

TEKNIK AUDIO FORENSIK DENGAN METODE MINKOWSKI UNTUK PENGENALAN REKAMAN SUARA PELAKU KEJAHATAN

Muhammad Azwar¹, Syarif Hidayat², Fietyata Yudha³

^{1,2,3}Jurusan Informatika Program Magister, Fakultas Teknologi Industri,
Universitas Islam Indonesia, Yogyakarta, Indonesia
Email: ¹muha.azwar@gmail.com, ²syarif@fti.uui.ac.id, ³yudha@uui.ac.id

(Naskah masuk: 04 Maret 2021, diterima untuk diterbitkan: 31 Mei 2021)

Abstrak

Tindak pidana yang dilakukan pelaku kejahatan sedikit tidaknya terdapat barang bukti digital yang ditinggalkan berupa rekaman suara yang dihasilkan dari percakapan menggunakan telepon, algoritma dalam menganalisis suara rekaman banyak beredar di internet salah satunya algoritma KNN yang biasanya digunakan untuk klasifikasi, identifikasi dan prediksi. Penelitian ini menggunakan algoritma tersebut dengan metode Minkowski untuk melakukan pengenalan sampel suara rekaman percakapan pelaku kejahatan dengan sampel suara tersangka. Melibatkan dua responden yang berperan sebagai tersangka dan tiga responden berperan sebagai pelaku. Masing-masing responden akan melakukan *dubbing* (perekaman suara) dan dipotong menjadi sebelas bagian sebagai data latih dan data uji. sampel suara percakapan akan di *extract* menggunakan teknik MFCC untuk mendapatkan nilai *spectrum* yang selanjutnya akan diproses menggunakan algoritma KNN dan metode *minkowski* menggunakan *Python*. Hasil dari penelitian ini adalah dari sebelas sampel suara tersangka pertama ada beberapa sampel suara yang identik dengan sampel suara pelaku dengan tingkat akurasi sebesar 0.63 terlihat dari jarak terkecil yang dihasilkan karena sampel suara pelaku yang mirip dengan sampel suara tersangka pertama diperankan oleh responden yang sama. Sedangkan tingkat akurasi yang didapatkan dari tersangka kedua sebesar 0.18 karena ada dua sampel suara pelaku yang mendekati kemiripan dengan suara tersangka kedua namun diperankan oleh responden yang berbeda, algoritma KNN dengan metode *Minkowski* dapat digunakan untuk melakukan pengenalan rekaman suara pelaku dengan tersangka karena menghasilkan jarak terkecil yang mendekati kemiripan sehingga barang bukti berupa rekaman suara dapat dipertanggungjawabkan dalam persidangan.

Kata kunci: *python, ekstraksi, Minkowski, spectrum, dubbing.*

AUDIO FORENSICS TECHNIQUE USING MINKOWSKI METHOD FOR VOICE RECOGNITION OF THE CRIME SUSPECT

Abstract

Criminal acts committed by criminals have at least digital evidence left in the form of sound recordings resulting from conversations using the telephone, algorithms in analyzing recorded sound circulating on the internet, one of which is the KNN algorithm which is usually used for classification, identification and prediction. This study uses the algorithm with the Minkowski method to identify voice samples recorded by the perpetrators of the conversation with the suspect's voice samples. Involving two respondents who acted as suspects and three respondents as perpetrators. Each respondent will do a dubbing (voice recording) and cut it into eleven parts as training data and test data. The speech voice sample will be extracted using the MFCC technique to get the spectrum value which will then be processed using the KNN algorithm and the Minkowski method using Python. The results of this study are that of the eleven voice samples of the first suspect, there are several sound samples that are identical to the perpetrator's voice sample with an accuracy of 0.63 seen from the smallest distance produced because the voice sample of the offender which is similar to the first suspect's voice sample is played by the same respondent. While the level of accuracy obtained from the second suspect is 0.18 because there are two samples of the perpetrator's voice that are close to the similarity of the second suspect's voice but played by different respondents, the KNN algorithm with the Minkowski method can be used to recognize the voice recording of the perpetrator and the suspect because it produces the smallest distance close to the similarity so that evidence in the form of sound recordings can be accounted for in court.

Keywords: *python, extraction, minkowski, spectrum, dubbing.*

1. PENDAHULUAN

Kasus pembunuhan Jamal Khashoggi pada tanggal 2 Oktober 2018 di Istanbul Turki, ditemukan sebuah rekaman suara percakapan pelaku yang kemudian dijadikan sebagai barang bukti digital yang selanjutnya dianalisis oleh tim investigator untuk mengenali pelaku kejahatan dalam pembunuhan berencana terhadap Jamal Khashoggi. hal ini menjadi salah satu tantangan investigator dalam melakukan proses identifikasi atau analisis barang bukti yang didapat (Tribunnews, 2019).

Mengidentifikasi seseorang dengan suaranya adalah sifat seorang manusia, sebagian besar diterima begitu saja dalam interaksi atau komunikasi antar sesama. Berbicara dengan seseorang melalui telepon biasanya dimulai dengan mengidentifikasi siapa yang berbicara dan, setidaknya dalam kasus pembicara yang akrab, verifikasi subjektif oleh pendengar bahwa identitasnya benar dan percakapan dapat dilanjutkan. Sistem pengenalan suara otomatis telah muncul sebagai sarana penting untuk memverifikasi identitas dalam banyak aplikasi serta dalam interaksi bisnis umum, forensik, dan penegakan hukum. Seorang ahli yang terlatih dalam pengenalan speaker forensik dapat melakukan tugas ini dengan lebih baik dengan memeriksa seperangkat karakteristik akustik, prosodi, dan linguistik dalam pendekatan umum yang disebut dengan mendengarkan terstruktur. Teknik-teknik dalam pengenalan *speaker* forensik telah dikembangkan selama bertahun-tahun oleh para ilmuwan dan ahli bahasa forensik untuk membantu mengurangi bias potensial atau pemahaman sebelumnya tentang validitas sampel audio yang tidak diketahui dan *template* referensi dari tersangka potensial (Hansen and Hasan, 2015).

Suara adalah bentuk dasar dari metode komunikasi yang efisien bagi seseorang untuk berinteraksi satu sama lain. komunikasi manusia didominasi oleh ucapan dan pendengaran informasi yang paling cepat ditransfer dari satu orang ke orang lain selalu dilakukan dengan ucapan, pengenalan suara adalah proses mengenali pembicara berdasarkan karakteristik yang terkandung dalam gelombang suara (Umar et al., 2019).

