JURNAL RISET REKAYASA ELEKTRO
Vol.4, No.2, Desember 2022, Hal. 39~50
P-ISSN: 2685 - 4341 E-ISSN: 2685 - 5313

Pengenalan Dialek Bahasa Daerah di Pulau Jawa menggunakan Metode *Mel-Frequency Cepstral Coefficients* dan *Adaptive Network-based Fuzzy Inference System*

Fajar Muhammad Fauzi[1], Latiful Hayat[2], Dian Nova Kusuma Hardani[3]
Program Studi S1 Teknik Elektro, Universitas Muhammadiyah Purwokerto
Fakultas Teknik dan Sains, Universitas Muhammadiyah Purwokerto

Informasi Makalah

Dikirim, 25 Agustus 2022
Direvisi, 20 Desember 2022
Diterima, 20 Desember 2022

Kata Kunci*:*

MFCC, ANFIS, Pengenalan Suara

INTISARI

Indonesia merupakan negara besar yang memiliki banyak keberagaman budaya dan suku sehingga memiliki banyak bahasa atau dialek yang berbeda-beda di setiap daerah. Berbagai penelitian dalam pengolahan sinyal suara telah banyak dikembangkan. Salah satu penelitian yang menarik untuk dikembangkan adalah Identifikasi Dialek. Dalam penelitian ini, dibuat suatu program atau aplikasi Speech Recognition dengan metode Mel-Frequency Ceptrums Coefficients (MFCC) dan Adaptive Neuro Fuzzy Inference System (ANFIS) untuk Bahasa di Pulau Jawa, yaitu Bahasa Betawi, Sunda, Banyumasan, dan Suroboyoan. MFCC untuk proses ekstraksi ciri dari sinyal ucapan pembicara dimana prosesnya akan mengubah sinyal suara menjadi beberapa vektor ciri yang kemudian akan ditampilkan dalam bentuk grafik. Analisis dan perancangan bentuk pola suara menggunakan software Phyton. Pengujian dilakukan dengan cara merancang perangkat lunak, melakukan pengambilan data, training sistem ANFIS, pembuatan GUI, dan pengujian dalam pengolahan data pengujian. Hasil dari penelitian ini menunjukkan MFCC mampu memberikan nilai ciri yang berbeda untuk setiap suara yang dimasukkan ke dalam sistem, dan parameter MFCC yang digunakan pada penelitian ini adalah Preemph = 0,99, Nfilt = 30, Nfft = 512, Winlen = 20ms, Winstep = 10ms, Numcep = 5, dan Lowfreq = 100. Model training yang digunakan dengan 120 data training dan 2 membership function tipe gausian menghasilkan nilai akurasi secara keseluruhan 32,5% pada proses pengujian (30% untuk Betawi, 50% untuk Sunda, 30% untuk Banyumasan, dan 20% untuk Suroboyoan).

Keyword*:*

MFCC, ANFIS, Speech Recognition

ABSTRACT

Because Indonesia is such a vast country with so much cultural and ethnic diversity, each region has its own set of languages or dialects. Much voice signal processing research has been widely established. Dialect Identification is an exciting topic that is currently being developed. Speech Recognition with Mel-Frequency Cepstrum Coefficients (MFCC) and Adaptive Neuro-Fuzzy Inference System (ANFIS) methods for Batavian, Sundanese, Banyumasan, and Suroboyoan dialects in Java island was developed in this study. MFCC denotes a feature extraction procedure from a speaker's speech signal, in which the voice signal is converted into numerous feature vectors and then shown graphically. Software Python was used to analyze and design sound pattern forms. The tests were carried out by designing software, collecting data, training the ANFIS system, creating a GUI, and testing data processing. The results of this study showed that MFCC could provide different feature values for each voice entered into the system, with Preemph = 0.99, Nfilt = 30, Nfft = 512, Winlen = 20ms, Winstep = 10ms, Numcep = 5, and Lowfreq = 100 as the MFCC parameters used in this study. The training model with 120 data training and two membership functions of the Gaussian type resulted in an overall accuracy value of 32.5% in the testing process (30% for Batavian, 50% for Sundanese, 30% for Banyumasan, and 20% for Suroboyoan).

Korespondensi Penulis:

Fajar Muhammad Fauzi
Program Studi Teknik Elektro
Fakultas Teknik dan Sains Universitas Muhammadiyah Purwokerto
JL. KH. Ahmad Dahlan, Purwokerto, 53182
Email: fajar.muslim.fauzi@gmail.com

1. PENDAHULUAN

Indonesia merupakan negara kepulauan terbesar di dunia. Menurut informasi Kementerian Dalam Negeri tahun 2018, jumlah pulau yang dimiliki oleh Negara Kesatuan Republik Indonesia (NKRI) yaitu 16.056 pulau [1]. Dengan banyaknya pulau menunjukkan bahwa Indonesia memiliki beragam suku, budaya, adat istiadat, dan bahasa. Bahasa dalam interaksi manusia merupakan sebuah alat komunikasi untuk bekerja sama dan mengidentifikasi diri sendiri dalam suatu masyarakat. Masyarakat terdiri atas budaya dan status sosial yang beragam, dimana keragaman dalam masyarakat tersebut berdampak pada variasi bahasa yang digunakan pada daerah tertentu atau yang biasa disebut dengan dialek.

