

**KLASIFIKASI SUARA BERDASARKAN USIA MENGGUNAKAN
MEL FREQUENCY CEPSTRAL COEFFICIENT (MFCC)
DAN K-NEAREST NEIGHBOUR (K-NN)
Classification Of Voice Based On Age Using Mel Frequency
Cepstral Coefficient and K- Nearest Neighbour (K-NN)**

Eva Susanti¹, Sudi Mariyanto Al Sasongko², I Gede Pasek Suta W.³

ABSTRA

Klasifikasi suara berdasarkan usia dibuat dengan tujuan agar komputer mampu mengenali suara laki-laki dan perempuan. Dengan kemampuan komputer yang mampu membedakan suara berdasarkan usia akan memperkuat tingkat suatu sistem keamanan yang menggunakan password dengan suara. Pencocokan password tidak hanya berdasarkan kata saja, namun ditambah dengan pencocokan karakteristik suara sehingga akan lebih aman. Pengenalan suara berdasarkan usia dilakukan dengan teknik ekstraksi ciri sinyal audio dengan menggunakan mfcc. Teknik yang digunakan pada Mel Frequency Cepstral Coefficient diambil berdasar pendekatan pada pendengaran manusia karena itu metode ini sangat baik dalam pengolahan suara manusia.

Pada tugas akhir ini telah dirancang dan direalisasikan suatu sistem yang dapat mengidentifikasi suara manusia, untuk diketahu usianya. Sistem klasifikasi suara ini terdiri dari ekstraksi ciri dan pengklasifikasian suara, dengan metode pengklasifikasiannya adalah K-NEAREST NEIGHBOUR (K-NN). Ciri yang telah didapatkan untuk membedakan suara adalah ada 14 koefisien yang mewakili dari suara yang digunakan pada penelitian ini.

Hasil dari pengklasifikasian suara ini dapat membedakan suara berdasarkan usia, dengan nilai rata-rata akurasi tertinggi 91,11 % dan FNR sebesar 6,49 % serta FPR 16,00 %. Hal ini dipengaruhi oleh nilai k dan banyaknya data yang digunakan. Pada penelitian ini nilai k yang terbaik ada pada k=3 dan nilai yang terburuk ada pada k=13. Secara keseluruhan kinerja sistem dalam mengklasifikasikan suara berdasarkan usia dapat dikatakan berhasil.

Kata Kunci: *Suara; Mel Frequency Cepstral Coefficient; K-NEAREST NEIGHBOUR*

ABSTRACT

Voice classification by age made that computer can identify men and women voices. Computer ability to classify by age would improve security system that using voice password. Password matching not solely based on word, yet improved with voice characteristic matching that will be more secure. Voice recognition by age runs with audio signal characteristic extraction technique using mfcc. Technique that used on Mel Frequency Cepstral Coefficient based on human hearing approach therefore this method is good for processing human voices.

This research devise and realized a system that can identified human voices by age. This voice classification system consists of characteristic extraction and sound classification, with K Nearest Neighbor (K-NN) classification method. There are 14 characteristic coefficient obtained to classifying the voice that represent voice that used in this research.

This voice classification system can classify voice by age, with highest average accuracy value 91,11% and FNR in amount of 6,49% with FPR in amount of 16,00%. This value affected by k value and the amount of data used. In this research, best k value lies on k=3 and the worst value lies on k=13. Overall the system performance to classifying voices by age has succeed.

Key Words: *Voice; Mel Frequency Cepstral Coefficient; K- Nearest Neighbour.*

PENDAHULUAN

Password suara yang sudah banyak diimplementasikan mendeteksi apakah password yang diucapkan sama dengan password yang disimpan untuk membuka suatu sistem keamanan. Namun sebenarnya masih terdapat celah pada sistem keamanan

ini. Ketepatan password tidak cukup menunjukkan bahwa orang yang dimaksud adalah benar. Oleh karena itu dibutuhkan penambahan fitur karakteristik suara yang mampu membedakan warna suara, salah satunya warna suara berdasarkan pada jenis

¹Jurusan Teknik Elektro, Fakultas Teknik, Universitas Mataram Jl. Majapahit 62, Mataram, Lombok – Indonesia
Email : Email : 1.eva.21susanti98@gmail.com, 2.mariyantosas@gmail.com, 3.gpsutawijaya@gmail.com

kelamin dan usianya sehingga mampu mengurangi masalah tersebut (Prasetya, 2008).