Rekaman audio memberikan beberapa keuntungan potensial untuk penyelidikan dibandingkan dengan film, video dan saksi mata seperti kemampuan mengumpulkan informasi dari semua sumber. rekaman audio menyediakan sebuah sekuensial waktu merekam dari acara sebagai sebuah objektif pengamatan. Audio forensik adalah cabang dari bidang ilmu forensik. ilmu forensik umumnya mengacu pada evaluasi bukti yang akhirnya dapat digunakan di pengadilan atau sebagai bagian dari beberapa penyelidikan formal lainnya. Forensik audio yang mengacu pada akuisisi, analisis dan interpretasi rekaman audio sebagai bagian dari investigasi, seperti dalam persiapan untuk pengadilan perdata atau pidana atau sebagai bagian dari

investigasi kecelakaan atau insiden lain yang melibatkan bukti audio (Maher and Hoerr, 2018)

Algoritma KNN (*K-Nearest Neighbor*) adalah salah satu algoritma yang paling umum digunakan dalam melakukan klasifikasi data. KNN (*K-Nearest Neighbor*) mengklasifikasikan sampel yang tidak diketahui berdasarkan klasifikasi yang diketahui dari tetangganya. Setiap sampel perlu diklasifikasikan sehingga mirip dengan sampel yang ada pada sekitarnya. Fungsi jarak memainkan peran penting dalam keberhasilan (Rajesh, 2011).

Tugas utama klasifikasi adalah untuk memprediksi label poin data uji dengan menginduksi semua poin data pelatihan. banyak sekali metode klasifikasi yang sudah dikembangkan dalam aplikasi nyata di antaranya klasifikasi KNN (*K-Nearest Neighbor*) telah dianggap sebagai salah satu dari 10 algoritma penambangan data karena kesederhanaan dan efisiensinya. Dengan demikian metode KNN (*K-Nearest Neighbor*) telah berhasil dikembangkan dalam aplikasi penambangan data, seperti klasifikasi, regresi, dan imputasi nilai yang hilang. Gagasan kunci dari metode KNN (*K-Nearest Neighbor*) standar adalah untuk memprediksi label titik data uji oleh aturan mayoritas, yaitu, label titik data uji diprediksi dengan kelas utama dari k titik data pelatihan yang paling mirip di fitur tersebut (Zhang et al., 2017).

Pada algoritma KNN (*K-Nearest Neighbor*) terdapat beberapa metode yang digunakan yaitu *Manhattan*, *Euclidean Minkowski*, dan *Hamming Distance*, pada penelitian yang dilakukan oleh (Umar et al., 2019) tentang identifikasi pengenalan suara, menggunakan algoritma KNN (*K-Nearest Neighbor*) dengan metode *Eucledian* untuk mencari jarak terkecil dalam pencocokan bukti digital berupa rekaman suara pelaku kejahatan. Namun metode tersebut masih kurang bisa mendekati kemiripan rekaman suara pelaku kejahatan. sehingga pada penelitian ini penulis melakukan pengujian suara dengan algoritma KNN dan Metode *Minkowski* sehingga proses uji forensik dapat melakukan pencocokan rekaman sampel suara dengan jarak terkecil yang mendekati kemiripan.

Tujuan penelitian ini adalah melakukan pencocokan sampel suara tersangka dengan suara pelaku dengan menggunakan algoritma KNN dan metode *Minkowski* untuk mendapatkan nilai atau jarak terkecil sebagai pengambilan keputusan untuk menentukan barang bukti yang kuat berupa rekaman suara tersangka yang nantinya bisa dipertanggung jawabkan dalam persidangan.

2. DASAR TEORI

2.1. Suara

Suara dihasilkan melalui proses *Generation* dan *Filtering*. *Generation* adalah proses dimana suara pertama kali diproduksi melalui bergetarnya pita

suara (*vocal cord* atau *vocal ford*) yang berada di laring (*larynx*) untuk menghasilkan bunyi periodik. Bunyi periodik yang bersifat konstan tersebut kemudian di filterisasi melalui *vocal tract* (disebut juga dengan istilah *resonator* suara atau *articulator*) yang mencakup lidah (*tongue*), gigi (*teeth*), bibir (*lips*), langit-langit (*palate*) dan lain-lain sehingga bunyi tersebut dapat menjadi bunyi keluaran berupa bunyi vokal (*vowel*) dan atau bunyi konsonan (*consonant*) yang membentuk kata-kata yang memiliki arti (Muhammad Nuh Al-Azhar, 2012).

2.2. Minkowski Distance

Minkowski Distance atau metrik minkowski adalah metrik dalam ruang vektor bernorma yang dapat dianggap sebagai generalisasi jarak *Euclidean* dan jarak *Manhattan*.

Dalam pengukuran jarak objek menggunakan minkowski distance biasanya digunakan $p = 1$ atau $p = 2$ sesuai dengan jarak manhattan dan jarak euclidean. Dalam kasus pembatasan p yang mencapai tak terhingga, maka disebut sebagai jarak *Chebyshev*. *Minkowski Distance* dapat didefinisikan ke dalam rumus (Nishom, 2019).

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

2.3. MFCC

Langkah pertama dalam sistem pengenalan suara adalah mengekstrak fitur, yaitu mengidentifikasi

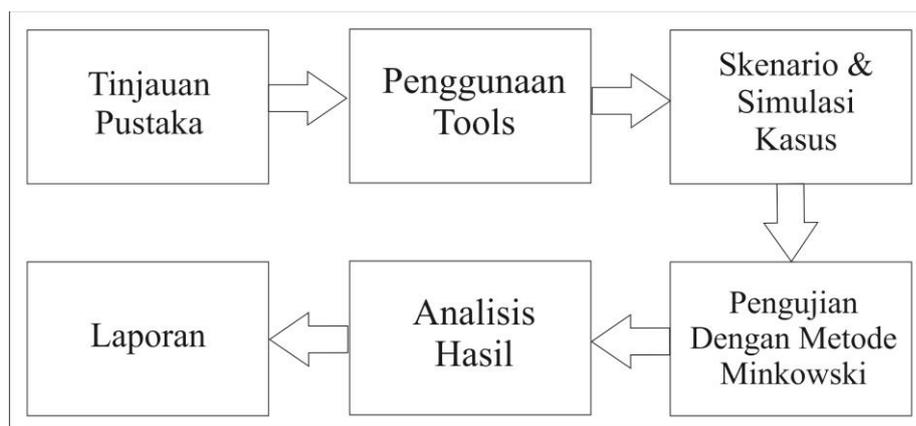
komponen sinyal audio yang baik untuk mengidentifikasi konten linguistik dan membuang semua hal lain yang membawa informasi seperti kebisingan latar belakang dan lain-lain (Lyons, 2015).

Tahapan-tahapan MFCC antara lain sebagai berikut:

- Pre-Emphasis*, *Pre-emphasis* berfungsi untuk menstabilkan nilai magnitudo dari sinyal suara.
- Framing*, berfungsi membagi sinyal suara menjadi beberapa frame dengan panjang sampel tertentu.
- Windowing*, berfungsi meredam noise yang muncul di kedua ujung *frame*.
- Fast Fourier Transform*, berfungsi mengubah sinyal dari domain waktu ke domain frekuensi.
- Mel-Frequency Filter Bank*, Berfungsi untuk mengubah domain frekuensi menjadi domain frekuensi mel.
- Transformasi Non Linear*, berfungsi untuk mengambil nilai logaritma natural dari setiap *Mel-Filter Bank*.
- Discrete Cosine Transform (DCT)*, berfungsi untuk mengambil sinyal suara pada domain frekuensi ke domain waktu sehingga didapatkan koefisien *cepstrum*.

3. METODE PENELITIAN

Dalam menyelesaikan penelitian ini perlu disusun langkah-langkah penyelesaian penelitian secara sistematis yang disebut metodologi penelitian. Metodologi penelitian yang diusulkan untuk menyelesaikan penelitian ini seperti terlihat pada gambar 1.