Saat ini perkembangan zaman yang semakin maju menjadikan beberapa dialek bahasa daerah berada di ambang kepunahan. Jika kita mengabaikan hal tersebut, maka kita akan kehilangan sebagian kekayaan budaya bangsa yang tidak ternilai harganya [2]. Salah satu upaya untuk menjaga kelestarian bahasa adalah dengan melakukan penelitian yang terkait dengan dialek bahasa yang ada di Indonesia. Pulau Jawa adalah salah satu pulau di Indonesia yang penduduknya paling banyak, setidaknya ada 56% total penduduk Indonesia atau 149 juta penduduk yang tinggal di pulau Jawa [3]. Pulau Jawa terdiri dari beberapa provinsi antara lain DKI Jakarta, Jawa Tengah, Jawa Barat, Jawa Timur, Daerah Istimewa Yogyakarta, dan Banten. Pulau Jawa terdiri dari beberapa provinsi antara lain DKI Jakarta, Jawa Tengah, Jawa Barat, Jawa Timur, Daerah Istimewa Yogyakarta, dan Banten.

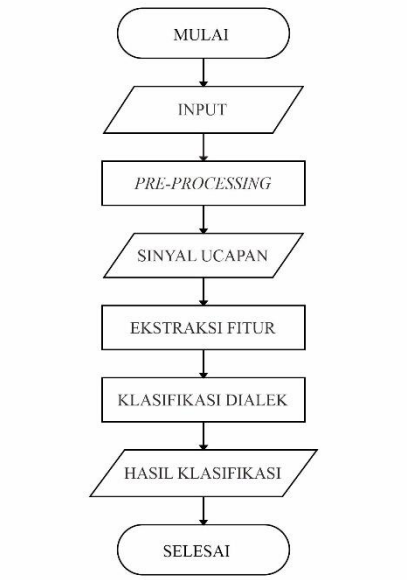
Seiring berkembangnya teknologi, saat ini banyak peneliti melakukan penelitian mengenai pengolahan sinyal suara. Dalam mengidentifikasi pola sinyal suara pada dialek bahasa daerah, muncul perbedaan-perbedaan yang menjadi ciri pada setiap sinyal suara yang dikeluarkan. *Speech recognition* merupakan salah satu dari teknologi yang memanfaatkan sinyal suara manusia sebagai input dan mampu dibaca oleh sistem. *Speech recognition* adalah kemampuan program untuk mengidentifikasi kata dan frasa dalam bahasa lisan dan mengkonversikannya ke format yang dapat dibaca oleh mesin [4]. Metode untuk ekstraksi suara yang sering digunakan adalah *Mel-Frequency Ceptrums Coefficients* (MFCC). MFCC banyak digunakan karena dianggap paling sesuai dalam memodelkan frekuensi suara manusia, sedangkan metode dalam mengidentifikasi sinyal suara salah satunya adalah dengan menggunakan metode *Adaptive Neuro Fuzzy Inference System* (ANFIS). Metode ini banyak digunakan untuk menganalisis suatu permasalahan dan juga digunakan sebagai metode pembelajaran untuk sistem.

Penelitian dari Putra (2021) berhasil mngidentifikasi dialek di Sumatra Selatan menggunakan ekstraksi mel spectrogram dan STFT. Akurasi tertinggi dicapai dalam mengenali dialek Beliti, yaitu 72,7% dan dialek Palembang 71,4 % jika ekstrasi ciri yang digunakan adalah mel spectrogram. Sedangkan untuk Bahasa Indonesia, akurasi tertinggi adalah dengan menggunakan ekstraksi ciri STFT, yaitu 71,4%. Berdasarkan latar belakang tersebut, maka dibuat sistem *speech recognition* dengan dialek yang ada di Pulau Jawa dengan menerapkan *Mel-Frequency Ceptrums Coefficients* (MFCC) untuk mengekstraksi ciri suara ke suatu sistem yang kemudian diidentifikasi melalui *Adaptive Neuro Fuzzy Inference System* (ANFIS) agar dimengerti oleh sistem tersebut dan cara menampilkan hasil keluaran dalam bentuk teks angka presentase serta menentukan dialek yang diucapkan.

