

## **Klastering Suara Berdasarkan Gender Menggunakan Algoritma K-Means Dari Hasil Ekstraksi FFT (*Fast Fourier Transform*)**

**Nailul Izzah<sup>1)</sup>**

<sup>1)</sup>Sekolah Tinggi Teknik Qomaruddin Gresik  
[nailul322@gmail.com](mailto:nailul322@gmail.com)

*Diterima : 8 Februari 2018, Direvisi : 4 Maret 2018, Disetujui : 5 Maret 2018*

### **Abstract**

*Speech recognition process is kind of applied Technique of digital sign process that is widely used for many applications, for example technology in the field of telecommunications which isn't only able to provide for serving for sending text data but also it can serve for sending data using sound. From technology development of signal process emerges new idea to make application program by creating new software to display sound signal characteristic based on frequency and highest magnitude. Clustering is part of pattern recognition science which is made for system that can perform into a groups. In this research, the researcher will distinguish between male voice or female voice. Mechanism process uses Collecting voice samples, then the feature extraction using FFT that produce two main features that is maximum value of frequency and maximum value of magnitude. After that, k-means alogarithm process is used for grouping voice of male clusteror female cluster. The result of this research uses 20 training data and 20 testing data that will produce 75% level of accuracy for training data and 100% for testing data.*

**Keywords:** *Speech Recognition, K-Means, FFT, Frequency, Magnitude.*

### **1. PENDAHULUAN**

Perkembangan teknologi saat ini semakin pesat, terutama dalam bidang *Artificial Intelligence*. AI (*Artificial Intelligence*) atau dalam bahasa Indonesia disebut kecerdasan buatan yaitu teori dan pengembangan sistem komputer yang mampu melakukan tugas yang biasanya membutuhkan kecerdasan manusia seperti persepsi visual, *decision making*, translasi antar bahasa dan lain-lain. Istilah *Artificial Intelligence* pertama kali dikemukakan pada tahun 1956 di konferensi Darthmouth [1]. Sejak saat itu penelitian, teori serta prinsip terkait *Artificial Intelligence* terus berkembang. Berbicara tentang *Artificial Intelligence*, Biometrik juga erat kaitannya dengan pengembangan Kecerdasan Buatan. Biometrik adalah ilmu tentang pengenalan identitas manusia berdasarkan keunikan individu. Dimana masing-masing individu akan memiliki ciri yang berbeda dengan individu lain. Selain identitas wajah (*Facial Patterns*), sidik jari, DNA dan iris, suara juga merupakan salah satu identitas manusia.

Setiap manusia mempunyai karakteristik suara berbeda-beda dikarenakan resonansi dalam tenggorokan juga berbeda. Pada penelitian ini mengangkat tema di bidang ilmu Biometrik yang berfokus pada deteksi jenis kelamin berdasarkan suara. Sistem ini akan mengklasifikasi data suara ke dalam kluster laki-laki dan perempuan menggunakan algoritma K-Means berdasarkan hasil ekstraksi FFT (*Fast Fourier Transform*). Proses ekstraksi fitur suara harus dilakukan dengan

-----Jurnal Ilmiah : SOULMATH, Vol 6(1), Maret 2018, Halaman 47-58-----

cermat dan teliti karena hasil dari proses ekstraksi ini yang selanjutnya akan dipakai untuk clustering suara laki-laki atau perempuan. Ada beberapa hal yang mungkin akan menjadi kendala pada saat proses ekstraksi suara yaitu berasal dari variabilitas suara saat kondisi seseorang sakit, emosi, dialog asing dan lingkungan. Lingkungan melatar belakangi kebisingan, gema dan mikrofon. Untuk itu di lakukan proses filtering.

Proses filtering ini sangat penting dilakukan, gunanya adalah memfilter suara untuk menghilangkan *noise-noise* pada suara. Filtering suara dilakukan sebelum proses ekstraksi suara. Adapun proses filter yang digunakan pada penelitian ini adalah *Pre-Emphasis, Frame Blocking dan Windowing*. Adapun pengembangan dari hasil penelitian ini dapat diimplementasikan pada robot atau mesin penjawab otomatis. Dengan adanya pengenalan gender ini diharapkan nantinya robot atau mesin penjawab otomatis akan dapat menggunakan kata sapa dengan tepat berdasarkan jenis kelamin seseorang, seperti: Bapak, Ibu, Mr, Mrs. Sehingga terkesan lebih interaktif.

