahkan hukum. Dengan kemajuan Teknologi informasi kebutuhan manusia akan tersedia dengan mudah, praktis dan tanpa batas. Disisi lain, berkembangnya Teknologi informasi menimbulkan kekhawatiran pada perkembangan tindak pidana yang berhubungan dengan kejahatan. Pada beberapa kasus biasanya terdapat barang bukti yang ditinggalkan seperti rekaman suara yang dihasilkan dari percakapan menggunakan telepon. Suara rekaman ini dapat digunakan sebagai pendukung dalam penyidikan saat dipersidangan untuk mengidentifikasi pelaku kejahatan. Penelitian untuk mengidentifikasi suara masih terdapat kendala yaitu kebisingan yang terjadi saat merekam suara. Metode MFCC merupakan metode yang dapat digunakan untuk mengekstraksi fitur serta mengurangi kebisingan atau noise. Penelitian menggunakan algoritma dalam menganalisis suara rekaman banyak beredar di internet, salah satunya algoritma KNN yang dapat digunakan untuk klasifikasi, identifikasi, dan prediksi. Identifikasi suara rekaman menggunakan algoritma KNN dengan metode minkowski dilakukan untuk pengenalan suara dalam menentukan identik atau tidaknya antara suara rekaman pelaku kejahatan dengan sampel suara tersangka. Penelitian ini melibatkan dua responden sebagai tersangka dan tiga responden sebagai pelaku. Setiap responden akan melakukan perekaman suara, dimana rekaman ini nanti akan dipotong menjadi sebes bagian yang terdiri dari data latihan dan data uji. Hasil yang diharapkan dapat melakukan pengenalan rekaman suara pelaku terhadap suara tersangka dengan mendapatkan jarak terkecil yang mendekati kemiripan sehingga barang bukti rekaman suara dapat dipertanggung jawabkan dalam persidangan.

**Kata Kunci:** Teknologi, Identifikasi, Ekstraksi, KNN, Minkowski

#### PENDAHULUAN

Rekaman suara merupakan bagian dari alat bukti digital yang dapat ditemukan dalam kasus pidana ataupun perdata (Umar et al., 2019) Pada beberapa kasus, rekaman

suara menjadi barang bukti yang sering diperdebatkan terutama dalam hal mengukur dan mendeteksi keakuratan suara. Dengan demikian, diperlukan

dukungan keputusan dalam memverifikasi barang bukti audio (Renza et al., 2018).

Rekaman suara dapat berisi percakapan wawancara, rekaman interogasi, atau penyadapan. Hasil (Maher, 2018). Identifikasi rekaman suara apakah identik atau tidak dapat menggunakan metode dengan menerapkan algoritma (Mataram & Sugiantoro, 2009). KNN merupakan algoritma yang berfungsi untuk melakukan klasifikasi data berdasarkan data pembelajaran dimana dataset training tersimpan.

Klasifikasi dapat diperoleh dengan membandingkan record yang paling cocok dengan dataset (Iswanto et al., 2021). Selain untuk klasifikasi, KNN dapat digunakan untuk mendapatkan nilai akurasi dan estimasi (Ali et al., 2019). Pada algoritma KNN terdapat beberapa metode yang dapat digunakan untuk menghitung jarak yaitu eucidean, manhattan, hamming distance, dan minkowski (Azwar et al., 2021).

Terdapat Penelitian yang membahas tentang identifikasi pengenalan suara berdasarkan jenis keamin menggunakan KNN, naïve bayes, mutiayer perceptron, random forest, dan support vector machine. Hasil yang didapat menunjukkan bahwa SVM adalah pengklasifikasian terbaik diantara lima skema lainnya (Ahmad et al., 2016). Sementara itu Penelitian lain dengan membandingkan suara manusia berdasarkan jenis keamin menggunakan algoritma KNN dengan metode eucidean memiliki nilai jarak yang tinggi dan tingkat keakuratan dalam perbandingan rekaman suara yang rendah dengan menggunakan dataset yang sedikit yaitu terdiri dari 5 pembicara (Umar et al., 2019). Penelitian lain tentang pengenalan suara dalam menentukan pelaku kejahatan menggunakan algoritma KNN dengan metode minkowski, memiliki nilai jarak terkecil yang mendekati kemiripan dengan nilai akurasi dari sampel suara identik masih rendah yaitu sebesar 0.63 dan 0.18 (Azwar et al., 2021). Penelitian dengan menggunakan algoritma KNN dan metode minkowski dapat digunakan untuk Penelitian tentang

identifikasi suara sehingga pada Penelitian ini penulis menggunakan algoritma KNN dan metode minkowski agar mendapatkan jarak terkecil yang mendekati kemiripan dalam pencocokan rekaman suara. Sebelum menerapkan algoritma KNN rekaman suara terlebih dahulu di ekstraksi menggunakan metode yang ada, salah satu metode yang dapat digunakan untuk mengekstraksi fitur serta mengurangi kebisingan atau noise yaitu MFCC karena pada Penelitian sebelumnya untuk nilai akurasi masih rendah.