Pengenalan Suara Pada Manusia. Suara manusia merupakan bunyi yang berasal dari pita suara dan bergetar manakala mendapat tekanan udara dari rongga dada. Getaran tadi bertemu dengan udara diluar rongga mulut sehingga menimbulkan bunyi yang kemudian dapat ditangkap oleh gendang telinga (Daryanto, 2008).

Suara manusia sendiri memiliki rentang frekuensi antara 20Hz sampai 20KHz. Pengenalan suara dapat diaplikasikan menjadi sistem autentikasi karena setiap pembicara memiliki gelombang suara yang unik. Sebagai contoh jika seseorang mengucapkan sebuah kata, maka polanya akan berbeda dengan pola suara orang lain. Pola suara juga akan berbeda jika dilihat dari jenis kelamin maupun kisaran umur. Jika seseorang hanya mendengarkan suara saja tanpa tahu orangnya, untuk suara dengan pola yang mirip, sebagian besar orang akan sulit mengenali kisaran usianya. Suara dalam bentuk analog harus dikonversi menjadi digital.

Hasilnya adalah sebuah representasi vektor dengan tidak menghilangkan ciri dari suara tersebut. Oleh karena itu, dibutuhkan sebuah teknik ekstraksi ciri untuk mengubah vektor suara menjadi vektor ciri tanpa mengurangi karakteristik suara tersebut.

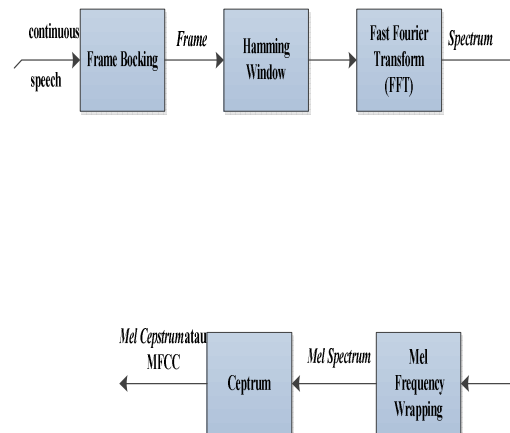
Dalam tugas akhir ini akan dirancang simulasi klasifikasi penutur. Sistem akan mengklasifikasikan suara penutur dari 30 orang penutur yang terdiri dari 5 suara anak perempuan dan anak laki-laki, 5 suara remaja perempuan dan laki-laki serta 5 suara orang tua perempuan dan laki-laki. suara diolah dengan menggunakan software.

Mel Frequency Cepstrum Coefficients (MFCC). *Feature extraction* (ekstraksi ciri) merupakan proses untuk menentukan satu nilai atau vektor yang dapat dipergunakan sebagai penciri obyek atau individu atau pembicara. Dalam pemrosesan suara, ciri yang biasa dipergunakan adalah *mel frequency cepstrum coefficients* (MFCC) yang menghitung koefisien *cepstrum* dengan mempertimbangkan persepsi sistem pendengaran manusia terhadap frekuensi suara. MFCC didasarkan pada variasi yang telah diketahui dari jangkauan kritis telinga manusia dengan frekuensi. MFCC memiliki 2 jenis filter dimana bersifat linear pada

frekuensi dibawah 1000 Hz dan bersifat logaritmik pada frekuensi diatas 1000 Hz. Beberapa keunggulan dari metode ini adalah : (Subramanian, 2004).

- Mampu untuk menangkap karakteristik suara yang sangat penting bagi pengenalan suara, atau dengan kata lain dapat menangkap informasi-informasi penting yang terkandung dalam *signal* suara.
 - Menghasilkan data seminimal mungkin, tanpa menghilangkan informasi-informasi penting yang dikandungnya.
- Mereplikasi organ pendengaran manusia dalam melakukan persepsi terhadap signal suara.