## 2. METODE

Pada penelitian ini dibuat sistem klastering suara berdasarkan gender dimana suara seseorang akan diklasifikasikan antara suara laki-laki dan perempuan berdasarkan algoritma K-Means. Dimulai dengan mengambil 10 data latih (5 suara laki-laki dan 5 suara perempuan) dan mengambil 10 data random (5 suara laki-laki dan 5 suara perempuan) sehingga total semua adalah 20 data untuk diuji yang terdiri dari 10 suara laki-laki dan 10 suara perempuan. Suara direkam dalam durasi 3 detik dan frekuensi sampling 8000Hz dengan perekaman suara kata “testing”.

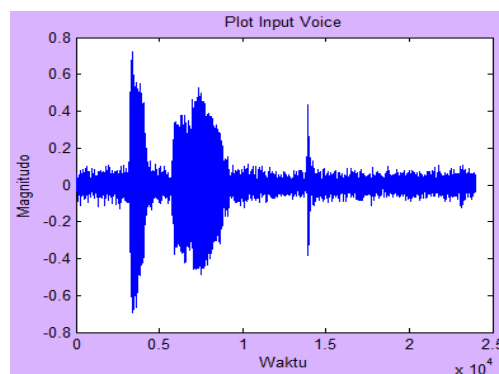
Data dalam penelitian ini adalah data suara diambil dari mahasiswa sekolah tinggi teknik qomaruddin gresik. Dari masing-masing suara akan diekstraksi menggunakan metode FFT (*Fast Fourier Transform*) .Berdasarkan hasil ekstraksi FFT akan diambil magnitudo maksimum dari masing-masing suara. Selanjutnya dilakukan klstering suara menggunakan algoritma KMeans, berdasarkan frekuensi dan magnitudo maksimum akan ditentukan suara tersebut tergolong suara laki-laki atau suara perempuan.

## 3. HASIL DAN PEMBAHASAN

Sistem ini diimplementasikan pada pemrograman MATLAB, yang memanfaatkan GUI matlab sebagai pengembangan desain aplikasinya. Berikut ini akan dijelaskan pengembangan sistem dalam proses ekstraksi fitur suara yang memiliki beberapa tahapan:

### 1. Perekaman Suara

Pada perekaman suara ini frekuensi sampling diset pada 8000 Hz, dimana dalam 1 detik sinyal tersebut terdapat 8000 titik sampling. Sebagai contoh, pada gambar 3 mempunyai sinyal sepanjang 3 detik disampling dengan frekuensi sampling sebesar 8000 Hz, akan menghasilkan Sampling sebanyak 240 titik.



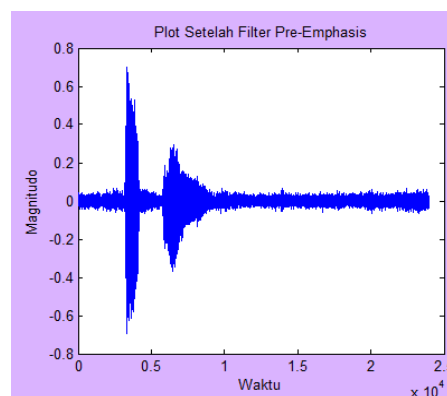
**Gambar 1** Sinyal Suara yang Disampling dengan Frekuensi 8000 Hz

### 2. Proses Filter Suara

Proses filter suara dilakukan menggunakan 3 metode yaitu *pre-emphasis*, *frame blocking* dan *windowing*:

#### a. *Pre-emphasis*

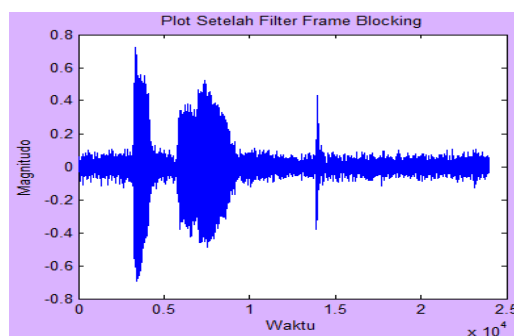
Dalam proses pengolahan sinyal wicara *Pre-Emphasis* filter diperlukan setelah proses sampling. *Pre-Emphasis* ini digunakan untuk meratakan sinyal suara dengan cara meloloskan sinyal frekuensi yang tinggi dan frekuensi rendah akan difilter. Pada gambar 2 adalah hasil filter *Pre-Emphasis* pada pemrograman MATLAB.