Metode MFCC merupakan fitur yang sering digunakan dalam pengenalan suara dan pembicara otomatis dengan merubah sinyal suara menjadi beberapa parameter (Li et al., 2020) (Siriwardena et al., 2022). Langkah awal dalam mengekstraksi fitur yaitu dengan mengidentifikasi konten dan membuang semua hal yang tidak berkaitan dengan informasi seperti kebisingan atau noise, emosi, dan lainnya. Berdasarkan latar belakang yang telah dijabarkan maka tujuan dari Penelitian ini adalah mengidentifikasi rekaman suara tersangka dengan rekaman suara pembanding dilihat dari identik atau tidaknya menggunakan algoritma KNN dengan metode minkowski, yang menghasilkan tingkat akurasi yang lebih tinggi dibandingkan dengan Penelitian sebelumnya serta mendapatkan nilai atau jarak terkecil untuk pengambilan keputusan dalam menentukan barang bukti berupa suara rekaman yang nantinya dapat digunakan dalam persidangan sebagai barang bukti yang kuat

## **METODE**

Penelitian ini diawali dengan identifikasi masalah yang mendasar. Penelitian, kajian pustaka, pengambilan sampel suara melalui rekaman, ekstraksi fitur, pengklasifikasian KNN dengan metode minkowski, analisis Hasil, dan kesimpulan. Tahapan Penelitian seperti pada Gambar 1 berikut, yang terdiri dari pengambilan sampel suara, proses ekstraksi fitur menggunakan metode MFCC, Klasifikasi KNN dengan metode

minkowski, analisis Hasil dan penarikan kesimpulan, membuat aporan Penelitian, serta publikasi Hasil Penelitian.



**Gambar 1. Tahapan Penelitian**

### Suara

Audio adalah suara yang dihasilkan oleh getaran suatu benda yang berupa sinyal analog dengan amplitudo yang terus berubah dari waktu ke waktu dan menyebar melalui udara (Imario et al., 2017). Suara dihasilkan melalui proses generation dan filtering. Generation adalah proses pertama menciptakan suara dengan menggetarkan pita suara (vocal cord) di kotak suara (aring) untuk menghasilkan suara periodik. (Azwar et al., 2021). Audio dalam bentuk rekaman suara memiliki karakter dalam pengucapan yang berbeda-beda dan kondisi yang tidak jelas. Rekaman suara berisi gelombang suara yang direkam dengan Teknologi perekaman digital (Mansyur & Manurung, 2017).

### Audio Forensik

Audio forensik merupakan penerapan ilmu pengetahuan yang berkaitan dengan penggunaan audio pada proses penyidikan dan mendapatkan fakta-fakta untuk persidangan. Rekaman suara yang dijadikan bukti dapat menunjukkan identitas dari subjek dengan menggunakan voice identification (Deva & Mardianto, 2019).

### Me Frequency Cepstra Coefficient (MFCC)

MFCC adalah salah satu metode yang dapat digunakan dalam melakukan ekstraksi fitur, dengan cara merubah sinyal suara

menjadi beberapa parameter. Ekstraksi fitur pada Automatic Speech Recognition (ASR) merupakan cara perhitungan urutan berdasarkan fitur vektor yang dapat menyimbolkan sinyal yang ada secara ide (Dave, 2013). Sama halnya pendapat (Putra et al., 2017). yang menyatakan bahwa MFCC merupakan suatu metode ekstraksi fitur yang keubarannya berupa feature vector yang dikenai dengan Cepstrum. Beberapa keunggulan dari metode MFCC adalah sebagai berikut:

1. Mampu menangkap karakteristik suara yang diperlukan untuk pengenalan suara.
2. Hasilkan data minima tanpa kehilangan informasi penting yang ada di dalamnya.
3. Mereplikasi sinyal sensorik organ pendengaran manusia.

Tahapan-tahapan ekstraksi suara menggunakan MFCC terdiri dari 7 tahapan yaitu Pre-Emphasize, Framing, Windowing, Fast Fourier Transform (FFT), Mel Frequency Warping, Discrete Cosine Transform, dan Cepstral filtering (Hidayat, 2022). Berikut ini penjelasan dari masing-masing tahapan tersebut:

#### a. Pre-Emphasize

Pre-Emphasize, berfungsi untuk menstabilkan nilai magnitudo dari sinyal suara. Pada proses ini dapat menghilangkan sedikit noise dari suara yang masuk, sehingga akurasi dari tahap ekstraksi fitur dapat ditingkatkan. Berikut proses pre-emphasize seperti pada persamaan.

$$y[n]=s[n]-\alpha s[n-1], 0.9 \leq \alpha < 1.0$$

Keterangan :

$y[n]$  = sinyal Hasil pre-emphasize filter

$s[n]$  = sinyal sebelum pre-emphasize filter

$\alpha$  = nilai alpha

#### b. Framing

Framing adalah fase di mana sampel suara dibagi menjadi beberapa bingkai atau slot. Framing berfungsi membagi sinyal suara menjadi beberapa frame dengan panjang yang umumnya pendek yaitu sekitar 20 hingga 40 ms. Keuntungan framing adalah memudahkan analisis dan mengurangi beban dari memori. Berikut proses framing seperti pada persamaan (2).

$$\text{Jumlah Frame} \frac{T_s}{M}$$

Keterangan:

$T_s$  = Durasi pengambilan suara

$M$  = Panjang frame (ms)

### c. Windowing

Windowing adalah proses pengurangan kesenjangan (discontinuitas) sinyal pada awal dan akhir suatu frame setelah proses pembokiran frame. Fungsi windowing yaitu mengurangi noise yang muncul di kedua ujung frame. Berikut proses windowing seperti pada persamaan.

$$x(n) = x_i(n)w(n)$$

Keterangan :