Blok diagram pemrosesan *mel frequency cepstrum coefficients* (MFCC) dapat dilihat pada gambar 2 dibawah ini :



Gambar 1. Blok Diagram Mel Frequency Cepstrum Coefficients

- Frame Blocking.** Pembagian sinyal *audio* menjadi beberapa *frame*. Satu *frame* terdiri dari beberapa sampel tergantung tiap berapa detik suara yang akan di-*sampling* dan berapa besar frekuensi *sampling*-nya (Akbar, 2011).
- Windowing.** Proses *windowing* dilakukan pada setiap *frame* dengan tujuan untuk meminimumkan diskontinuitas antar dua *frame* yang *adjacent*, khususnya pada bagian awal dan akhir. Apabila tidak dilakukan *windowing* maka terjadi diskontinuitas antar *frame* yang akan menyebabkan kehilangan suatu informasi pada suara yang diproses.
- Fast Fourier Transform (FFT).** *Fourier transform* bertujuan mengubah masing-masing *frame* *N* sampel dari domain

waktu menjadi domain frekuensi. *Fourier transform* menguraikan sinyal kedalam suatu fungsi eksponensial pada frekuensi.

4. **Mel Frequency Wrapping.** Proses pengfilteran dari spektrum setiap *frame* yang diperoleh dari tahapan sebelumnya, menggunakan sejumlah *M filter* segitiga dengan tinggi satu. *Filter* ini dibuat dengan mengikuti persepsi telinga manusia dalam menerima suara. Persepsi ini dinyatakan dalam *mel scale* yang mempunyai hubungan tidak linear dengan frekuensi suara. Pemetaan antara frekuensi dalam *hertz* dan *mel scale* adalah linear untuk frekuensi dibawah 1000 hz dan logaritmis untuk frekuensi diatas 1000 Hz. Dari pemetaan tersebut dapat dilihat perbedaan antara frekuensi dan *mel-scale*. Batasan untuk frekuensi *mel-scale* adalah *2Fs* sesuai aturan *nyquise*.

Selanjutnya dibentuk suatu susunan filter yang berisi sejumlah *M filter* segitiga dengan *M filter* segitiga yang digunakan adalah 20. Dari jumlah filter ini kemudian akan dibangkitkan sejumlah *filterbank*. Setelah *filterbank* terbentuk, langkah berikutnya adalah melakukan *mel frequency warping* untuk memperoleh nilai koefisien spektrum *mel (mel spectrum coefficients)*.

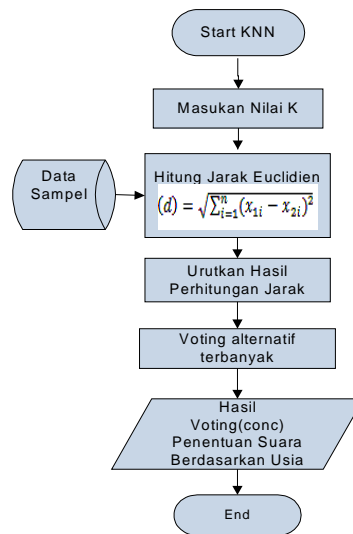
5. **Cepstrum.** *Cepstrum* adalah sebutan kebalikan untuk *spectrum*. *Cepstrum* biasa digunakan untuk mendapatkan informasi dari suatu sinyal suara yang diucapkan oleh manusia. Pada langkah terakhir ini, *mel spectrum* perlu diubah menjadi domain waktu menggunakan *Discrete Cosine Transform (DCT)*. Hasil dari proses ini dinamakan *Mel-Frequency Cepstrum Coefficients (MFCC)*. (Satrya, 2010)

K-Nearest Neighbor (KNN). Klasifikasi *K-Nearest Neighbor* didasarkan pada pembelajaran (*learning*) dengan membandingkan sebuah data uji dengan sejumlah data latih. Data latih terdiri dari *n* atribut. Tiap data merepresentasikan titik pada sebuah ruang berdimensi *n*, dengan begitu semua data latih disimpan di dalam ruang pola berdimensi *n*. Ketika diberikan sebuah data yang tidak diketahui kelasnya, *k-nearest neighbor* akan mencari pola ruang untuk data latih *k* yang terdekat (Yessivirna, 2011).