**Gambar 2** Sinyal suara 'testing' Setelah difilter *Pre-Emphasis*

b. *Frame blocking*

Pada frame blocking sinyal suara akan dipotong-potong dalam slot-slot waktu tertentu. Proses ini dilakukan agar sinyal suara yang dihasilkan memenuhi 2 syarat yaitu linear dan *time invariant*.



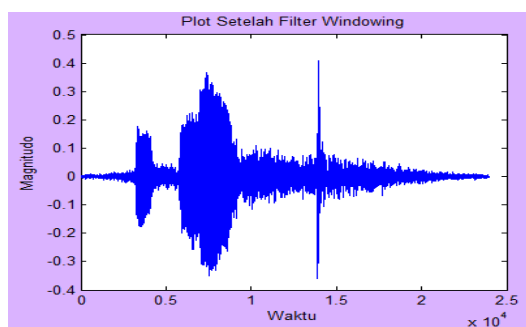
**Gambar 3** Sinyal suara ‘testing’ Setelah difilter *Frame Blocking*

Pada gambar 3 dilakukan filter *Frame Blocking* dengan, Frekuensi sampling = 8000 Hz. Berarti tiap satu detik terdiri dari 8000 sampel. Disampling tiap 200 ms = 0.02 detik

Jadi jumlah sampel tiap frame =  $(8000 \times 0.02) / 1 = 160$  sampel

c. *Windowing*

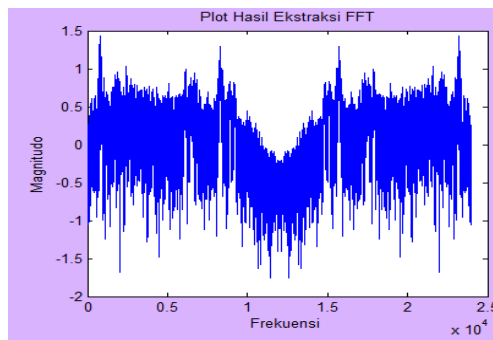
Sinyal suara yang dipotong-potong menjadi beberapa frame akan menyebabkan kesalahan data pada proses *fourier transform*. *Windowing* diperlukan untuk mengurangi efek diskontinuitas dari potongan-potongan sinyal. Sehingga jika diimplementasikan pada MATLAB akan menghasilkan gambar 4 sebagai berikut :



**Gambar 4** Sinyal suara ‘testing’ Setelah difilter *Hamming Window*

### 3. Ekstraksi FFT (*Fast Fourier Transform*)

FFT (*Fast Fourier Transform*) merupakan algoritma untuk memecahkan DFT (*Discrete Fourier Transform*) secara cepat dan efisien. Pada gambar 5 merupakan implementasi dari ekstraksi fitur suara pada pemrograman MATLAB.



**Gambar 5** Sinyal suara ‘testing’ Setelah Ekstraksi FFT

#### 4. Klustering Data Latih

Berikut ini adalah perhitungan dari uji coba sistem klustering suara laki-laki dan perempuan berdasarkan hasil ekstraksi fitur ciri FFT (*Fast Fourier Transform*). Mulanya dilakukan perhitungan pada data latih setelah didapat hasil akhir selanjutnya dilakukan perhitungan jarak terdekat antara data uji dengan centroid akhir dari data latih.

**Tabel 1.** Data Latih

No	Subyek	Magnitudo Max	Frekuensi Max
1	LK1	1130.84	1046
2	LK2	991.399	2160
3	LK3	1136.56	1038
4	LK4	549.633	1545
5	LK5	1070.75	2601
6	LK6	1046.4	1109
7	LK7	517.83	861
8	LK8	628.817	1764
9	LK9	1391.52	5652
10	LK10	1269.77	1314
11	PR1	864.728	1631
12	PR2	1073.34	844
13	PR3	1253.57	8174
14	PR4	1056.02	931
15	PR5	1294.25	1826
16	PR6	839.412	8783
17	PR7	1521.68	8310
18	PR8	1363.65	6643
19	PR9	1330.76	9394
20	PR10.	1155.2	6577

Keterangan : LK adalah Laki-laki, PR adalah Perempuan

Berikut adalah proses perhitungan K-Means pada data latih:

1. Penentuan Pusat Awal *cluster* dengan menggunakan rata-rata.

Diambil Data ke-8 sebagai pusat *cluster*

ke-1 : 628.817      1764

Diambil Data ke-18 sebagai pusat *cluster* ke-2 : 1363.65      6643

2. Perhitungan jarak pusat *cluster*

Menghitung jarak setiap data yang ada terhadap setiap pusat *cluster* dengan menggunakan rumus *Euclidian Distance*.