$x(n)$  = Nilai sampel sinyal Hasil windowing

$x_i(n)$  = Nilai sampel sinyal dari frame sinyal ke  $i$

$w(n)$  = Fungsi window

### d. Fast Fourier Transform (FFT)

FFT adalah tahapan untuk mengubah sinyal suara dari domain waktu menjadi domain frekuensi. Hasil dari sinyal FFT disebut spektrogram. Berikut proses windowing seperti pada persamaan.

$$f(n) = \sum_{k=0}^{N-1} y_k e^{-2\pi jkn} \quad n = 0, 1, 2, \dots, N-1$$

Keterangan :

$f(n)$  = Frekuensi

$N$  = Jumlah sampel pada masing-masing frame

$k = 0, 1, 2, \dots, (N-1)$

$j$  = Bilangan imajiner ( $\sqrt{-1}$ )

### e. Mel-Frequency Wrapping

Mel-frequency wrapping adalah proses yang dilakukan dengan cara filterbank. Filterbank digunakan untuk membungkus (wrapping) mel-frequency dan dipadatkan di domain waktu atau domain frekuensi. Hasil akhir dari tahap ini adalah mendapatkan beberapa mel-frequency yang di-filter bank. Nilai mel-frequency menunjukkan berapa banyak energi yang dimiliki setiap filter mel dalam rentang frekuensi. Berikut proses mel-frequency wrapping seperti pada persamaan.

$$\text{mel } f = \frac{2595 + \log_{10} \left( 1 + \frac{f}{700} \right)}{\frac{s_i}{2}}$$

Keterangan:

$s_i$  = Sinyal awal Hasil FFT

$f$  =  $f_0 - f_n$

### f. Discrete Cosine Transform

Discrete cosine transform merupakan tahapan yang menghasilkan citra yang bagus dari hasil spektrum suara dari perhitungan mel-frequency spectrum. Berikut proses discrete cosine transform seperti pada persamaan (6).

$$C_n = \sum_{k=1}^K (\log s_k) \cos \left[ n \left( k - \frac{1}{2} \right) \frac{\pi}{K} \right]; \quad n = 1, 2, \dots, K \quad (6)$$

Keterangan :

$s_k$  = Keuaran dari proses filter bank pada index  $k$

$K$  = Jumlah koefisien yang diharapkan

### g. Cepstral Filtering

Hasil dari proses utama ekstraksi ciri MFCC memiliki beberapa keamahan. Keamahan tersebut merupakan keamahan low order dari cepstral coefficients yang sangat sensitif terhadap spectra noise. Keamahan lainnya adalah keamahan high order yang sangat sensitif terhadap noise. Cepstral filtering adalah salah satu teknik standar yang digunakan untuk meminimalkan sensitivitas tersebut. Berikut proses discrete cosine transform seperti pada persamaan.

$$w[n] = \left\{ 1 + \sin \frac{1}{2} \sin \left( \frac{n\pi}{L} \right) \right\} \quad n = 1, 2, \dots, L$$

Keterangan:

- = Jumlah cepstra coefficients
- N = Index dari cepstra coefficients

**K-Nearest Neighbors (KNN)**

KNN adalah algoritma yang berfungsi untuk melakukan klasifikasi data berdasarkan data pembelajaran dimana dataset training disimpan. Klasifikasi diperoleh dengan membandingkan record yang paling cocok dengan dataset pelatihan. Selain untuk klasifikasi, KNN dapat digunakan untuk estimasi dan prediksi (Handayani, 2019)(Iswanto et al., 2021). Berikut langkah-langkah algoritma KNN adalah sebagai berikut:

- a. Menentukan parameter K (jumlah tetangga terdekat)
- b. Menghitung jarak (similarity) antara semua record training dan objek baru dengan rumus minkowski.
- c. Mengurutkan data berdasarkan nilai jarak dari terkecil ke terbesar.
- d. Mengambil data dari beberapa nilai K.
- e. Menentukan atribut yang paling sering muncul pada k catatan training yang paling dekat dengan objek.

**Minkowski**

Minkowski distance adalah suatu metrik pada ruang vektor bernorma (normed vector space) dan dianggap sebagai generalisasi dari jarak Euclidean dan jarak Manhattan (Nishom, 2019). Menghitung jarak dengan metode ini dapat menggunakan rumus pada persamaan berikut.

$$d(x, y) = \left( \sum_{i=1}^n |x_i - y_i|^p \right)^{\frac{1}{p}}$$

**Keterangan:**

d = jarak antara x dan y

n = jumlah data

x = data pusat cluster

$x_i$  = data pada cluster ke i

y = data pada atribut

$y_i$  = data pada setiap data ke i

I = setiap data

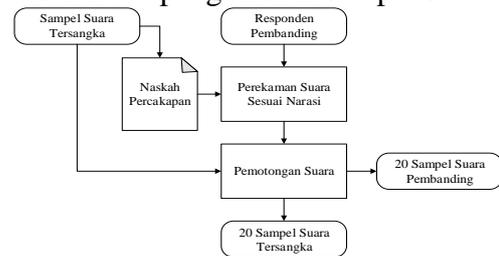
p = power

Jarak minkowski menggunakan pangkat p, dimana p adalah bilangan bulat

antara dua titik yaitu x dan y. Dalam pengukuran jarak objek dengan menggunakan minkowski distance biasanya menggunakan  $p = 1$  atau  $2$ , dimana angka ini sesuai dengan jarak Manhattan dan jarak Euclidean.