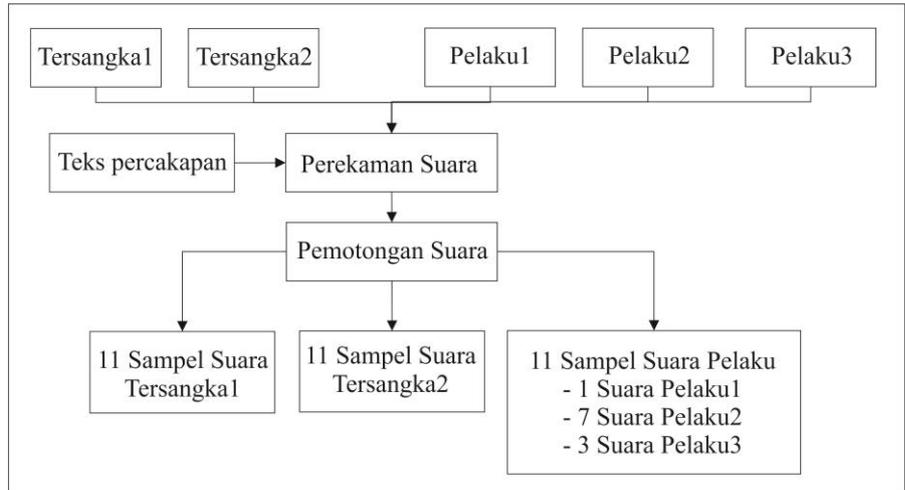
Gambar 1. Alur Metode Penelitian

Tahap pertama adalah tinjauan pustaka dilakukan untuk mengumpulkan bahan-bahan informasi mengenai topik penelitian yang dapat bersumber dari buku, artikel, paper, jurnal, makalah, yang berupa teori, laporan penelitian, atau penemuan sebelumnya dan mengunjungi beberapa situs yang terdapat pada internet terkait teori-teori tentang audio forensik,

barang bukti, sehingga dapat menunjang tujuan akhir yang dilakukan penelitian ini.

Tahapan kedua adalah penggunaan tools dimana proses yang dilakukan oleh penulis dalam mendapatkan data berupa sampel suara tersangka dan

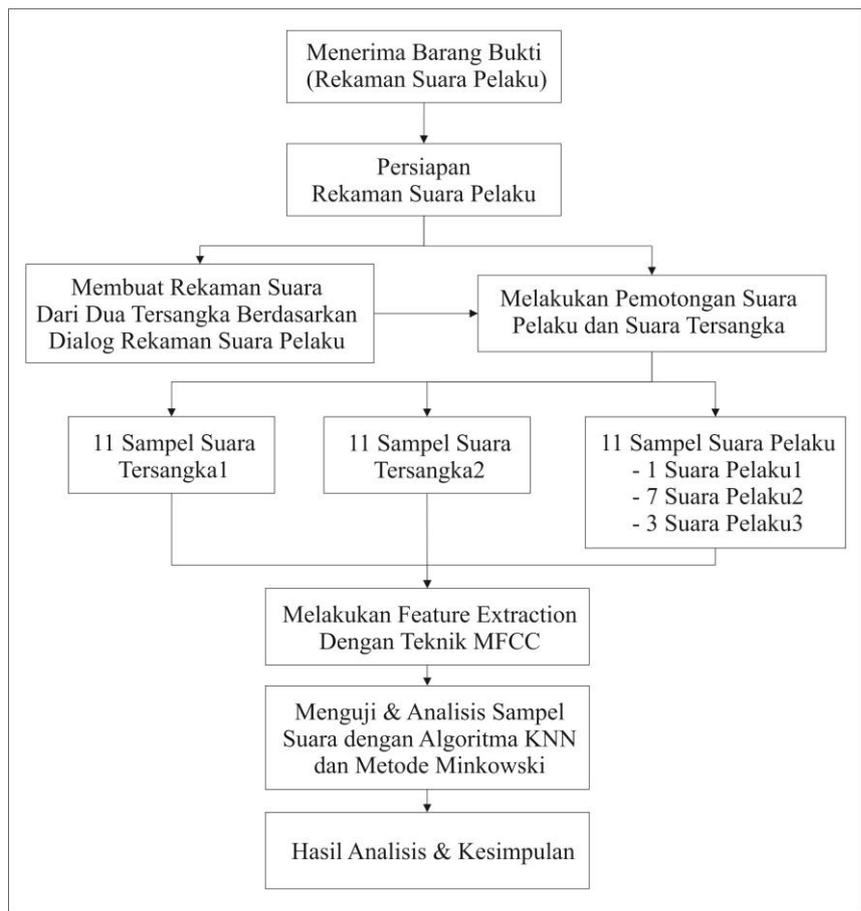
pelaku adapun proses yang dilakukan seperti terlihat pada gambar 2.



Gambar 2. Proses mendapatkan sampel suara

Tahapan ketiga adalah skenario dan simulasi kasus merupakan tahapan yang membuat simulasi kasus dalam pembunuhan Jamal Khashoggi dengan tiga

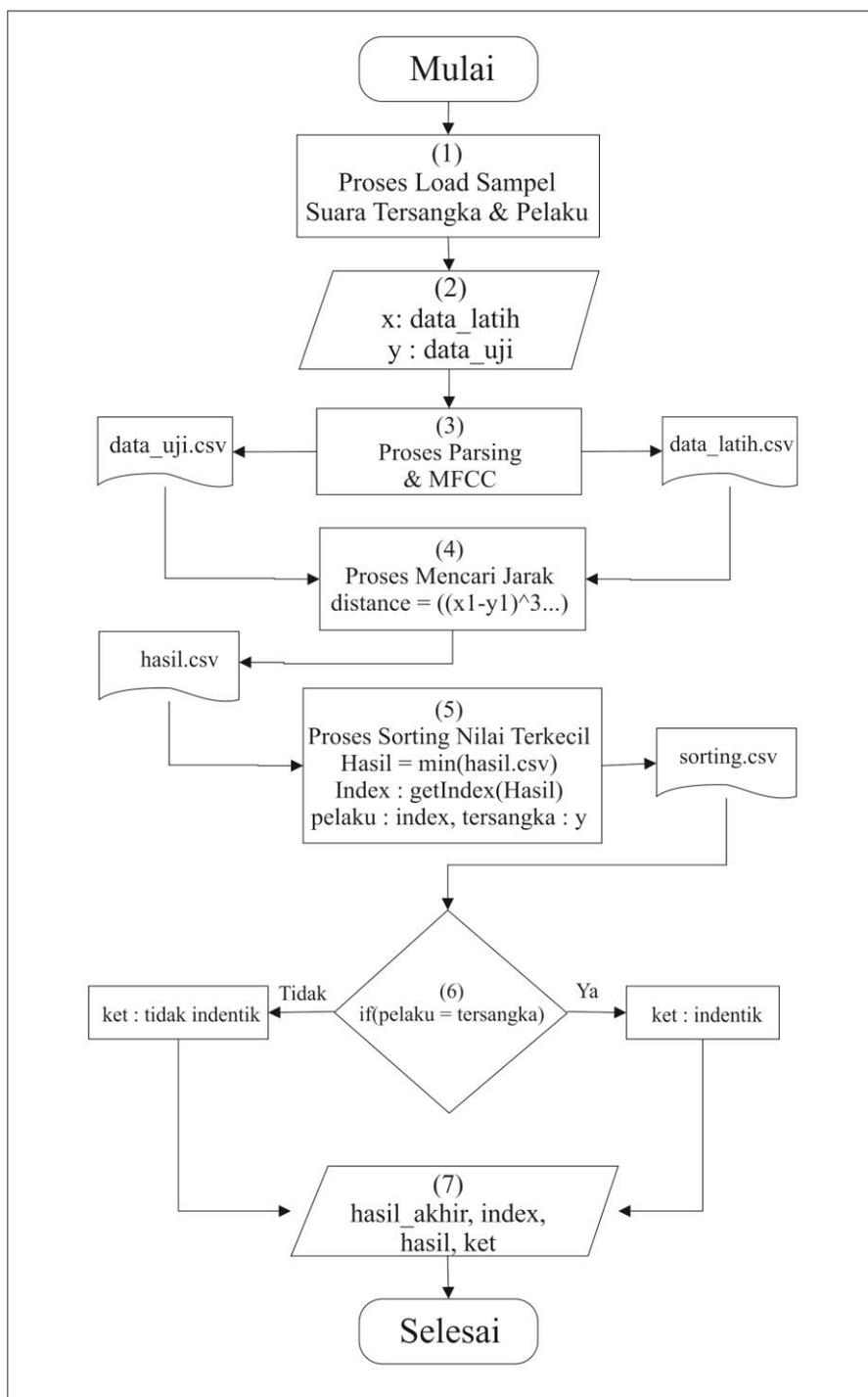
responden sebagai suara pelaku dan tersangka, adapun gambaran pada tahapan ini seperti terlihat pada gambar 3.