Data latih *k* ini merupakan *k "nearest neighbor"* dari data yang tidak diketahui tersebut. Kedekatan didefinisikan dengan ukuran jarak seperti *Euclidean distance*, di mana jarak dua data $X_1 = (x_{11}, x_{12}, \dots, x_{1n})$ dan $X_2 = (x_{21}, x_{22}, \dots, x_{2n})$ dinyatakan dalam persamaan 1.

$$dist(X_1, X_2) = \sqrt{\sum_{i=1}^n (x_{1i} - x_{2i})^2} \dots\dots\dots (1)$$

Metode dapat digambarkan dengan *flowchart K-Nearest Neighbor* dibawah ini :



Gambar 2. Flowchart K-Nearest Neighbor (KNN)

Keterangan *flowchart* pada gambar 2 adalah:

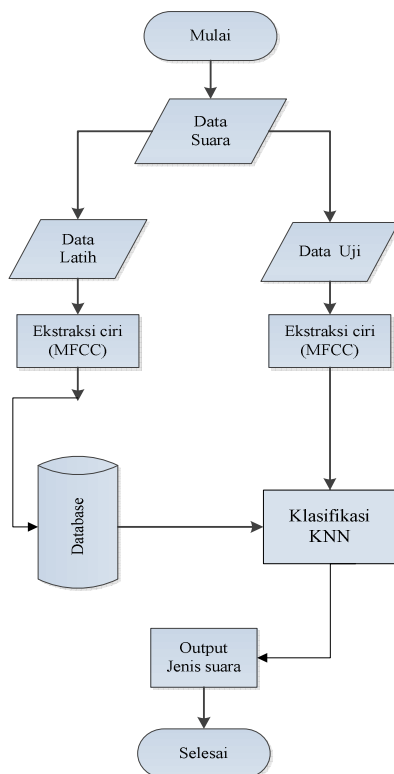
1. *Input* kan nilai *k*. Menginputkan suara *output* dari *mfcc (Mel frequency Cepstral Coefficient)* dan kemudian menginput nilai *k* yaitu *k* yang digunakan adalah ganjil (*k1, k3, k5, k7, k9, k11, k13, k15* dan *k17*).
2. Proses *Perhitungan Jarak*. Proses perhitungan jarak dilakukan berdasarkan fitur dari suara yaitu dari 14 koefisien dari hasil MFCC. Perhitungan jarak dilakukan dengan menggunakan *Euclidean distance* sesuai dengan persamaan 2.13.
3. Proses *Pengurutan (Sorting)*. Proses pengurutan bertujuan untuk memudahkan dalam pencarian tetangga terdekat. Pengurutan dilakukan dengan mengurutkan konten dari *array* berdasarkan jarak dari kecil ke besar.
4. Proses *Voting Alternatif (Weighted voting)*. Proses *weighted voting* dilakukan sesuai dengan nilai *k* yang dimasukkan. Dengan data yang sudah diurutkan berdasarkan jarak terdekat, sistem akan mengambil *k* tetangga terdekat untuk melakukan *voting*

dari tiap kelas anak-anak, remaja dan tua. Selanjutnya, sistem akan memilih jumlah jarak terbesar antara kelas berdasarkan usia sebagai kelas dari data yang diuji. Disini ada proses hitung probabilitas dari peluang id yang dikenali dan *Conclusi* atau kesimpulan dari id yang dikenali milik anak-anak,remaja ataupun tua.

METODOLOGI PENELITIAN

Metode yang digunakan dalam penelitian dibagi menjadi beberapa tahapan, yaitu pengumpulan data, pengolahan data, pengujian sistem, dan analisis data. Pengumpulan data berupa data suara perempuan dan laki-laki yang dibagi 3 katagori usia yaitu anak-anak, remaja dan tua. Pengolahan suara dilakukan setelah pengumpulan data selesai. Pengolahan dan pengujian sistem dilakukan dengan bantuan perangkat lunak. Setelah itu analisis data dengan melihat persentase keandalan sistem.

Perancangan Sistem. Proses perancangan suatu sistem dilakukan penelitian dan penganalisaan tentang sistem yang akan dibangun, berikut ini adalah gambar diagram alur penelitian .



Gambar 3. Diagram alur penelitian

Penjelasan dari gambar diatas :

- 1. Sampel suara.** "Sampel suara" *Input* dari sistem yang dirancang dan direalisasikan adalah sampel suara yang telah direkam dengan frasa " SEMANGAT " dan pada hasil tugas akhir ini diambil beberapa jenis sampel suara.
- 2. Proses Pelatihan (Learning).** Tahapan ini dilakukan beberapa proses dengan menggunakan data latih yang diambil dari dataset suara. Proses-proses yang dilakukan pada proses pelatihan antara lain proses *load* data uji dan data latih, perhitungan jarak *Euclidean*, dan perhitungan *weighted voting*.