**Pengambilan Sampel Suara**

Tahapan pengumpulan sampel suara digunakan untuk memahami teknik pengambilan data rekaman suara serta prosedur yang akan digunakan sesuai dengan Standard Operating Procedure (SOP). Sampel rekaman suara tersangka didapatkan dari Hasil penyadapan suara tersangka yang disimpan pada media penyimpanan dan diakusisi. Untuk pengambilan sampel yaitu dengan membandingkan sampel rekaman suara tersangka dengan rekaman suara dari dua responden yang berisi kata-kata yang sesuai naskah percakapan suara asli tersangka. Rekaman suara pembandingan yang didapatkan kemudian dipotong menjadi 20 sampel suara sesuai dengan prosedur dalam penanganan barang bukti suara, dimana dalam pengoahan data suara perkata tidak boleh kurang dari 19 kata. Berikut ini Gambar 2 alir pengambilan sampel suara.



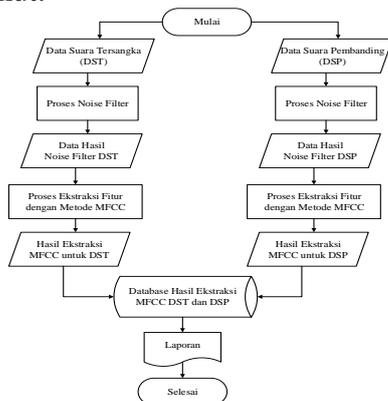
**Gambar 2. Alir Pengambilan Sampel Suara**

**2.7 Proses Ekstraksi Fitur Menggunakan MFCC**

Metode ekstraksi fitur yang digunakan yaitu MFCC. MFCC merupakan metode speech technology yang sering digunakan untuk pengenalan ucapan atau suara. Feature extraction pada MFCC bekerja dengan cara merubah sinyal suara menjadi parameter yang berbeda-beda. Tahapan ekstraksi fitur menggunakan metode MFCC diawali memisahkan antara data suara buatan dan data suara angsung,

selanjutnya meakukan proses noise filter atau menghiangkan derau.

Dari Hasil noise filter selanjutnya data suara buatan dan data suara angsung diekstraksi menggunakan metode MFCC, sehingga didapatkan Hasil ekstraksi yang akan dijadikan database Hasil ekstraksi MFCC untk data suara buatan dan data suara angsung. Untuk memudahkan dalam mengetahui aur ekstraksi fitur dapat diihat pada fowchart MFCC seperti pada Gambar 3 berikut.

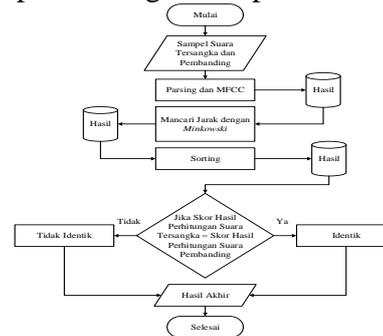


**Gambar 3. Fowchart MFCC**

**Pengkasifiasian KNN Dengan Minkowski**

Pada tahap ini sampel rekaman suara akan diubah menjadi nilai spectrum dengan metode MFCC, selanjutnya akan membandingkan rekaman suara tersangka dengan rekaman suara pembanding. Proses membandingkan sampel suara ini diawali dengan meakukan oad atau input data berupa sampel rekaman suara dan sampel rekaman suara pembanding. Selanjutnya meakukan parsing yaitu mengubah data frame rate sampel rekaman suara berbentuk array muti dimenasi ke satu dimensi agar data bisa di ekstraksi menggunakan MFCC. Proses ekstraksi dilakukan untuk merubah sampel suara menjadi nilai spectrum agar memudahkan dalam membandingkan sampel suara menggunakan metode minkowski. Perhitungan dengan metode minkowski menghasilkan nilai jarak yang akan disimpan dan diambil kembali pada proses sorting yaitu proses penentuan identik atau tidak identik, proses ini ditentukan dengan jika skor Hasil perhitungan suara tersangka sama dengan skor Hasil perhitungan suara pembanding

maka identik. Berikut Gambar 4 fowchart proses perbandingan sampel suara.

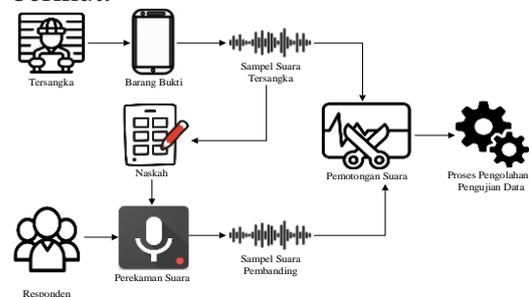


**Gambar 4. Proses Perbandingan Sampel Suara**

**HASIL DAN PEMBAHASAN**

Proses pengujian diawali dengan meakukan ekstraksi fitur dari sampel rekaman suara tersangka dan suara pembanding untuk mendapatkan nilai spectrum. Nilai spectrum ini dikemudian digunakan pada proses sorting menggunakan metode minkowski untuk mendapatkan nilai jarak dan mengkasifikasikan apakah rekaman suara tersebut identik atau tidak identik. Bahasan terakhir pada bab ini yaitu menganalisis Hasil dan akurasi dari Penelitian yang dilakukan.

Skenario kasus dibutuhkan untuk mengetahui dan menentukan jaannya proses pengujian dalam membandingkan sampel rekaman suara tersangka terhadap sampel rekaman suara pembanding, ha ini untuk menggambarkan suatu tindak kejahatan yang terjadi seperti pada Gambar 5 berikut.



**Gambar 5. Iustrasi Skenario Kasus**

Persiapan awa pada Penelitian ini adaah mempersiapkan sampel rekaman suara tersangka dengan suara pembanding. Berikut langkah-langkah dalam persiapan sampel suara rekaman.