**Tabel 2.** Perhitungan Jarak Pusat *Cluster*

Subyek	C1	C2	Jarak Terpendek
LK1	876.1	5601.8	876.1
LK2	536.92	4498.4	536.919
LK3	885.93	5609.6	885.934
LK4	232.88	5162.6	232.876
LK5	946.51	4052.6	946.506
LK6	776.79	5543.1	776.789
LK7	909.8	5843.5	909.795
LK8	0	4934	0
LK9	3962.1	991.39	991.392
LK10	783.15	5329.8	783.148
PR1	270.82	5036.8	270.819
PR2	1021.8	5806.3	1021.76
PR3	6440.4	1535	1534.95
PR4	936.16	5720.3	936.158
PR5	668.32	4817.5	668.315
PR6	7022.2	2203.3	2203.28
PR7	6606.6	1674.5	1674.47
PR8	4934	0	0
PR9	7662.2	2751.2	2751.2
PR10	4841.7	218.65	218.649

Keterangan:

1. C1 adalah *cluster* laki-laki yang telah dihitung jarak terdekatnya dengan rumus *Euclidian Distance*.
2. C2 adalah *cluster* perempuan yang telah dihitung jarak terdekatnya dengan rumus *Euclidian Distance*.
3. Jarak Terpendek didapat dari jarak yang paling mendekati pusat *cluster* yaitu dengan mengambil nilai terkecil antara C1 dan C2.

Sehingga didapat pengelompokan data yang dapat dilihat pada tabel 3

**Tabel 3.** Pengelompokan Data

No	C1	C2
1	*	
2	*	
3	*	
4	*	
5	*	
6	*	
7	*	
8	*	
9		*
10	*	
11	*	
12	*	
13		*
14	*	
15	*	
16		*
17		*
18		*
19		*
20		*

Keterangan:

Dimana \* adalah *centroid* yang merupakan jarak terdekat dan menjadi letak *cluster*. Sehingga jika \* berada di C1 maka menjadi *cluster* laki-laki dan jika \* berada pada C2 maka suara masuk pada *cluster* perempuan.

## 5. Penentuan Cluster Baru

Mencari pusat *cluster* baru dengan merata-rata data yang masuk pada masing-masing *cluster* sebelumnya , sehingga didapat *cluster* baru yang terlihat pada tabel 4 berikut:

**Tabel 4.** Penentuan Pusat *Cluster* Baru

<i>Cluster</i> baru	
971.5643846	1265.11314
1436.153846	7647.57143

Keterangan:

1. *Cluster* baru magnitudo laki-laki didapat dari rata-rata C1 data ke-  
 $1+2+3+4+5+6+7+8+10+11+12+14+15 = 971.5643846$
  2. *Cluster* baru Frekuensi laki-laki didapat dari rata-rata C2 data ke-  
 $1+2+3+4+5+6+7+8+10+11+12+14+15 = 1436.153846$
  3. *Cluster* baru magnitudo perempuan didapat dari rata-rata C1 data ke-  
 $9+13+16+17+18+19+20 = 1265.11314$
  4. *Cluster* baru Frekuensi perempuan didapat dari rata-rata C2 data ke-  
 $9+13+16+17+18+19+20 = 7647.57143$
- Cluster* baru yang ke-1 971.564385 ; 1436  
*Cluster* baru yang ke-2 1265.11314 ; 7648

**Tabel 5.** Menghitung Ulang Jarak Pusat *Cluster*

Nama	C1	C2	Jarak Terpendek
LK1	421.41	6602.9	421.413
LK2	724.12	5494.4	724.118
LK3	430.99	6610.8	430.987
LK4	435.74	6144.4	435.745
LK5	1169.1	5050.3	1169.06
LK6	335.6	6542.2	335.604
LK7	732.58	6827.6	732.582
LK8	474.3	5917.9	474.298
LK9	4236.7	1999.6	1999.57
LK10	322.25	6333.6	322.255
PR1	222.21	6029.9	222.214
PR2	600.84	6806.3	600.836
PR3	6743.7	526.56	526.555
PR4	512.17	6719.8	512.165
PR5	506.07	5821.6	506.069
PR6	7348	1212.6	1212.61
PR7	6895.8	710.38	710.379
PR8	5221.6	1009.4	1009.39
PR9	7965.9	1747.7	1747.66
PR10	5144.1	1076.2	1076.2