2. METODE PENELITIAN

Pada penelitian ini diperlukan bahan suara sebagai masukan dengan 40 sampel untuk setiap dialek. Kemudian data diolah pada *pre-processing* untuk mendapatkan bentuk sinyal yang dapat di proses lebih lanjut oleh sistem. Hasil dari *pre-processing* kemudian dilakukan proses ekstraksi fitur untuk diambil cirinya mengunakan metode MFCC. Keluaran dari proses ekstraksi fitur kemudian dinormalisasi datanya agar data yang digunakan dapat dimasukkan kedalam *database* ciri. Proses klasifikasi dengan metode ANFIS, yang

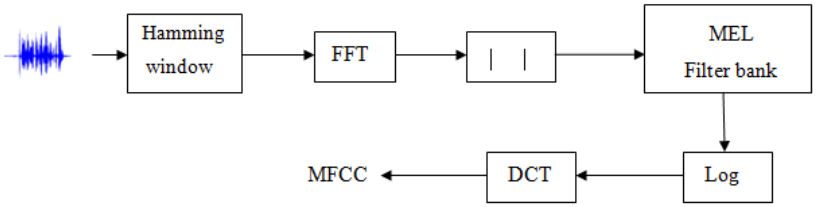
pertama dilakukan adalah melatih sistem agar mendapatkan hasil keluaran yang sesuai. Sistem yang sudah dilatih kemudian diuji dengan memberikan beberapa sample uji untuk meihat besarnya nilai akurasi sistem. Proses uji sistem dibangun dalam sistem berbasis aplikasi desktop/GUI dan hasil klasifikasi akan ditampilkan pada aplikasi tersebut guna mempermudah penggunaan. Gambar 1 adalah rangkaian alur identifikasi dari penelitian ini.



Gambar 1. Alur perancangan sistem

2.1. Ekstraksi Fitur

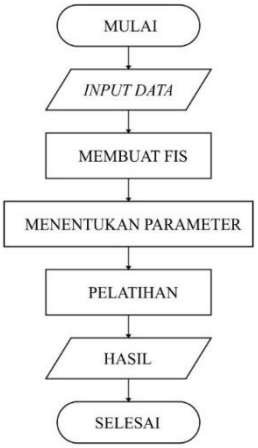
Ekstraksi ciri atau feature extraction dilakukan pada dua proses, yaitu ektraksi ciri untuk pembuatan database sebagai template dan ekstraksi ciri masukan data uji. *Mel-Frequency Cepstral Coefficients* (MFCC) adalah metode ekstraksi fitur yang mendekati sistem pendengaran manusia. MFCC sering digunakan untuk ekstrasi fitur dalam pengolahan ucapan yang bertujuan untuk mengekstrak karakteristik penting dari sinyal ucapan yang unik pada setiap kata untuk membedakan antara serangkaian kata yang berbeda. MFCC dianggap sebagai metode standar untuk ekstrasi fitur dalam pengenalan suara dan sering digunakan untuk pemrosesan suara [5]. Proses ekstraksi fitur penelitian ini dapat dilihat pada Gambar 2.



Gambar 2. Proses Ekstraksi Fitur dengan MFCC [7]

2.2 Klasifikasi Dialek

Pada tahap ini sistem klasifikasi dibuat dengan menggunakan metode ANFIS. *Adaptive Neuro-Fuzzy Inference System* (ANFIS) adalah penggabungan *Fuzzy Inference System* yang digambarkan dalam arsitektur jaringan syaraf. Proses awal klasifikasi dilakukan dengan cara menggabungkan semua nilai yang diperoleh pada tahapan ekstraksi ciri untuk dimasukan ke dalam suatu *database* yang kemudian digunakan dalam proses *training* ANFIS. Setelah ANFIS di-*training* oleh *database* maka akan diperoleh suatu sistem ANFIS yang dapat digunakan sebagai alat untuk mengidentifikasi dan mengklasifikasi dialek dengan dasar klasifikasi dari *database* tersebut. langkah-langkah yang dilakukan dalam pembuatan sistem ANFIS dapat dilihat pada Gambar 3.



Gambar 3. Diagram Alur Proses Klasifikasi

Pada tahap pertama data yang didapat dari hasil ekstraksi ciri semua komponen akan dimasukan ke database, kemudian dibangun sistem FIS. Langkah selanjutya adalah menentukan parameter-parameter ANFIS yang diperlukan agar sistem ANFIS dapat melakukan pengambilan keputusan sesuai dengan yang diinginkan. Kemudian dilakukan proses training sistem ANFIS untuk kemudian dilakukan pengujian terhadap sistem tersebut.