1. Sampel rekaman suara pembanding berisi rekaman suara responden yang mengucapkan kata-kata sesuai naskah, dimana isi dari naskah berdasarkan kata yang diucapkan pada rekaman suara tersangka. Selanjutnya sampel rekaman suara tersangka dan sampel suara pembanding dipotong menjadi 20 kata dimana kata-kata tersebut seperti terlihat pada Gambar 6.

Apakah mungkin memasukan mayat ke dalam tas tidak terlalu sulit, karena sudah biasa mengerjakannya, saya tau cara memotong dengan sangat baik.

**Gambar 6. Naskah Percakapan**

2. Sampel rekaman suara tersangka dan rekaman suara pembanding yang sudah dipotong kemudian dilakukan proses pengujian dan membandingkan sampel rekaman pada tahap selanjutnya.

Proses pengujian adalah proses yang dilakukan untuk mengolah sampel rekaman suara tersangka dengan dua sampel rekaman suara pembanding dimana proses yang pertama kali dilakukan adalah mengubah sampel rekaman suara ke dalam nilai *spectrum* dengan menggunakan metode MFCC dan membandingkan sampel rekaman suara tersangka dengan sampel rekaman suara peaku yang berupa nilai *spectrum* dengan menggunakan metode *minkowski* dan diproses menggunakan *toos orange*. Tahapan pengujian diawali dengan proses *extraction* sampel rekaman suara. Proses *extraction* adalah proses untuk mengkonversi sampel rekaman suara tersangka (*suspect*) dengan suara pembanding kedalam nilai *spectrum* dengan menggunakan metode MFCC. Adapun Hasil konversi yang berupa nilai *spectrum* sampel rekaman suara tersangka seperti terlihat pada Tabe 1 dan Tabe 2 berikut:

**Tabel 1. Nilai Spectrum Sampel Rekaman Suara Tersangka (1)**

Suara	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10
00	13.690	17.324	18.617	16.920	16.214	20.380	20.339	20.049	15.344	16.788
01	5.695	8.072	9.527	10.019	11.729	12.742	12.936	11.220	10.007	13.108
02	-11.816	5.505	12.309	12.914	11.759	11.704	13.509	14.951	10.559	12.768
03	16.643	17.761	19.083	18.179	16.909	16.433	14.245	15.506	17.015	16.620
04	-3.438	5.955	13.010	18.860	18.153	15.497	13.373	12.125	12.669	15.063
05	-2.172	-2.842	4.536	11.735	14.646	12.220	9.576	8.209	9.543	10.793
06	13.010	15.041	16.560	16.608	16.904	17.217	14.444	13.170	13.992	14.975
07	7.788	7.984	6.556	2.383	6.207	10.057	3.169	8.109	2.489	3.703
08	8.656	5.309	14.107	-15.080	-2.788	19.153	9.835	15.697	17.966	14.586
09	6.623	9.535	12.210	15.356	15.508	13.763	12.496	6.485	9.391	4.214
10	10.365	11.020	13.110	16.750	17.663	17.448	15.589	15.466	16.747	14.896

**Tabel 2. Nilai Spectrum Sampel Rekaman Suara Tersangka (2)**

Suara	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20
00	19.431	19.837	18.950	18.689	20.443	19.413	18.379	16.233	10.029	9.752
01	16.352	15.709	15.614	16.073	13.632	8.549	20.451	-4.966	6.873	14.735
02	14.417	14.965	15.806	16.200	11.818	4.683	8.631	13.197	15.492	17.944
03	17.110	18.252	17.675	18.479	19.483	19.589	16.697	12.074	12.901	14.478
04	16.881	16.023	10.301	13.389	17.174	20.328	20.908	20.589	21.082	20.399
05	11.927	12.545	14.004	14.875	14.217	13.971	14.874	15.356	13.974	12.245
06	14.497	14.957	14.743	13.293	13.885	12.902	10.302	9.189	12.420	15.384
07	14.666	15.278	6.251	6.100	6.313	7.368	17.732	18.334	18.955	18.295
08	16.990	18.060	14.304	9.849	10.455	7.290	8.582	10.402	13.372	14.096
09	10.718	12.272	11.636	4.083	5.694	3.691	12.427	11.924	3.468	4.308
10	11.219	5.324	3.848	3.157	10.442	6.601	17.015	13.871	13.397	4.403

Selanjutnya proses *extraction* sampel rekaman suara pembanding, pada proses ini menggunakan *source code* yang sama seperti yang digunakan pada proses *extraction* sampel rekaman suara peaku, yang membedakan dari *source code* untuk *extraction* sampel suara pembanding pada parameter harus berisi variabel pembanding, Hasil *extraction* yang didapatkan dari sampel suara pembanding seperti terlihat pada Tabe 3 dan Tabe 4 berikut:

**Tabel 3. Nilai Spectrum Sampel Rekaman Suara Pembandingan (1)**

Suara	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10
00	14.208	16.789	13.881	11.080	11.954	19.106	20.195	20.006	15.842	16.851
01	6.985	10.322	12.146	11.256	12.429	13.779	13.589	11.659	10.466	13.636
02	-11.846	5.504	12.308	12.912	11.806	11.857	13.320	13.239	9.665	12.740
03	17.238	18.230	19.600	18.719	17.575	17.314	15.632	16.092	17.524	17.465
04	3.615	6.041	13.020	18.265	17.366	14.818	12.684	12.155	12.692	14.788
05	1.979	2.432	4.365	10.336	11.580	9.961	8.296	7.770	9.210	10.325
06	13.671	15.374	17.278	17.358	17.579	18.078	16.196	15.599	16.009	17.068
07	8.727	7.741	8.482	5.913	2.686	4.619	8.726	6.203	4.145	5.074
08	8.657	5.309	14.104	-15.080	-2.787	19.153	9.835	15.697	17.969	14.984
09	6.633	9.545	12.219	15.375	15.522	13.773	12.506	6.496	11.114	3.822
10	9.017	10.577	12.646	16.363	17.323	17.068	15.110	14.990	16.398	14.668