**Tabel 6.** Pengelompokan Data

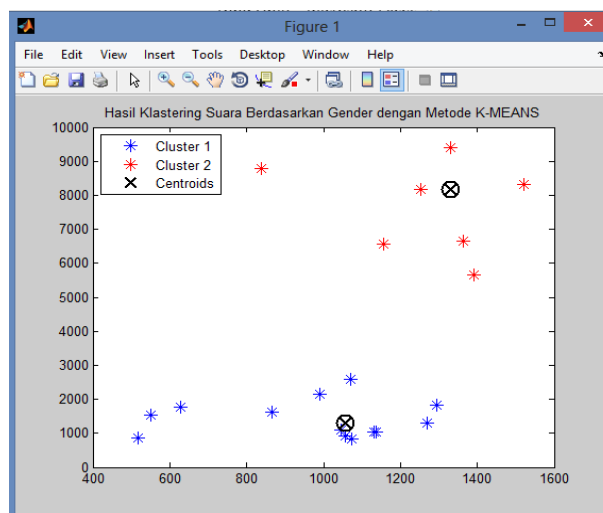
No	C1	C2
1	*	
2	*	



3	*	
4	*	
5	*	
6	*	
7	*	
8	*	
9		*
10	*	
11	*	
12	*	
13		*
14	*	
15	*	
16		*
17		*
18		*
19		*
20		*

Karena *Centroid* tidak berubah, maka perhitungan dianggap selesai. Dari tabel 6 dapat dilihat hasil pengelompokkan data 10 suara laki-laki masuk pada *cluster* laki-laki sebanyak 13 suara. Dan 10 suara perempuan masuk pada *cluster* perempuan sebanyak 7 suara. Sehingga, didapat tingkat akurasi pada data latih klasterisasi ini  $\frac{15}{20} \times 100\% = 75\%$ .

Berikut pada gambar 6 ini adalah klastering suara dilakukan dengan metode K-Means Euclidean Distance yang diimplementasikan pada pemrograman MATLAB:



**Gambar 6.** Hasil Klastering Suara Data Latih

Pada gambar 6 merupakan data latih berasal dari 10 data suara laki-laki dan 10 data suara perempuan yang memiliki lambang \*merah mewakili suara perempuan dan \*biru adalah laki-laki. Dari data latih yang telah dihitung sebelumnya dengan aplikasi yang diterapkan pada pemrograman MATLAB sama-sama didapat suara yang masuk pada *cluster* laki-laki 13 dan *cluster* perempuan 7 suara. Sehingga tingkat keakuratan sama yaitu 75%.

## 6. Proses Data Uji

Pada proses pengujian tidak dilakukan perhitungan mulai awal menentukan pusat *cluster* secara acak melainkan dengan cara menghitung jarak terdekat antara data uji dengan pusat *cluster* akhir dari perhitungan data latih, menghitung jarak terdekat menggunakan rumus *Euclidean Distance*.

Pada tabel 1 merupakan data yang akan digunakan untuk pengujian pengelompokan suara berdasarkan *gender*. Data uji diproses dengan cara menghitung jarak setiap data yang ada pada data uji terhadap pusat *cluster* akhir yang dihasilkan oleh data latih, yaitu dengan menghitung jarak terdekat berdasarkan rumus *Euclidian Distance*. Sehingga frekuensi dan magnitude dari data uji yang dekat dengan *centroid* perempuan akan dimasukkan dalam *cluster* perempuan, begitu juga dengan frekuensi dan magnitude suara laki-laki yang dekat dengan *centroid* laki-laki akan dimasukkan dalam *cluster* laki-laki.

**Tabel 7.** Proses Pada Data Uji

Subyek	C1	C2	Jarak Terpendek
LK1	421.413	6602.9	421.412797
LK2	724.118	5494.4	724.117854
LK3	430.987	6610.8	430.987283
LK4	435.745	6144.4	435.744855
LK5	1169.06	5050.3	1169.06131
LK6	602.973	5732	602.972926
LK7	930.79	5393.6	930.789534
LK8	495.956	5814.6	495.956311
LK9	890.049	5373.6	890.049246
LK10	489.719	6499.4	489.719045
PR1	7348.03	1212.6	1212.60855
PR2	6895.82	710.38	710.378887
PR3	5221.59	1009.4	1009.39252
PR4	7965.95	1747.7	1747.66194
PR5	5144.12	1076.2	1076.19891
PR6	8532.22	2350.3	2350.29866
PR7	6168.02	282.72	282.720537
PR8	7236.01	1036.5	1036.4543

PR9	4742.09	1535.8	1535.7549
PR10	7489.33	1273	1272.99916

Menghitung jarak setiap data yang ada terhadap setiap pusat *cluster* dengan menggunakan rumus *Euclidian Distance*. Lebih jelasnya untuk pengelompokan data uji dapat dilihat pada tabel 8.