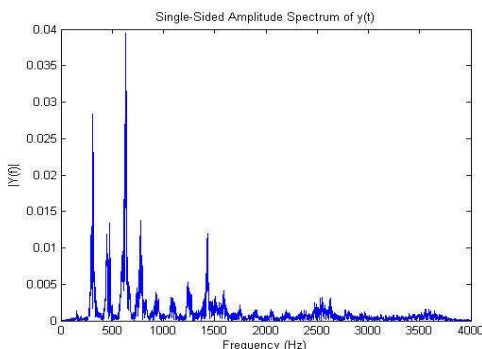
Metode *K- Nearest Neighbor* dan *Euclidian Distance* digunakan dalam dua tahap. Penggunaan yang pertama adalah pada proses pembelajaran atau *learning*, dalam tahapan pengklasifikasian sampel suara. Disini jumlah sampel suara yang digunakan dalam proses pembelajaran adalah 50 sampel suara orang anak-anak, 50 sampel suara orang remaja dan 50 sampel suara tua, dengan total jumlah sampel yang digunakan untuk proses pembelajaran adalah 150 sampel suara. Ekstraksi cirri dari 150 sampel suara tersebut yang akan dijadikan sampel latih pada proses pengklasifikasian.

- 3. Proses Pengujian (Testing).** Tahapan berikutnya yang dilakukan adalah proses pengujian atau *testing* untuk menentukan *output* yang dihasilkan oleh sampel suara yang diuji. Sampel suara yang akan diuji berjumlah 150 pada masing-masing jenis suara. Parameter yang digunakan dalam proses pengujian didapatkan dari proses pembelajaran yang telah dilakukan sebelumnya, yaitu berdasarkan ciri yang telah didapatkan pada masing-masing sampel suara.

- 4. Proses Pengujian (Testing).** Tahapan berikutnya yang dilakukan adalah proses pengujian atau *testing* untuk menentukan *output* yang dihasilkan oleh sampel suara yang diuji. Sampel suara yang akan diuji berjumlah 150 pada masing-masing jenis suara. Parameter yang digunakan dalam proses pengujian didapatkan dari proses pembelajaran yang telah dilakukan sebelumnya, yaitu berdasarkan ciri yang telah didapatkan pada masing-masing sampel suara.

5. Ekstraksi ciri. Ekstraksi ciri suara dilakukan untuk mendapatkan karakteristik atau ciri dari sampel suara yang akan diproses yang dijadikan input dari sistem. Dari pemrosesan ini diharapkan akan didapatkan suatu ciri dari setiap jenis suara yang akan menjadi pembeda terhadap sampel suara lainnya.

- Input*, yaitu masukan suara yang berasal dari tiap pembicara.
- Sampling Rate (FS)*, yaitu banyaknya nilai yang diambil dalam satu detik. Dalam penelitian ini digunakan sampling rate sebesar 8000 Hz.
- Time Frame (TS)*, yaitu waktu yang diinginkan untuk satu *frame* (dalam milidetik). *Time frame* yang digunakan adalah 30 ms.
- Lap (M)*, yaitu *overlapping* yang terdapat dari $N/2$.
- Cepstrum coefficient*, yaitu jumlah *cepstrum* yang diinginkan sebagai *output*. *Cepstrum coefficient* yang digunakan sebanyak 14 didapat dari nilai *spectrum* pada gambar 4 yang mencakup nilai frekuensi dari data suara.



Gambar 4. Hasil rekaman suara di FFT

Hasil ekstraksi ciri menggunakan MFCC memiliki hasil berupa matriks ciri $n \times k$, n adalah jumlah frame dan k adalah koefisien. Untuk menghasilkan matriks yang berukuran sama di setiap suara, yaitu matriks berukuran $l \times k$, dilakukan perata-rataan koefisien pada setiap baris.

6. Database. Database digunakan untuk menyimpan ciri dari data pelatihan.

Pada KNN, terdapat 4 tahapan proses. Tahapan awal dimulai dengan memasukkan data fitur suara dan nilai k . Kemudian perhitungan jarak, pengurutan (*sorting*), dan dilanjutkan dengan proses *weighted voting* sesuai dengan nilai k yang dimasukkan, sehingga dihasilkan output

berupa hasil klasifikasi suara berdasarkan usianya.

Output dari proses ini adalah jenis suara yang telah diklasifikasikan, jenis suara yang akan terdeteksi adalah sesuai dengan jumlah k yang akan dipilih.