Gambar 3. Skenario dan Simulasi Kasus

Tahap keempat adalah pengujian proses pengujian berisi aktivitas yaitu merubah sampel suara tersangka dan pelaku kedalam nilai *spectrum* dengan teknik MFCC dan menghitung jarak

menggunakan algoritma KNN dan metode *Minkowski*, berikut *flowchart* yang menggambarkan secara detail alur dari tahapan pengujian seperti terlihat pada gambar 4.



Gambar 4. Flowchart Pengenalan Suara

Adapun penjelasan Gambar 4 sebagai berikut:

1. Memasukan atau *load* sampel suara tersangka dan pelaku.
2. Deklarasi variabel diaman variabel x berisi data latih yaitu sampel suara pelaku dan variabel y berisi data uji yaitu sampel suara tersangka.
3. Parsing data *array* berupa nilai *frame rate* suara kedalam bentuk *array* satu dimensi dan Melakukan *Extraction Feature* menggunakan

4. Teknik MFCC yaitu merubah sampel suara menjadi nilai *Spectrum*.
5. Proses Perhitungan untuk mencari nilai berupa jarak, setelah proses perhitungan maka hasil berupa jarak akan disimpan ke dalam *file Hasil.csv*
6. Langkah selanjutnya adalah melakukan proses *sorting* dengan *meload file hasil_minkowski.csv* untuk mencari nilai terkecil pada *matrix* *hasil_minkowski.csv*

berdasarkan *index* kemudian hasil *sorting* akan disimpan ke dalam *file* dengan nama Hasil_Sorting.csv

6. Mengambil nama sampel suara berdasarkan *Indexs* yang memiliki nilai terkecil dan enentukan nilai keterangan Indentik dan Tidak Identik dengan membandingkan nama sampel suara dari hasil *sorting* dengan *file* asli adalah sama maka *output* keterangan adalah indentik jika tidak sama maka input keterangan adalah tidak indentik.
7. Menampilkan hasil akhir yang diambil dari *file sorting.csv* dan diambil dari proses kelima.

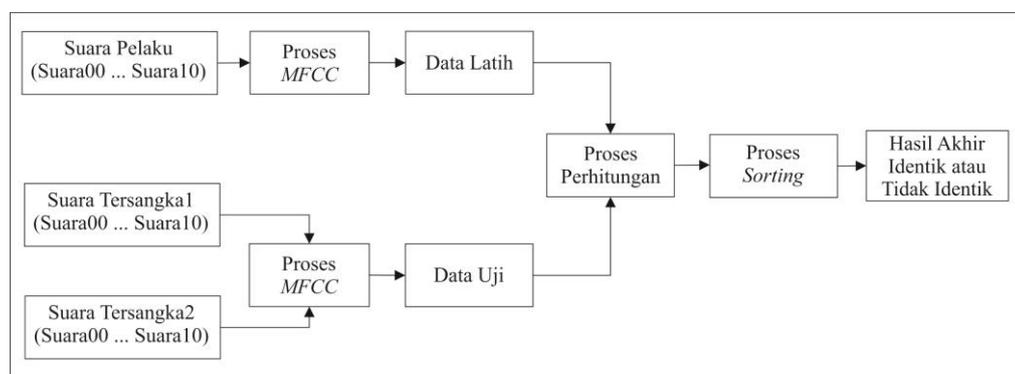
Tahap kelima adalah analisis yaitu menganalisis hasil yang didapatkan dari proses pengujian berupa

nilai jarak dan kecocokan antara sampel suara tersangka dan suara pelaku.

Tahap keenam adalah pembuatan laporan tentang hasil yang didapatkan dari proses analisis rekaman suara tersangka dan pelaku sehingga bisa dijadikan sebagai barang bukti yang sah dan bisa dipertanggung jawabkan dalam persidangan.

4. HASIL DAN PEMBAHASAN

Dalam bab ini membahas tentang percobaan dalam menganalisa sampel suara rekaman pelaku kejahatan menggunakan bahasa pemrograman *Python* seperti terlihat pada gambar 5



Gambar 5. Diagram Proses Percobaan

Diagram pada gambar 5 menunjukkan proses atau langkah-langkah yang dilakukan penulis untuk menganalisis sampel suara rekaman suara pelaku kejahatan.

Penjelasan tahapan proses pengujian:

1. Terdapat dua sampel suara tersangka dengan nama tersangka1 dan tersangka2 serta satu sampel suara pelaku.
2. Masing-masing sampel suara tersangka dan pelaku dipotong menjadi 11. proses ini bertujuan agar mudah dicocokkan dalam proses pengujian dengan algoritma KNN dan *Minkowski*.
3. Sampel Suara pelaku dalam pengujian ini disebut sebagai data latih dan sampel suara tersangka disebut sebagai data uji.
4. Masing-masing sampel suara tersangka dan pelaku akan dikonversi atau melalui tahap *extraction* data menggunakan Teknik MFCC untuk mendapatkan nilai numerik berupa nilai *spectrum*.
5. Selanjutnya data latih dan data uji yang sudah dikonversi akan melalui tahap pengujian untuk mendapatkan hasil berupa jarak.
6. Setelah mendapatkan jarak tahap selanjutnya dilakukan proses *sorting* untuk mengurutkan jarak terkecil sehingga terlihat hasil pencocokan antara sampel suara pelaku dan sampel suara

tersangka yang mendekati kecocokan atau identik dan tidak identik

4.1. Proses Pengambilan Sampel Data

Dalam proses pengambil data ini memuat tahapan pada tinjauan pustaka dan penggunaan tools, dalam melakukan perekaman suara pelaku dan tersangka penulis menggunakan tools yang dapat diunduh di Playstore dengan nama *Voice Recorder* untuk mendapatkan suara rekaman dengan tipe.WAV dan memanfaatkan Handphone Android untuk merekam suara responden sebagai pelaku dan tersangka, adapun teks percakapan yang diubah kedalam suara rekaman percakapan pelaku seperti terlihat pada tabel 1.