**Tabel 3. Nilai Spectrum Sampel Rekaman Suara Pembandingan (2)**

Suara	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20
00	18.650	19.222	18.990	18.612	19.699	18.379	17.635	15.903	10.755	10.620
01	16.366	15.510	15.780	14.813	12.337	10.798	6.714	3.322	8.126	14.165
02	14.456	14.950	15.765	15.881	11.768	5.392	8.546	12.828	15.179	17.809
03	18.401	19.190	18.466	19.035	19.932	20.175	18.053	13.319	14.069	15.899
04	16.329	15.500	10.024	12.835	16.345	18.111	16.769	15.410	16.941	16.354
05	11.164	11.111	9.542	7.834	6.883	6.977	8.356	10.007	10.399	10.114
06	17.141	17.338	17.121	15.730	15.819	16.745	17.548	12.573	13.720	16.672
07	8.921	11.579	9.656	6.470	4.073	4.138	7.773	13.635	16.622	16.584
08	16.386	16.839	14.370	10.085	10.568	10.370	10.487	11.449	13.297	14.154
09	6.887	7.324	6.742	6.483	6.751	6.481	5.781	5.709	6.725	6.957
10	11.136	5.758	6.278	9.841	7.852	2.358	13.212	7.621	21.434	2.922

Setelah nilai *spectrum* sampel rekaman suara tersangka dan sampel rekaman suara pembanding telah didapat, selanjutnya melakukan proses *sorting* untuk mendapatkan nilai jarak yang dihasilkan metode *minkowski*. Nilai jarak sampel rekaman suara pembanding seperti terlihat pada Tabe 5 dan Tabe 6 berikut.

**Tabel 5. Nilai jarak Sampel Rekaman Suara Pembandingan (1)**

Suara	N1	N2	N3	N4	N5	N6	N7	N8	N9	N10
00	0.518	0.534	4.736	5.840	4.260	1.274	0.344	0.043	0.498	0.063
01	1.290	2.250	2.619	1.237	0.700	1.037	0.652	0.438	0.458	0.528
02	0.029	0.001	0.001	0.002	0.047	0.152	0.190	1.713	0.895	0.028
03	0.595	0.469	0.517	0.540	0.666	0.881	1.387	0.587	0.510	0.845
04	7.053	0.086	0.010	0.595	0.787	0.679	0.689	0.030	0.023	0.275
05	4.151	5.275	0.171	1.399	3.066	2.258	1.279	0.439	0.332	0.468
06	0.661	0.333	0.718	0.750	0.675	0.860	1.752	2.430	2.017	2.093
07	0.940	0.243	1.926	3.530	3.521	5.439	5.557	1.906	1.655	1.371
08	0.001	0.000	0.004	0.000	0.000	0.000	0.000	0.000	0.003	0.399
09	0.010	0.010	0.010	0.019	0.014	0.010	0.010	0.011	1.722	0.393
10	1.348	0.442	0.464	0.387	0.340	0.380	0.479	0.476	0.349	0.228

**Tabel 6. Nilai jarak Sampel Rekaman Suara Perbandingan (2)**

Suara	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20
00	0.782	0.615	0.040	0.078	0.743	1.033	0.745	0.329	0.726	0.868
01	0.015	0.199	0.166	1.260	1.295	2.249	13.737	8.289	1.253	0.570
02	0.039	0.015	0.040	0.319	0.049	0.709	0.085	0.369	0.313	0.135
03	1.291	0.938	0.791	0.556	0.449	0.587	1.356	1.245	1.168	1.422
04	0.552	0.523	0.277	0.554	0.829	2.217	4.140	5.178	4.141	4.045
05	0.764	1.434	4.462	7.041	7.335	6.994	6.517	5.349	3.575	2.131
06	2.644	2.381	2.378	2.437	1.935	3.843	7.246	3.384	1.300	1.287
07	5.745	3.700	3.404	0.370	2.240	3.230	9.959	4.699	2.333	1.712
08	0.604	1.220	0.066	0.236	0.133	3.080	1.905	1.048	0.075	0.059
09	3.831	4.949	4.894	2.400	1.057	2.790	6.646	6.215	3.258	2.649
10	0.083	0.433	2.430	6.684	2.589	4.242	3.803	6.250	8.037	1.480

Proses sorting merupakan proses dalam pengurutan nilai jarak yang terkecil hingga nilai jarak yang terbesar, pada proses sorting terdapat fungsi yang digunakan untuk membandingkan antara jarak yang didapatkan sampel suara tersangka dengan jarak yang didapatkan oeh sampel suara perbandingan. Dengan ketentuan yaitu jika kedua sampel tersebut memiliki nilai jarak yang sama maka dapat dikatakan identik atau sebaiknya maka dikatakan tidak identik.