3. HASIL DAN PEMBAHASAN

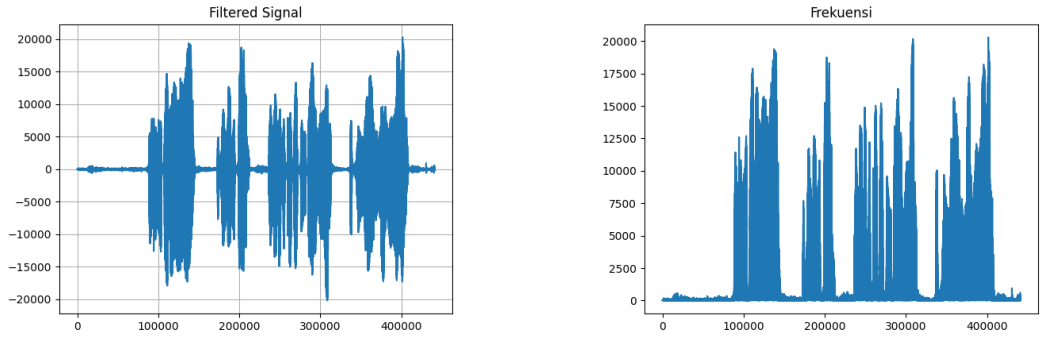
Pada penelitian ini, data yang digunakan berasal dari sampel suara dari dialek setiap daerah di pulau Jawa dalam bentuk format digital. Sampel suara didapat dari fitur rekaman suara *Voice Note* aplikasi WhatsApp dengan jumlah 160 sampel suara. Dengan 40 data untuk setiap dialek yang terdiri dari 30 data untuk data latih dan 10 data untuk data uji. Format yang digunakan aplikasi WhatsApp adalah .otf, .ogg dan .mp4 sehingga perlu diselaraskan menjadi format file .wav dan dipotong menjadi 10 detik. Setelah didapatkan data suara, selanjutnya adalah memproses data tersebut pada pre-processing melalui software PyCharm dengan bahasa pemrograman python.

3.1. Pre-processing

Setelah proses akuisisi data dan pengkondisian sinyal serta telah didapatkan data yang ideal, maka proses selanjutya adalah *pre-prosessing* atau pra-pengolahan yang berfungsi untuk mempersiapkan data sebelum dimasukan ke proses ekstraksi ciri. Pada penelitian ini proses *pre-prosessing* dibagi dalam beberapa tahapan yaitu :

1) *Filtering*

Pada proses *filtering* data yang telah diolah pada akuisisi data disaring (*filter*) pada rentang frekuensi 1 Hz – 16.384 Hz. Rentang maksimum frekuensi yang dipilih adalah sebesar 16.384 Hz, karena nilai tersebut adalah batas frekuensi yang mampu didengar oleh telinga manusia. *Output* dari *filtering* tersebut dapat dilihat pada Gambar 4.

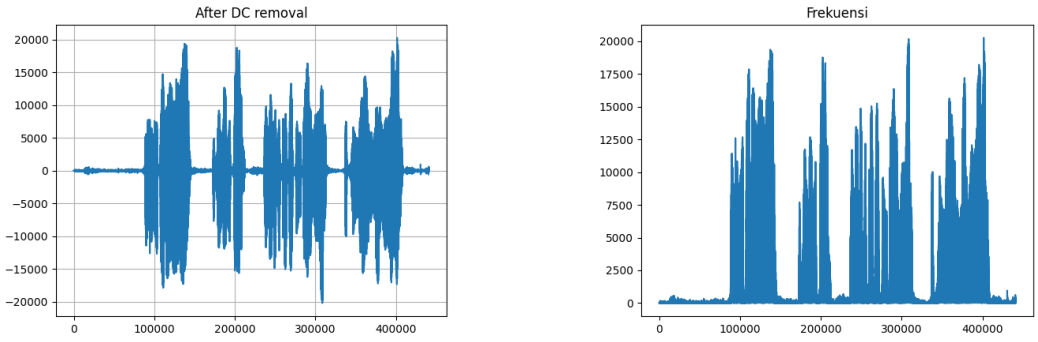


(a) (b)

Gambar 4 Output suara setelah proses filtering domain waktu (a) dan frekuensi (b)

2) *DC Removal*

Proses konversi sinyal analog ke digital menyebabkan *noise* dengan frekuensi yang sangat rendah dengan rentang 0 – 5 Hz. *Noise* yang terbentuk dari proses tersebut merupakan komponen DC. Untuk menghilangkan komponen DC diperlukan proses DC *Removal* yang dapat dilakukan dengan mengurangi sinyal dengan nilai rata-rata sinyal. Seperti prosedur *High Pass Filter*, pada DC *Removal* nilai frekuensi diatas nilai tertentu diloloskan sehingga *noise* pada rentang frekuensi rendah teredam. Pada *PyCharm*, sebagai pengganti DC *Removal* digunakanlah *High Pass Filter* dengan frekuensi *cut-off* 5 Hz. *Output* dari proses DC *Removal* dapat dilihat pada Gambar 5.

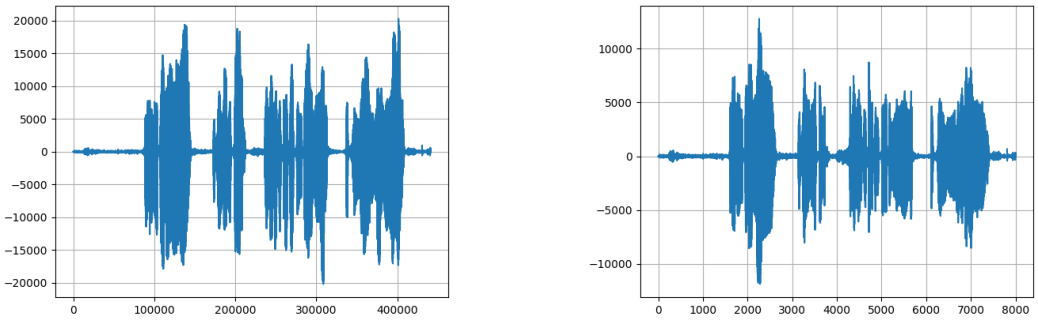


(a) (b)