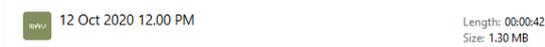
Tabel 1. Sampel Teks Percakapan

Suara	Pelaku	Teks Percakapan
Suara 00	1	Apakah mungkin memasukan mayat ke dalam tas
Suara 01	2	Tidak terlalu sulit, karena saya sudah biasa mengerjakannya
Suara 02	2	Saya tau cara memotong dengan sangat baik.
Suara 03	2	Kamu menunggu di luar biar saya yang menunggu di dalam.
Suara 04	3	Saya akan mengawasi wilayah sekitar
Suara 05	2	Saya biasanya memakai <i>earphone</i> untuk mendengarkan musik.

Suara	Pelaku	Teks Percakapan
Suara 06	2	Dan memotong mayat sambil menghisap rokok
Suara 07	3	Masih ada orang yang masuk tunggu lima menit
Suara 08	3	Sepertinya dia ke arah yang lain situasinya sudah aman dan tidak ada orang yang masuk.
Suara 09	2	Setelah saya memotongnya anda akan membungkusnya ke dalam kantong plastik
Suara 10	2	Memasukkannya ke dalam koper dan membawanya keluar.

4.2. Skenario dan Simulasi Kasus

- Dalam skenario dan simulasi kasus ini dimana proses pertama adalah menerima barang bukti berupa suara percakapan para pelaku pembunuhan Jamal Kasshogi sebanyak satu sampel. Seperti terlihat pada pada gambar 6.



Gambar 6. Barang bukti rekaman suara pelaku

- Setelah barang bukti diterima selanjutnya penulis melakukan perekaman suara dua responden yang berperan sebagai tersangka1 dan tersangka2. Berikut rincian hasil perekaman suara tersangka seperti terlihat pada gambar 7.



Gambar 7. File suara rekaman tersangka

- Setelah mendapatkan sampel suara pelaku dan sampel suara tersangka maka langkah selanjutnya melakukan pemotongan suara sebanyak sebelas bagian bertujuan agar mudah diproses dalam tahap pengujian dalam melakukan pencocokan suara, dalam melakukan pemotongan penulis menggunakan aplikasi *Audacity*.

- Pemotongan suara pelaku sebagai data latihan Untuk mengetahui lebih jelasnya nama responden dalam perekaman sampel suara pelaku yang dijadikan sebagai barang bukti dapat dilihat pada tabel 2:

Tabel 2. Peran Respoden Sebagai Pelaku

Nama Responden	Posisi	Suara pelaku Setelah dipotong
Ahmad Jamaah	Pelaku 1	Suara_00
Khaerudin	Pelaku 2	Suara_01, Suara_02, Suara_03, Suara_05,

Nama Responden	Posisi	Suara pelaku Setelah dipotong
		Suara_06, Suara_09, Suara_10
Munawir	Pelaku 3	Suara_04, Suara_07, Suara_08

Hasil yang didapatkan setelah sampel suara pelaku dipotong menjadi sebelas bagian seperti terlihat pada gambar 8:



Gambar 8. Sampel Suara Pelaku

- Pemotongan suara tersangka sebagai data uji

Untuk mengetahui lebih jelasnya nama responden dalam perekaman sampel suara tersangka dapat dilihat pada tabel 3:

Tabel 3. Peran Responden sebagai tersangka

Nama Responden	Posisi	Suara tersangka Setelah dipotong
Khaerudin	Tersangka 1	Suara_00, Suara_01, Suara_02, Suara_03, Suara_04, Suara_05, Suara_06, Suara_07, Suara_08, Suara_09, Suara_10
Sunardin	Tersangka 2	Suara_00, Suara_01, Suara_02, Suara_03, Suara_04, Suara_05, Suara_06, Suara_07, Suara_08, Suara_09, Suara_10

Hasil yang didapatkan setelah sampel suara pelaku dipotong menjadi sebelas bagian seperti terlihat pada gambar 9:



Gambar 9. Sampel Suara Tersangka

- Sampel suara pelaku dan tersangka yang sudah dipotong maka selanjutnya *melakukan feature extraction* menggunakan teknik MFCC dengan bahasa pemrograman Python untuk mendapatkan nilai *Spectrum* dari masing-masing *file* yang sudah dipotong agar mudah diproses pada bagian pengujian.

4.3. Hasil dan Analisis Kecocokan Sampel Suara Pelaku dan Tersangka

Hasil *extraction* dengan teknik MFCC sampel suara pelaku dan sampel suara tersangka seperti terlihat pada gambar 10.

Suara	N0	N1	N2	N3	N4	N5	N6	N7	N8	
0	-212.6751	121.8672	-40.8379	12.8930	-11.8982	11.0701	-4.1995	2.5099	9.1227	
1	-354.3129	136.2039	6.2251	25.1000	11.3747	-8.4541	-1.1940	-1.7780	14.4264	
2	-352.6378	140.5218	9.7760	20.5039	4.5805	-6.8049	-5.0776	-2.5020	9.1758	
3	-353.9185	147.2373	11.5496	22.3316	11.0282	-10.3986	-2.2880	-5.4667	2.0661	
4	-317.2667	147.4409	-10.0249	28.7917	1.2714	4.3623	-2.0716	5.9220	8.1348	
5	-328.2464	141.1319	7.9278	18.4244	5.1278	-10.9873	-7.1575	5.1737	1.4256	
6	-347.8517	144.2545	12.3964	11.6466	3.0026	-4.4044	-3.0856	2.5670	-0.3276	
7	-333.3961	163.8985	-2.6312	21.8317	2.6949	7.6708	-1.3059	-0.3635	-0.1793	
8	-320.8038	146.8308	1.5115	30.9528	1.0644	4.4434	-4.4135	8.2882	3.2492	
9	-298.3864	138.0110	0.1365	5.1157	-5.4285	-2.0978	-16.5018	-1.5564	-1.6767	
10	-297.0852	150.0441	-2.2450	-5.7257	-0.1947	-6.4339	-15.5519	0.7143	-1.1057	
N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19
-9.9647	0.2494	-6.5544	5.7200	-9.9860	-1.5205	5.4357	-7.1651	9.3446	-3.2219	2.3358
1.2383	4.5941	4.5062	2.0633	3.3722	-10.5399	3.7073	-0.9410	1.9711	2.2348	0.9040
9.6001	5.1068	-1.4398	0.3143	4.0563	-9.3936	2.1927	0.1322	0.2285	4.1564	-0.9287
10.0526	1.2122	0.5891	1.0775	4.8849	-10.6372	2.8781	-0.5540	0.6794	3.6094	-1.1351
8.5770	-2.9471	1.8265	4.4382	6.0550	-4.8583	3.4577	8.2951	-6.5753	7.6826	-2.1365
9.4608	0.7927	0.8169	2.1047	-0.6487	-8.4314	1.1707	-0.7431	-0.3359	2.7975	-0.4915
9.5664	2.6108	1.4658	2.8916	0.1917	-9.4880	1.9259	-0.7514	-0.0765	3.0623	0.5649
7.1287	-5.1722	1.1386	5.8040	3.1136	-1.8477	0.6372	8.3750	-5.4154	5.8292	-1.4817
8.8997	-3.7703	2.6846	4.5907	2.9982	-2.7050	0.6149	10.3294	-7.6424	5.6383	-0.3662
6.9366	-3.0229	-2.7914	0.5521	0.3945	-13.7936	0.8898	-2.7646	-0.9893	2.6708	0.4745
8.7023	0.0750	-5.5821	-4.0445	0.0564	-11.8142	-0.0852	-0.8715	0.0974	0.4691	1.3991