**Tabel 8.** Pengelompokan Data

Subyek	C1	C2
1	*	
2	*	
3	*	
4	*	
5	*	
6	*	
7	*	
8	*	
9	*	
10	*	
11		*
12		*
13		*
14		*
15		*
16		*
17		*
18		*
19		*
20		*

Berdasarkan pengelompokan data pada tabel 8 didapat 10 *centroid* masuk pada *cluster* laki-laki dan 10 *centroid* masuk pada *cluster* perempuan. Sehingga, didapat tingkat akurasi pada data uji ini  $\frac{20}{20} \times 100\% = 100\%$ .

#### 4. KESIMPULAN

Berdasarkan hasil implementasi dan pengujian sistem dapat disimpulkan perhitungan jarak terdekat *cluster* dilakukan sebanyak 2 kali perulangan dan didapat magnitude serta frekuensi pusat *cluster* pada suara laki-laki (971.564; 1436.154) dan suara perempuan dengan pusat *cluster* (1265.113; 7647.571), sedangkan akurasi pada data uji 100 % dan data latih 75%.

## 5. DAFTAR PUSTAKA

- [1] Agustini, Ketut, “BIOMETRIK SUARA DENGAN TRANSFORMASI WAVELET BERBASIS ORTHOGONAL DAUBENCHIES,” Gematek Jurnal Teknik Komputer, 2007, Volume 9 Nomor 2.
- [2] Gunawan, Dadang, Juwono, & Filbert Hilman,” PENGOLAHAN SINYAL DIGITAL DENGAN PEMROGRAMAN MATLAB,” Graha Ilmu Yogyakarta, 2012.
- [3] Hadi Putra, Prabowo,” PENGGOLONGAN SUARA BERDASARKAN USIA DENGAN MENGGUNAKAN METODE K-MEAN,” Institut Teknologi Sepuluh Nopember Surabaya, 2012.
- [4] Kurnia Handani, Putri dan Setiawan Arif,” KLASSTERING SUARA BERDASARKAN GENDER DENGAN EKSTRAKSI CIRI BERBASIS DOMAIN WAKTU,” Seminar Nasional Teknologi Informasi dan Komunikasi Terapan Semarang, 2012.
- [5] Musthofa, Ali, “ SISTEM PENGENALAN PENUTUR DENGAN METODE MEL-FREQUENCY WRAPPING,” Universitas Brawijaya Malang, 2007.
- [6] Rismawan, Tedy dan Sri, Kusumadewi, “APLIKASI K-MEANS UNTUK PENGELOMPOKAN MAHASISWA BERDASARKAN NILAI BODY MASS INDEX (BMI) DAN UKURAN KERANGKA,” Seminar Nasional Aplikasi Teknologi Informasi Yogyakarta, 2008.
- [8] Riyanto, Eko, “ SISTEM PENGENALAN PENGUCAP MANUSIA DENGAN EKSTRAKSI CIRI MFCC DAN ALGORITMA JARINGAN SARAF TIRUAN PERAMBATAN BALIK SEBAGAI PENGENALANYA,” JSIB, 2013.
- [9] Riyanto, Sugeng. Purwanto, Agus. Dan Supardi,” ALGORITMA FAST FOURIER TRANSFORM (FFT) DECIMATION IN TIME (DIT) DENGAN RESOLUSI 1/10 HERTZ, “ Prosiding Seminar Nasional Penelitian, Pendidikan, dan Penerapan MIPA Fakultas MIPA, Universitas Negeri Yogyakarta, 2009.
- [10] Sianipar, R. H. 2015. *Pemrograman Matlab*. Yogyakarta: Andi, 2015.
- [11] V. Oppenheim, Alan. S. Willsky, Alan, dan Nawab, S. Hamid, “ SINYAL DAN SISTEM,” Jakarta, Erlangga, 2000.