Tahap analisis adaah tahap penjabaran secara engkap makna dari Hasil yang didapatkan seperti tingkat akurasi dari metode yang digunakan dalam proses membandingkan sampel rekaman suara tersangka dengan kedua sampel rekaman suara perbandingan kemudian menganalisis nilai jarak dan nilai signfikasi, berikut ini Tabe 7 Hasil sorting nilai jarak suara yaitu sebagai berikut:

**Tabel 7. Hasil Sorting Nilai jarak Minkowski**

No	Jarak <i>Minkowski</i>	Keterangan
00	7,300	Identik
01	14,759	Tidak Identik
02	1,839	Identik
03	2,622	Identik
04	8,926	Identik
05	12,444	Identik
06	8,371	Identik
07	12,333	Identik
08	3,403	Identik
09	9,738	Identik
10	10,718	Identik

Tabe 7 menampilkan nilai jarak minkowski dari Hasil sorting perbandingan sampel rekaman suara antara sampel suara tersangka dan sampel suara perbandingan. Dari 11 sampel suara terdapat 10 data yang

identik, Hasil ini didapatkan dari Hasil jarak terkecil menggunakan metode minkowski.

Analisis tingkat akurasi metode minkowski dalam membandingkan rekaman suara tersangka dengan rekaman suara perbandingan dapat menggunakan tabe confusion matrix dengan data yang didapatkan dari Hasil sorting yang terdapat pada Tabe 8, adapun penjeasan singkat mengenai tabe confusion matrix adaah sebagai berikut:

**Tabel 8. Hasil Confusions matrix**

		Nilai Aktual	
		TRUE	FALSE
Nilai Prediksi	TRUE	TP ( <i>True Positive</i> )	FN ( <i>False Negative</i> )
	FALSE	FP ( <i>False Positive</i> )	TN ( <i>True Negative</i> )

Berdasarkan Tabe 8 (Luque et al., 2019) untuk mencari nilai akurasi menggunakan *confusion matrix* dengan meihat kesesuaian nilai dari Hasil *sorting* adaah sebagai berikut:

1. *True Positive* adaah jumlah dari sampel suara yang identik (data positif yang terkasifikasi benar oeh sistem)
2. *Fase Negative* adaah jumlah sampel suara yang tidak identik (data positif yang terkasifikasi saah oeh sistem)
3. *True Negative* adaah jumlah dari sampel suara yang tidak identik identik (data negatif yang terkasifikasi benar oeh sistem)
4. *Fase Negative* adaah jumlah sampel suara yang identik (data negatif tetapi terkasifikasi saah oeh sistem)

Untuk menguatkan kesesuaian nilai pada tabe 4.7 dengan nilai dari Hasil *sorting* maka penjabaran nilai aktua dan nilai prediksi adaah sebagai berikut:

- a. Nilai aktua yaitu nilai sebenarnya baik *True* maupun *Fase*, pada Penelitian ini sampel suara yang digunakan sebanyak 11 sampel suara, dari 11 data tersebut diambil 5 data suara tersangka sebagai perbandingan dan 6 data sebagai suara tersangka yang asi.
- b. Nilai prediksi yaitu Hasil yang didapatkan dari program dengan niainya positif dan negatif. Hasil akhir dari perbandingan antara sampel rekaman suara tersangka dengan sampel rekaman

suara pembanding menghasilkan 10 sampel suara yang identik dan 1 sampel suara tidak identik.

Setelah mengetahui nilai akurasi dan nilai prediksi, selanjutnya nilai tersebut akan dihitung menggunakan rumus:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

Dengan menggunakan rumus di atas maka nilai akurasi yang didapatkan adalah sebagai berikut:

$$Akurasi = \frac{10 + 11}{10 + 11 + 0 + 1} = 0.95$$

Berdasarkan nilai dari akurasi yang didapatkan sebesar 0.95 maka dapat disimpulkan nilai akurasi dari Penelitian ini sebesar 95% yang artinya kinerja metode Minkowski dalam membandingkan antara rekaman suara tersangka dengan suara pembanding lebih baik dibandingkan dengan perbandingan rekaman suara menggunakan metode Eucidean pada Penelitian yang dilakukan oleh (Umar et al., 2019) yang menghasilkan tingkat akurasi sebesar 40%.

Dari implementasi metode Minkowski dapat dianalisis kelebihan dan kekurangan dari proses pengujian dengan membandingkan antara rekaman suara tersangka dan rekaman suara peaku antara lain.

1. Kelebihan :
  - a. Tingkat akurasi yang dihasilkan dengan metode Minkowski lebih tinggi dibandingkan metode Eucidean.
  - b. Nilai jarak yang diperoleh melalui metode Minkowski lebih kecil dibandingkan nilai jarak dengan metode Eucidean.
2. Kekurangan :
 

Persamaan Minkowski hanya menggunakan nilai jarak atau nilai power (p) = 3, dan disesuaikan secara manual

## SIMPULAN

Adapun kesimpulan yang dapat diambil dari Penelitian ini sesuai dengan rumusan masalah adalah sebagai berikut:

1. Tingkat akurasi kinerja algoritma KNN dengan rumus Minkowski dalam

membandingkan sampel rekaman suara tersangka dengan sampel suara pembanding mendapatkan tingkat akurasi sebesar 95% artinya terdapat peningkatan nilai akurasi yang didapatkan metode KNN dengan rumus Minkowski dibandingkan dengan menggunakan rumus Eucidean dengan tingkat akurasi sebesar 40%.