Gambar 5. Output suara setelah proses DC removal pada domain waktu (a) dan frekuensi (b)

3) *Resampling*

Proses selanjutnya adalah *Resampling*, dimana data yang telah di *filter* memiliki nilai *sampling* sebesar 44.100 Hz diturunkan menjadi sebesar 8.000 Hz. Penurunan *sampling* ini berfungsi untuk mengurangi beban kerja sistem. Hasil dari proses *down-sampling* dapat dilihat pada Gambar 6.



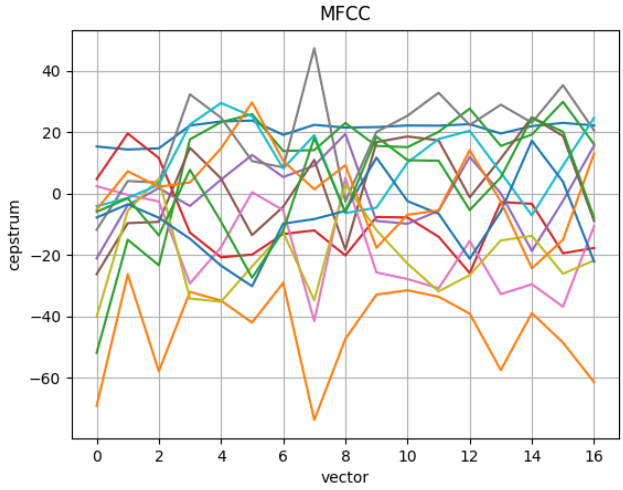
(a) (b)

Gambar 6. Bentuk sinyal sebelum (a) dan sesudah (b) proses resampling

Pada Gambar 6 data yang diperoleh dari proses *down-sampling* terlihat penurunan *sampling* dari panjang data semula 44,1 x 10^4 menjadi 8 x 10^4. Perubahan ini terjadi dikarenakan nilai dari *sampling rate* yang digunakan hanya 8.000 sampel maka panjang data menjadi 10 detik x 8.000 sampel = 80.000 sampel. Bentuk sinyal setelah diturunkan tidak berubah, yang berarti menandakan bahwa data dari suara telah diturunkan sampelnya.

3.2. Ekstraksi Ciri

Proses selanjutnya adalah data di ekstraksi cirinya untuk mendapatkan bagian unik dari setiap datanya. MFCC menjadi salah satu metode untuk mendapatkan ciri dengan nilai cepstrum. *Phyton* memiliki *library* untuk memproses sinyal suara menggunakan metode MFCC. Dalam modul MFCC terdapat dalam *library Phyton_speech_features* beberapa fungsi yang ada dalam library tersebut: M*el Frequency Cepstral Coefficients*, *Filterbank Energies*, *Log Filterbank Energies*, *Spectral Subband Centroids*. Hasil keluaran pada proses ekstraksi ciri dapat dilihat pada gambar 7.