Gambar 10. Hasil MFCC sampel suara pelaku

Hasil MFCC suara pelaku adalah hasil yang didapatkan dari sebelas sampel pelaku yang di *Extract* dimana *Index* 0 – 10 adalah nama suara sedangkan hasil dari *extraction*

outputnya sebanyak 20 yaitu N0 – N19 dan menjadi kriteria dalam proses perhitungan nanti.

Suara	N0	N1	N2	N3	N4	N5	N6	N7	N8
0	-281.6080	163.9904	-9.7306	3.0962	0.9657	4.3333	1.0551	14.6810	10.6388
1	-294.9082	154.5125	-6.4379	9.0448	4.4440	-3.0540	3.5194	15.7176	10.4793
2	-294.0775	169.1063	-18.0304	5.0880	-3.1037	0.1398	3.2237	16.0536	16.3041
3	-323.3325	167.4116	-5.7781	20.6560	0.9343	-0.8865	10.6472	7.6678	11.0405
4	-299.1239	156.4790	-12.8393	16.0451	1.7169	0.2834	8.1232	15.4286	9.0920
5	-307.8640	156.3210	-7.1561	15.0955	3.1702	0.2785	6.0130	13.5827	11.7071
6	-302.1855	167.8687	-13.7541	12.5092	-4.2623	2.6219	7.4042	10.6484	13.0043
7	-329.7350	163.0683	-3.6914	14.7756	10.0977	7.2412	7.6111	8.2970	10.7715
8	-300.3516	155.3746	-11.1743	18.0925	1.2483	0.1386	6.7542	10.2895	8.5298
9	-299.0923	167.2598	-15.4277	7.7934	3.6866	0.0099	7.4583	10.4830	8.5109
10	-310.1244	182.0076	-18.3182	8.0669	5.5058	-3.6078	6.7683	10.2544	12.1382

	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19
3.0207	-0.5864	-3.3382	7.9932	3.9827	-5.3670	12.5632	-2.3399	4.5369	1.7816	-0.9096	
2.3641	-2.5205	-0.2815	9.8524	3.0789	-3.7858	10.8147	-2.7083	5.3377	0.9981	-0.9657	
2.4923	-2.6717	-4.1649	7.3590	5.1827	-2.7449	12.9215	0.0892	5.6627	0.3699	-2.3717	
2.4685	-5.0110	-1.1138	9.2041	3.7615	-2.1810	10.6575	-1.6115	3.2942	0.4640	-3.5302	
2.1076	2.7100	-2.9147	9.7647	4.2478	-3.6510	12.9532	-3.9466	2.0274	1.0959	-1.6553	
2.1985	0.5973	-2.8712	7.0421	4.9607	-0.8590	10.7309	-1.2161	1.4072	0.6887	-0.5209	
3.4618	1.3199	-1.8060	6.0083	3.3910	-2.5469	13.1902	-1.0288	1.4614	1.2982	-1.9131	
4.6499	1.7973	3.0753	9.2733	7.4165	-0.4645	8.2937	-1.5087	3.3555	2.5111	-1.8403	
1.3711	-2.1032	1.8104	6.6470	5.9435	-2.3457	10.3830	-2.8110	1.0995	1.4251	-3.9065	
2.6991	-2.3236	0.9343	7.5604	4.5273	-3.8390	9.4236	-0.0720	-0.2524	2.5236	-3.3451	
5.1901	-3.3824	-1.6576	6.4967	7.2783	-3.5035	9.2110	1.1885	-3.1566	4.1466	-1.9290	

Gambar 11. Hasil MFCC sampel suara tersangka pertama

Hasil MFCC suara tersangka pertama adalah hasil yang didapatkan dari sebelas sampel suara tersangka pertama yang di *Extract*, dimana *Index* 0 – 10 adalah nama suara

sedangkan hasil dari *extraction output* sebanyak 20 yaitu N0 – N19 dan menjadi kriteria dalam proses perhitungan nanti.

Suara	N0	N1	N2	N3	N4	N5	N6	N7	N8
0	-281.6080	163.9904	-9.7306	3.0962	0.9657	4.3333	1.0551	14.6810	10.6388
1	-294.9082	154.5125	-6.4379	9.0448	4.4440	-3.0540	3.5194	15.7176	10.4793
2	-294.0775	169.1063	-18.0304	5.0880	-3.1037	0.1398	3.2237	16.0536	16.3041
3	-323.3325	167.4116	-5.7781	20.6560	0.9343	-0.8865	10.6472	7.6678	11.0405
4	-299.1239	156.4790	-12.8393	16.0451	1.7169	0.2834	8.1232	15.4286	9.0920
5	-307.8640	156.3210	-7.1561	15.0955	3.1702	0.2785	6.0130	13.5827	11.7071
6	-302.1855	167.8687	-13.7541	12.5092	-4.2623	2.6219	7.4042	10.6484	13.0043
7	-329.7350	163.0683	-3.6914	14.7756	10.0977	7.2412	7.6111	8.2970	10.7715
8	-300.3516	155.3746	-11.1743	18.0925	1.2483	0.1386	6.7542	10.2895	8.5298
9	-299.0923	167.2598	-15.4277	7.7934	3.6866	0.0099	7.4583	10.4830	8.5109
10	-310.1244	182.0076	-18.3182	8.0669	5.5058	-3.6078	6.7683	10.2544	12.1382