Perhitungan Tingkat Pengenalan. Tingkat pengenalan sistem akan dihitung dengan menggunakan ROC untuk mengevaluasi hasil penelitian. Untuk setiap data yang diuji, akan dilihat apakah data tersebut mengklasifikasi dengan benar atau tidak.

$$\text{Akurasi Sistem} = \frac{TP + TN}{TP + FP + FN + TN} \times 100\%$$

$$FPR = \frac{FP}{(FP + TN)} \times 100$$

$$FNR = \frac{FN}{(FN + TN)} \times 100$$

Ket:

True Positive (TP)

False Negative (FN)

False Positive (FP)

True Negative (TN)

FPR (*False Positive Real*) serta *FNR*

(*False Negative Real*).

HASIL DAN PEMBAHASAN

Pengolahan data suara. Pengujian klasifikasi suara berdasarkan usia menggunakan *Mel-Frequency Cepstrum Coefficients* MFCC dan *K-Nearest Neighbor* digunakan dataset suara hasil ekstraksi ciri terdiri dari 300 *record*. Data yang digunakan dalam sistem yaitu data dengan ciri masing-masing suara diwakili dengan 14 koefisien, dengan ukuran matriks dari semua suara ada 300×14 . Uji coba dilakukan dengan menggunakan prosentase jumlah data latih sebesar 50% dan data uji sebesar 50% dengan ukuran matriks masing-masing 150×14 , serta menggunakan 9 nilai k yang berbeda.

Jumlah dataset sebanyak 300 suara, maka akan digunakan data latih sebanyak 150 (25 orang perempuan dan laki-laki masing-masing kelas) dan data uji sebanyak 150 (25 orang perempuan dan laki-laki masing-masing kelas) data. Pengujian ini diberi nilai k ganjil dari 1, 3, 5, 7, 9, 11, 13, 15, 17. Setelah melakukan pengujian setiap penutur dengan nilai k yang berbeda, maka didapatkan data hasil pengujian tersebut. Dalam proses pengujian akan dikenali berdasarkan ID yang telah ditentukan, Setiap penutur diwakili dengan ID untuk ID 1

mewakili anak-anak, ID 2 mewakili remaja dan ID 3 mewakili tua.

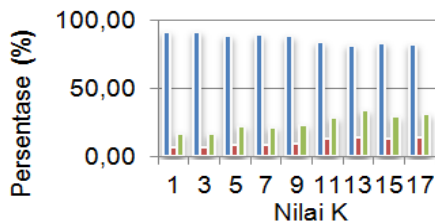
Membuktikan *performance* dari penelitian sistem klasifikasi ini, berikut skenario pengujian yang dilakukan:

1. Skenario 1 yaitu pengujian data kelas anak-anak terhadap semua data latih sebanyak 150 data pada kelas anak-anak, remaja, dan tua.
2. Skenario 2 yaitu pengujian data kelas remaja terhadap semua data latih sebanyak 150 data(anak-anak, remaja, dan tua).
3. Skenario 3 yaitu pengujian data kelas tua terhadap semua data latih sebanyak 150 data(anak-anak, remaja, dan tua).

Tabel 1 Tingkat Rata-rata Akurasi, FNR dan FPR

No	Nilai K	Rata-Rata		
		Akurasi (%)	FNR (%)	FPR (%)
1	1	91,11	6,47	16,67
2	3	91,11	6,49	16,00
3	5	88,22	8,43	21,67
4	7	88,67	8,37	21,00
5	9	87,78	9,03	22,33
6	11	83,56	12,59	28,00
7	13	81,33	13,86	33,33
8	15	82,89	12,93	28,67
9	17	82,00	13,57	30,67

Berdasarkan tabel 1 dapat dilihat bahwa rata-rata tingkat akurasi paling bagus ada pada k3 = 91,00%, begitu pula dengan nilai FNR dan FPR yang paling bagus ada pada k3 yaitu 6,49% dan 16,00%. Hal ini didasari pada kondisi ideal dimana akurasi paling baik adalah 100% sedangkan untuk nilai FNR 0% dan nilai FPR 0%.