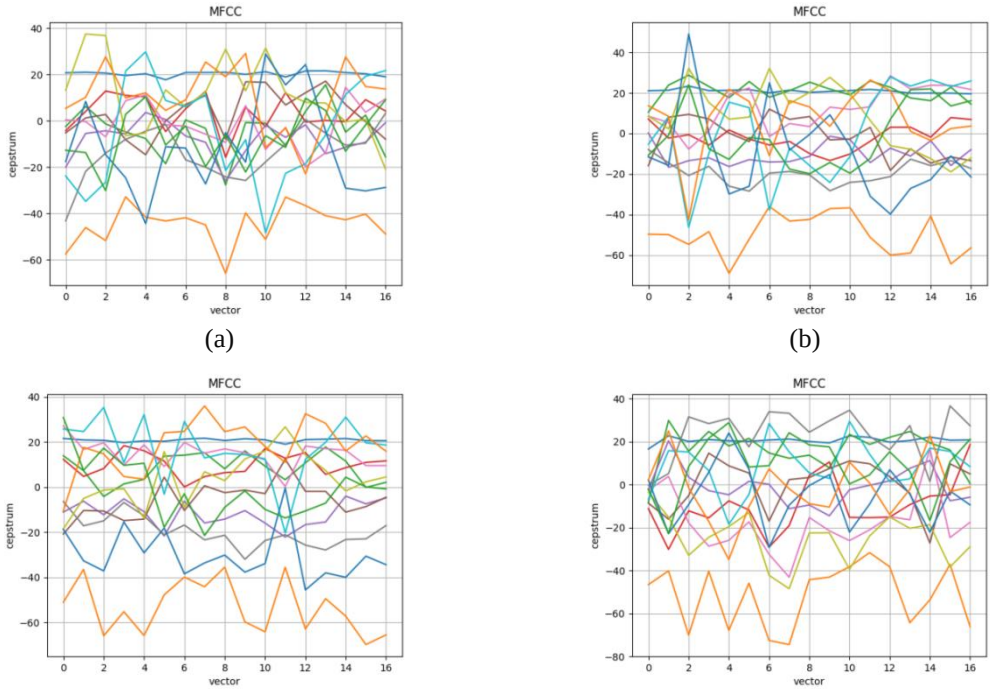


Gambar 7. Hasil dari proses ekstraksi ciri menggunakan MFCC

Pada penelitian ini, nilai MFCC akan dicari yang paling optimal dengan mengubah nilai pada parameter yang ada pada modul MFCC. Bentuk grafik dan nilai akan ditampilkan agar terlihat perbedaan pada setiap dialek. Parameter yang akan diubah adalah jumlah koefisien *cepstral*, batasan frekuensi rendah dan tinggi, ketetapan *preemphasis*, dan durasi frame.

1) Percobaan 1

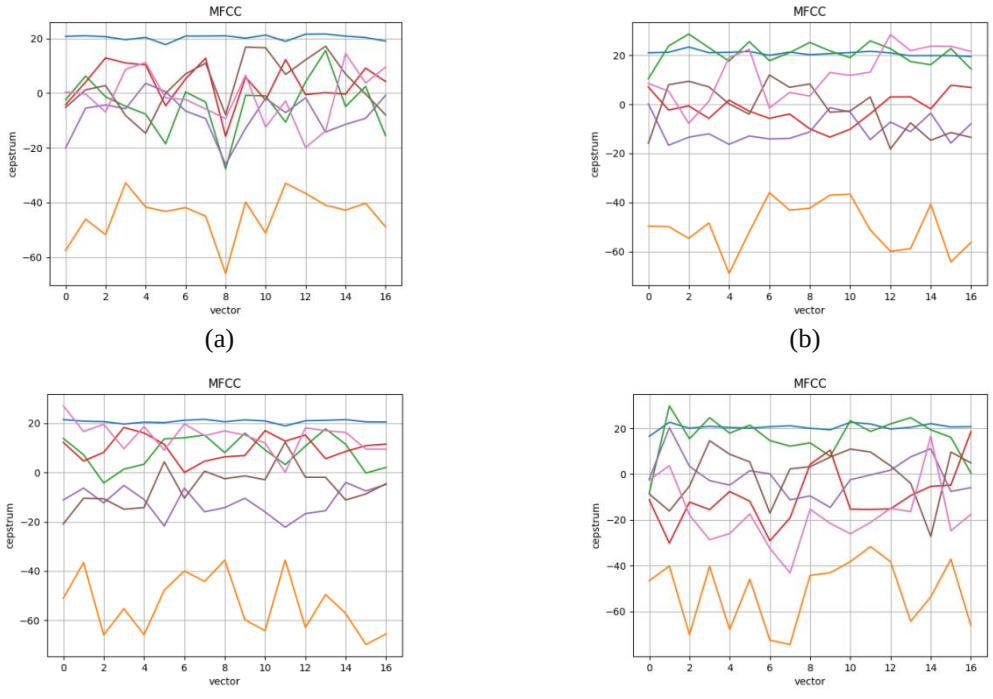
Pada percobaan 1 ini tidak ada nilai yang diubah, semua parameter yang digunakan adalah sesuai dengan default dari modul MFCC, hanya memasukan nilai sinyal dan *samplerate* yang datanya telah diolah sebelumnya. Maka hasil dari perhitungan dengan menggunakan percobaan 1 ini didapat grafik ciri seperti Gambar 8.



Gambar 8. Hasil ekstraksi percobaan 1 dialek (a) betawi, (b) ngapak, (c) sunda, (d) suroboyoan

2) Percobaan 2

Pada percobaan 2 ini hanya dikurangi untuk nilai koefisien *cepstral* menjadi 7, hal ini dilakukan untuk mengurangi beban kerja dari sistem ANFIS. Kemudian parameter lain akan diatur *default* sesuai dengan yang telah ditentukan oleh modul MFCC. Bentuk grafik dari percobaan 2 adalah seperti Gambar 9.



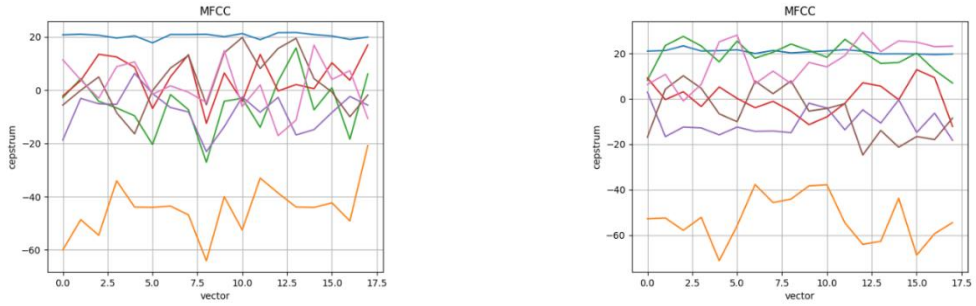
(a) (b)