	N9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19
3.0207	-0.5864	-3.3382	7.9932	3.9827	-5.3670	12.5632	-2.3399	4.5369	1.7816	-0.9096	
2.3641	-2.5205	-0.2815	9.8524	3.0789	-3.7858	10.8147	-2.7083	5.3377	0.9981	-0.9657	
2.4923	-2.6717	-4.1649	7.3590	5.1827	-2.7449	12.9215	0.0892	5.6627	0.3699	-2.3717	
2.4685	-5.0110	-1.1138	9.2041	3.7615	-2.1810	10.6575	-1.6115	3.2942	0.4640	-3.5302	
2.1076	2.7100	-2.9147	9.7647	4.2478	-3.6510	12.9532	-3.9466	2.0274	1.0959	-1.6553	
2.1985	0.5973	-2.8712	7.0421	4.9607	-0.8590	10.7309	-1.2161	1.4072	0.6887	-0.5209	
3.4618	1.3199	-1.8060	6.0083	3.3910	-2.5469	13.1902	-1.0288	1.4614	1.2982	-1.9131	
4.6499	1.7973	3.0753	9.2733	7.4165	-0.4645	8.2937	-1.5087	3.3555	2.5111	-1.8403	
1.3711	-2.1032	1.8104	6.6470	5.9435	-2.3457	10.3830	-2.8110	1.0995	1.4251	-3.9065	
2.6991	-2.3236	0.9343	7.5604	4.5273	-3.8390	9.4236	-0.0720	-0.2524	2.5236	-3.3451	
5.1901	-3.3824	-1.6576	6.4967	7.2783	-3.5035	9.2110	1.1885	-3.1566	4.1466	-1.9290	

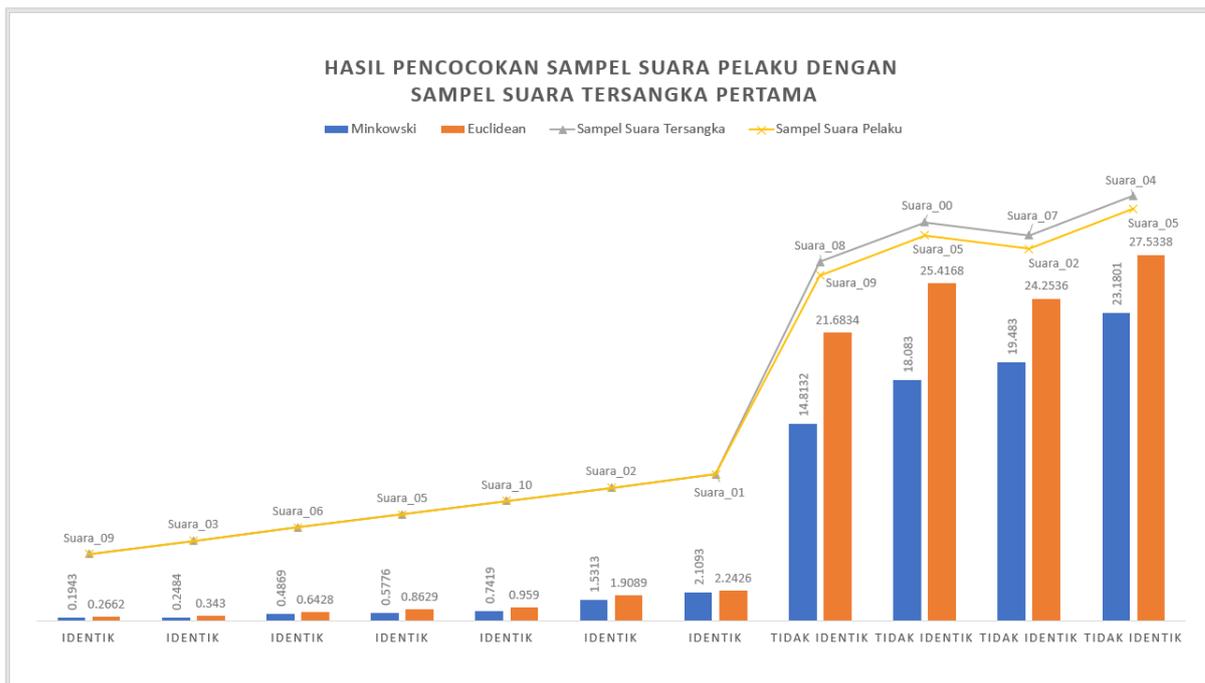
Gambar 12. Hasil MFCC Sampel Suara Tersangka Kedua

Suara	S0	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10
0	122.7740	141.8526	141.2540	144.5165	107.4110	118.9617	138.1655	135.7608	98.6016	89.7198	89.4611
1	23.1257	2.1093	11.4803	16.3063	38.9016	28.1574	20.0392	20.5369	48.5098	58.0311	60.9210
2	22.7497	11.4408	1.5313	10.4421	38.1074	25.2642	12.5868	19.4830	46.3499	55.1099	57.6056
3	25.2821	16.0936	10.8205	0.2484	41.9274	26.9245	14.2715	25.3187	48.5300	57.0202	59.3360
4	22.8611	38.0661	37.1683	40.6299	31.4851	24.1130	36.0333	36.2685	27.1730	30.8093	38.3950
5	18.0830	26.1610	23.7721	26.7183	23.1801	0.5776	20.1014	23.3007	23.0029	31.8568	35.9631
6	22.1433	19.7242	12.6875	14.2034	36.0100	20.5622	0.4869	22.1305	41.8937	50.0920	51.8287
7	29.2181	34.0713	30.5026	29.1085	46.7377	28.1705	26.2357	42.2498	41.0164	41.7342	42.8244
8	22.9036	34.8212	33.0142	36.2316	31.5242	20.7396	31.4469	33.6467	26.5748	33.1231	41.0009
9	38.9338	56.2052	53.9871	57.2083	26.7228	31.8237	50.0446	49.1228	14.8132	0.1943	15.4407
10	41.4752	59.3984	56.8074	59.7379	36.8580	36.1186	51.9701	53.2833	26.5659	15.1634	0.7419

Gambar 13. Jarak minkowski tersangka pertama

Hasil *sorting* merupakan hasil akhir yang didapatkan dalam melakukan pencocokan sampel suara pelaku dengan sampel suara tersangka pertama setelah

mendapatkan nilai jarak dengan metode *minkowski* dan *euclidean* dari nilai *spectrum* dimana hasil akhir seperti terlihat pada gambar 14.



Gambar 14. Hasil Akhir Pencocokan Sampel Suara Pelaku dengan Tersangka Pertama

Pada gambar 14 terlihat bahwa ada tujuh sampel suara tersangka pertama yang cocok dengan suara pelaku dimana sampel suara pelaku dengan garis berwarna kuning berada pada titik yang sama dengan sampel suara tersangka dengan garis berwarna abu-abu sehingga menghasilkan nilai keterangan identik yaitu pada sampel suara_09, suara_03, suara_06, Suara_05, suara10, suara_02 dan suara_01 dan menghasilkan jarak terkecil menggunakan metode

minkowski dibandingkan jarak yang dihasilkan oleh metode *euclidean*.