Gambar 5. Pengaruh nilai k terhadap parameter uji (akurasi, FNR dan FPR) pada sistem

Berdasarkan pada grafik 1 nilai persentase cenderung fluktuatif. Nilai yang paling bagus ada pada k=3 untuk FNR, FPR dan akurasi. Hal ini mengacu pada kondisi ideal yaitu nilai FNR = 0%, FPR = 0%,

dan Akurasi = 100% . Berikut ini akan dijelaskan pengaruh parameter uji terhadap nilai k:

- Pengaruh *k* terhadap akurasi berdasarkan penelitian ini adalah semakin tinggi atau semakin besar nilai *k* maka tingkat akurasi cenderung menurun. Berikut nilainya pada tiap-tiap *k*. K1 = 91,11 %, K3 = 91,11 %, K5 = 88,22 %, K7 = 88,67 %, K9 = 87,78 %, K11 = 83,56 %, K13 = 81,33 %, K15 = 82,89 % dan K17 = 82,00%.
- Pengaruh *k* terhadap FPR berdasarkan penelitian nilai FPR cenderung fluktuatif, tapi semakin besar nilai *k* maka nilai FPR akan semakin buruk.
- Pengaruh *k* terhadap FNR yaitu nilai FNR cenderung naik dengan semakin bertambahnya nilai *k* yang diberikan.

KESIMPULAN

Berdasarkan hasil penelitian tentang klasifikasi suara berdasarkan *usia* dengan menggunakan metode *Mel Frequency Cepstral Coefficient* (MFCC) dan *K-Nearest Neighbor* (KNN), dapat disimpulkan bahwa:

1. Proses ekstraksi dengan MFCC (*Mel Frequency Cepstral Coefficient*) yang terdiri dari *frame blocking*, *Hamming Window*, *Fast Fourier Transform*, *Mel frequency wrapping*, dan *Cepstrum* digunakan 14 koefisien, dimana nilai tersebut mewakili suara setiap pengucap dari masing-masing kelas anak-anak, remaja dan tua.
2. Secara keseluruhan kinerja sistem dalam mengklasifikasikan suara berdasarkan usia dapat dikatakan berhasil, dengan rata-rata persentase akurasi tertinggi 91,11% , FNR sebesar 6,49 % dan FPR 16,00 %.
3. Parameter KNN untuk pengenalan sistem klasifikasi suara berdasarkan usia memiliki nilai *k* terbaik yaitu pada k=3 sebesar 91,11% dan nilai terjelek ada pada k=13 sebesar 81,33 %.

DAFTAR PUSTAKA

- Akbar, F., 2011, "*Konversi Nada-nada Akustik Menjadi Chord Menggunakan Pitch Class Profile*", Jurusan Teknik Informatika Institut Teknologi Sepuluh Nopember, Surabaya.

- Daryanto, T., 2008, *Sistem Multimedia Pertemuan ke-2 Audio dan Suara*, <http://kk.mercubuana.ac.id/files/92052-2-132860772042.pdf> .
- Prasetya, B. W., 2008, "*Identifikasi Suara Pria dan Wanita Berdasarkan Frekuensi Suara*", *Jurnal Informatika*. Universitas Kristen Duta Wicana, Yogyakarta.
- Satrya, R., 2010, "*Sistem Identifikasi Suara Pria Dan Suara Wanita Berdasarkan Usia Menggunakan Mel Frequency Cepstral Coefficient Dan K-Mean Clustering*", Fakultas Elektro dan Komunikasi Institut Teknologi Telkom Bandung, Bandung.
- Setiawan, A., 2011, *Klasifikasi Suara Berdasarkan Gender dengan Algoritma K-Means Prosiding Seminar Nasional Pengaplikasian Telematika 2011 (SINAPTIKA 2011)*, Universitas Mercu Buana, Jakarta.
- Subramanian, H., 2004, *Audio Signal Classification*. http://www.ee.iitb.ac.in/~esgroup/es_mtech04_sem/es_sem_04_paper_04307909.pdf.
- Yessivirna,R., 2011, *Klasifikasi Suara Berdasarkan Gender (Jenis Kelamin) Dengan Metode K-Nearest Neighbor (KNN)*, Jurusan Ilmu Komputer, Fakultas Program Teknologi Informasi dan Ilmu Komputer, Universitas Brawijaya, Malang.



Eva Susanti, lahir di Nyarinying, Sumbawa pada tanggal 09 Agustus 1992, Menempuh Pendidikan Program Strata 1 (S1) di Fakultas Teknik Universitas Mataram sejak tahun 2010.