2. Tingkat akurasi yang didapatkan dari metode Minkowski dalam membandingkan rekaman suara tersangka dengan suara pembanding menggunakan tabel *confusion matrix* dengan menghitung nilai prediksi dan nilai aktual dimana untuk nilai prediksi didapat dari *output sistem* dan nilai aktual didapatkan dari penggunaan jumlah sampel suara yang sebenarnya

## DAFTAR PUSTAKA

- Ahmad, J., Fiaz, M., Kwon, S., Sodanil, M., Vo, B., & Baik, S. W. (2016). *Gender Identification using MFCC for Telephone Applications - A Comparative Study*. <https://doi.org/https://doi.org/10.48550/arXiv.1601.01577>
- Ali, N., Neagu, D., & Trundle, P. (2019). Evaluation of k-nearest neighbour classifier performance for heterogeneous data sets. *SN Applied Sciences*, 1(12), 1559. <https://doi.org/10.1007/s42452-019-1356-9>
- Azwar, M., Hidayat, S., & Yudha, F. (2021). Teknik Audio Forensik Dengan Metode Minkowski Untuk Pengenalan Rekaman Suara Pelaku Kejahatan. *Cyber Security Dan Forensik Digital*, 4(1), 1–12. <https://doi.org/10.14421/csecurity.2021.4.1.2372>
- Dave, N. (2013). Feature Extraction Methods LPC, PLP and MFCC In Speech Recognition. *International Journal for Advance Research in Engineering and Technology*, 1(Vi), 1–5.

- Deva, B. S., & Mardianto, I. (2019). Teknik Audio Forensik Menggunakan Metode Analisis Formant Bandwidth, Pitch dan Analisis Likelihood Ratio. *ULTIMATICS*, 10(2), 67–72. <https://doi.org/10.31937/ti.v10i2.936>
- Handayani, I. (2019). Application of K-Nearest Neighbor Algorithm on Classification of Disk Hernia and Spondylolisthesis in Vertebral Column. *Indonesian Journal of Information Systems*, 2(1), 57–66. <https://doi.org/10.24002/ijis.v2i1.2352>
- Hidayat, R. (2022). Frequency Domain Analysis of MFCC Feature Extraction in Children's Speech Recognition System. *JURNAL INFOTEL*, 14(1), 30–36. <https://doi.org/10.20895/infotel.v14i1.740>
- Imario, A., Sudiharto, D. W., & Ariyanto, E. (2017). Uji Validasi Suara Berbasis Pengenalan Suara (Voice Recognition) Menggunakan Easy VR 3.0. 801–806.
- Iswanto, I., Tulus, T., & Sihombing, P. (2021). Comparison of Distance Models on K-Nearest Neighbor Algorithm in Stroke Disease Detection. *Applied Technology and Computing Science Journal*, 4(1), 63–68. <https://doi.org/10.33086/atcsj.v4i1.2097>
- Li, Q., Yang, Y., Lan, T., Zhu, H., Wei, Q., Qiao, F., Liu, X., & Yang, H. (2020). MSP-MFCC: Energy-Efficient MFCC Feature Extraction Method with Mixed-Signal Processing Architecture for Wearable Speech Recognition Applications. *IEEE Access*, 8, 48720–48730. <https://doi.org/10.1109/ACCESS.2020.2979799>
- Luque, A., Carrasco, A., Martín, A., & de las Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, 91, 216–231. <https://doi.org/10.1016/j.patcog.2019.02.023>
- Maher, R. C. (2018). *Principles of Forensic Audio Analysis*.
- Mansyur, M., & Manurung, R. A. P. (2017). Alat Bukti Rekaman Suara Dalam Pembuktian Perkara Tindak Pidana Korupsi. *Jurnal Komunikasi Hukum (JKH)*, 3(1), 105. <https://doi.org/10.23887/jkh.v3i1.9246>
- Mataram, A. S. S., & Sugiantoro, B. S. B. (2009). Analisis Rekaman Suara Voice Changer dan Rekaman Suara Asli Menggunakan Metode Audio Forensik. ... *Journal of Networking and Security (IJNS ...)*, January. <http://ijns.org/journal/index.php/ijns/article/view/39>
- Nishom, M. (2019). Perbandingan Akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada Algoritma K-Means Clustering berbasis Chi-Square. *Jurnal Informatika: Jurnal Pengembangan IT*, 4(1), 20–24. <https://doi.org/10.30591/jpiti.v4i1.1253>
- Putra, D. K., Triasmoro, I. I., Atmaja, R. D., Iwut, I., & Atmaja, R. D. (2017). Simulasi Dan Analisis Speaker Recognition Menggunakan Metode Mel Frequency Cepstrum Coefficient (MFCC) dan Gaussian Mixture Model (GMM). *EProceedings of Engineering*, 4(2), 1766–1772. <http://libraryproceeding.telkomuniversity.ac.id/index.php/engineering/article/view/487/460>
- Renza, D., Ballesteros L., D. M., & Lemus, C. (2018). Authenticity verification of audio signals based on fragile watermarking for audio forensics. *Expert Systems with Applications*, 91, 211–222. <https://doi.org/10.1016/j.eswa.2017.09.003>
- Siriwardena, Y. M., Attia, A. A., Sivaraman, G., & Espy-Wilson, C.

(2022). *Audio Data Augmentation for Acoustic-to-articulatory Speech Inversion using Bidirectional Gated RNNs*.

<http://arxiv.org/abs/2205.13086>

Umar, R., Riadi, I., Hanif, A., & Helmiyah, S. (2019). Identification of speaker recognition for audio forensic using k-nearest neighbor. *International Journal of Scientific and Technology Research*, 8(11), 3846–3850.