Tingkat Akurasi Kemiripan Sampel Suara Tersangka Kedua

Pada proses *extraction* sampel suara dengan teknik MFCC, proses tersebut untuk mendapatkan nilai rata-rata dari masing-masing sampel suara yang menghasilkan 20 kriteria berupa nilai *spectrum*, maka dalam perhitungan akurasi digunakan

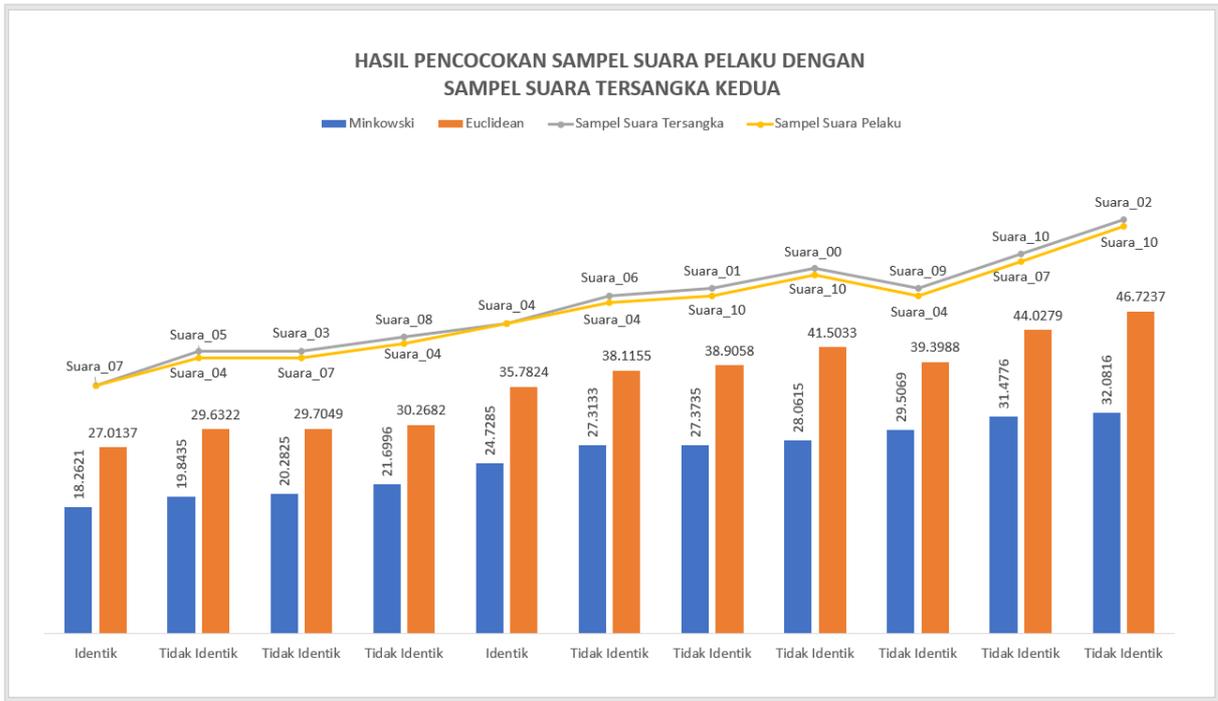
Confusion Matrix dengan jumlah kriteria (N) sebagai jumlah TP (True Positive):

$$Akurasi = \frac{TP \text{ (Jumlah Kriteria *suara identik)}}{\text{total kriteria}} \quad (2)$$

Hasil yang diperoleh dalam menghitung tingkat akurasi tersangka pertama dalam proses pencocokan sampel suara dengan sampel suara pelaku sebagai berikut :

$$Akurasi = \frac{20 * 7}{220} = \frac{140}{220} = 0.63 \quad (3)$$

Untuk tersangka kedua dalam proses mencocokkan dengan sampel suara pelaku menggunakan rumus yang sama dengan yang digunakan pada tersangka pertama. Hasil *sorting* yang merupakan hasil akhir yang didapatkan dari tersangka kedua seperti terlihat pada gambar 15.



Gambar 15. Hasil Akhir Pencocokan Sampel Suara Pelaku dengan Tersangka Kedua

Dari gambar 15 terlihat bahwa terdapat dua sampel suara tersangka yang berada pada titik yang sama pada garis berwarna kuning dengan sampel suara pelaku yang bergaris warna abu-abu sehingga menghasilkan keterangan identik yaitu pada suara_07 dan suara_04.

Tingkat Akurasi Kemiripan Sampel Suara Tersangka Kedua

Tingkat akurasi yang didapatkan dari proses pencocokan sampel suara tersangka kedua dengan suara pelaku berdasarkan hasil akhir pada gambar 15 seperti terlihat pada perhitungan dibawah ini.

$$Akurasi = \frac{20 * 2}{220} = \frac{40}{220} = 0.18 \quad (4)$$

5. KESIMPULAN DAN SARAN

Penelitian dilakukan dengan menggunakan 11 sampel suara pelaku dan 11 sampel suara tersangka dengan tersangka berjumlah dua responden yaitu pada percobaan dalam pencocokan sampel suara tersangka pertama menggunakan metode *minkowski* dan *euclidean* menampilkan tujuh sampel suara yang

sama atau identik dengan sampel suara pelaku pada suara_09, suara_03, suara_06, suara_05, suara_10, suara_02, suara_01 dengan nilai jarak terkecil dibandingkan dengan tersangka kedua suara yang mendekati kemiripan pada suara_07 dan suara_04 dengan nilai jarak yang masih besar dibandingkan nilai jarak yang dihasilkan pada tersangka pertama, sedangkan tingkat akurasi yang ditampilkan pada tersangka pertama dalam pencocokan suara sebesar 0.63 dibandingkan responden kedua yaitu 0.18 artinya tingkat kecocokan pada responden pertama lebih baik dibandingkan responden kedua.

Saran penelitian selanjutnya dalam melakukan pengenalan suara yaitu menggunakan algoritma yang bisa melakukan pengenalan suara dengan mengabaikan awal dan akhir *frame* yang terpotong dari bukti digital berupa suara rekaman sebelum melakukan pengujian sehingga bisa berfokus pada suara percakapan karena pada penelitian ini algoritma yang digunakan dalam melakukan pengujian harus mempunyai *frame* sama panjang antara suara tersangka dengan suara pelaku untuk mendapatkan hasil yang maksimal.

DAFTAR PUSTAKA

- TRIBUNNEWS, 2019. Rekaman Rahasia Pembunuhan Jamal Khashoggi. [online]. Available at: <<http://www.tribunnews.com/international/2019/10/01/rekaman-rahasia-pembunuhan-wartawan-arab-saudi-jamal-khashoggi.html>> [Accessed: 01 Agust 2020]
- HANSEN, J.H.L. and HASAN, T., 2015. *Speaker recognition by machines and humans: A tutorial review. IEEE Signal Processing Magazine*, .
- LYONS, J., 2015. Mel Frequency Cepstral Coefficient (MFCC) tutorial. *Practical Cryptography*.
- MAHER, R.C. and HOERR, E.R., 2018. Audio forensic gunshot analysis and multilateration. In: *145th Audio Engineering Society International Convention, AES 2018*.
- MUHAMMAD NUH AL-AZHAR, 2012. *Digital Forensic Practical Fuidelines For Computer Investigation. BMC Public Health*, .
- 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), pp.20–24.
- RAJESH, D., 2011. Application of Spatial Data mining for Agriculture. *International Journal of Computer Applications*, 15(2).
- UMAR, R., RIADI, I., HANIF, A. and HELMIYAH, S., 2019. Identification of speaker recognition for audio forensic using k-nearest neighbor. *International Journal of Scientific and Technology Research*, 8(11), pp.3846–3850.
- ZHANG, S., LI, X., ZONG, M., ZHU, X. and CHENG, D., 2017. Learning k for kNN Classification. *ACM Transactions on Intelligent Systems and Technology*, 8(3).