(c) (d)

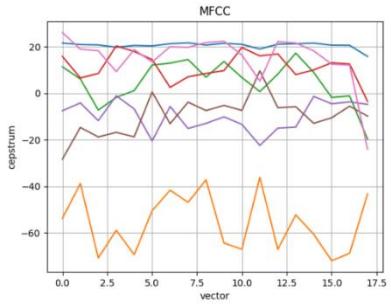
Gambar 9. Hasil ekstraksi percobaan 2 dialek (a) betawi, (b) ngapak, (c) sunda, (d) suroboyoan

3) Percobaan 3

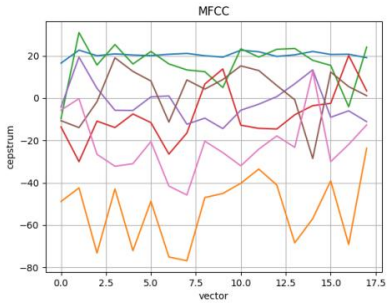
Pada percobaan 3 ini jumlah koefisien *cepstral* dikurangi menjadi 7, agar mengurangi beban kerja pada saat menjadi parameter nilai dalam ANFIS. Kemudian parameter yang lain akan disamakan sesuai penelitian yang dilakukan oleh Rusydi Umar, Imam Riadi, dan Abdullah Hanif yang berjudul "*Analisis Bentuk Pola Suara Menggunakan Ekstraksi Ciri Mel-Frequency Cepstral Coefficients (MFCC)"* dimana pada penelitian tersebut berhasil membedakan bentuk pola suara dari masing-masing sampel. Parameter yang digunakan adalah *Preemph* = 0,99, *Nfilt* = 30, *Winlen* = 20, dan *Lowfreq* = 100. Bentuk grafik dari percobaan 3 adalah seperti Gambar 10.



(a) (b)



(c)

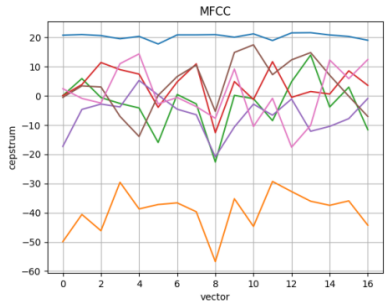


(d)

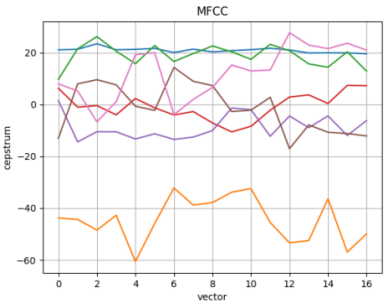
Gambar 10. Hasil ekstraksi percobaan 3 dialek (a) betawi, (b) ngapak, (c) sunda, (d) suroboyoan

4) Percobaan 4

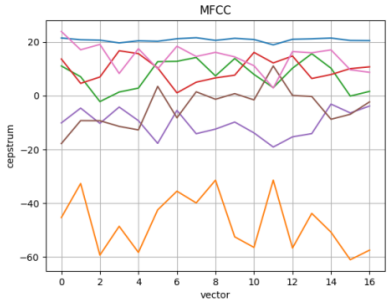
Pada percobaan 4 ini jumlah koefisien *cepstral* dikurangi menjadi 7, agar mengurangi beban kerja pada saat menjadi parameter nilai dalam ANFIS. Kemudian parameter yang lain akan disamakan sesuai penelitian yang dilakukan oleh Siti Helmiyah, Abdul Fadlil, dan Anton Yudhana yang berjudul "*Pengenalan Pola Emosi Manusia Berdasarkan Ucapan Menggunakan Ekstraksi Fitur Mel-Frequency Cepstral Coefficients (MFCC)*" dimana pada penelitian tersebut, MFCC mampu mengenali pola emosi manusia berdasarkan ucapan dari sampel yang ada. Parameter yang digunakan adalah *Preemph* = 0,97, *Nfilt* = 20, *Winlen* = 25, *Winstep* = 10, dan *Numcep* = 7. Bentuk grafik dari percobaan 4 adalah seperti Gambar 11.



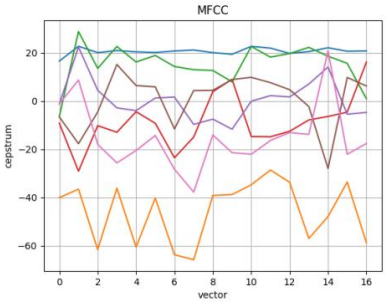
(a)



(b)



(c)



(d)

Gambar 11. Hasil ekstraksi percobaan 4 dialek (a) betawi, (b) ngapak, (c) sunda, (d) suroboyoan

3.3. Training ANFIS

Proses *training* diawali dengan pembuatan *database training* yang digunakan untuk melatih sistem ANFIS agar mampu melakukan proses klasifikasi. Data yang dimasukkan kedalam *database training* ini adalah ketujuh ciri yang dihasilkan dari proses ekstraksi ciri menggunakan metode *Mel-Frequency Cepstral Coefficients* (MFCC). Kemudian setelah didapatkan *database training*, maka selanjutnya adalah melakukan proses *training*. Proses *training* ini memerlukan dua modul yaitu ANFIS dan *membership function*, dimana kedua modul ini telah tersedia pada *library python*. Proses pemanggilan *database* pada *training* ini menggunakan perintah :

```
ts = numpy.loadtxt("database_MFCC.txt",
usecols=[1,2,3,4,5,6,7,8])
X = ts[:,0:7]
Y = ts[:,7]
```

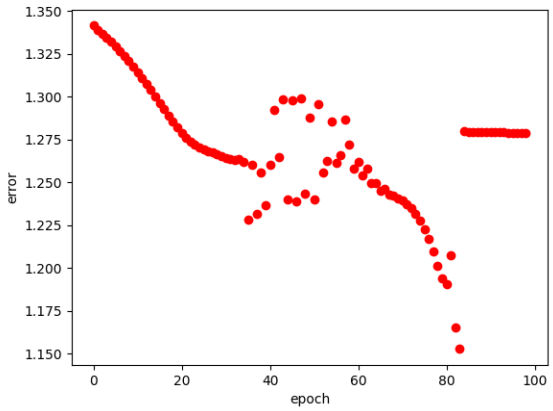
Dimana pada perintah tersebut, kolom yang akan digunakan pada *database* adalah dari kolom ke-1 sampai ke-8. Dan untuk *input* parameter *training* ditunjukkan pada variabel X, sedangkan variabel Y digunakan untuk menentukan *output* training ideal. Kemudian pada proses *training* ini menggunakan 2 *membership function* untuk tiap variabelnya. Dikarenakan keterbatasan komputasi menggunakan *python* dengan spesifikasi perangkat yang digunakan oleh penulis, maka ditentukan hanya menggunakan 2 *membership function*. Setelah ditentukan jenis *membership function*-nya selanjutnya adalah membuat FIS *(Fuzzy Inference System*) menggunakan modul ANFIS menggunakan perintah :

```
anf = ANFIS.ANFIS(X, Y, mfc)
anf.trainHybridJangOffLine(epochs=5)
```

Variabel anf menjadi parameter didalam program ANFIS ini. Pada modul ANFIS ini menggunakan metode pembelajaran *hybrid* yang merupakan metode pebelajaran yang terdiri dari dua bagian yaitu arah maju dan arah mundur. Dimana pada arah maju menggunakan metode LSE (*Least Square Estimator*) dan pada arah mundur menggunakan *backpropagation-error*.

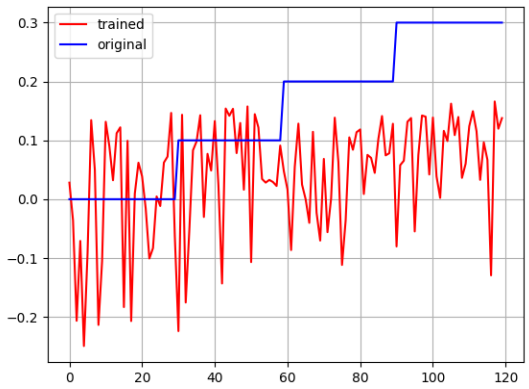
Proses *training* ini dilakukan dengan memasukan nilai *epoch* atau nilai iterasi. Nilai *epoch* adalah nilai iterasi atau pengualangan pada proses *training*, dimana pada setiap pengulangan akan dilakukan perubahan nilai bobot pada *neural network* untuk membuat sistem lebih mampu beradaptasi pada keluaran yang dikehendaki. Sejauh ini masih belum ditemukan metode yang mampu menentukan nilai *epoch* yang ideal pada suatu *neural networks*. Oleh karena itu penulis melakukan beberapa kali percobaan untuk menentukan nilai dari *epoch* yang ideal.

Pada percobaan pertama menggunakan 7 variabel input pada *database*, sistem memproses nilai *error* dengan sangat lama. Karena keterbatasan perangkat yang digunakan oleh penulis, maka menulis mencoba mengurangi parameter input menjadi 5. Dengan mengurangi jumlah variabel maka sistem bekerja dengan baik. Oleh karena itu nilai *epoch* yang akan dipilih menggunakan 5 variabel *input* saja, yaitu mulai dari kolom 0 hingga 5 pada variabel X. Pada percobaan ini penulis menentukan nilai *epoch* pada *training* diawali dengan 100. Kemudian akan dlihat titik mana nilai *error* yang paling kecil. Gambar plot nilai *error* dengan niai *epoch* 100 dapat diamati pada Gambar 12.



Gambar 12. Plot data hasil training dengan nilai epoch 100

Dari percobaan dengan nilai *epoch* 100 yang ditampilkan pada Gambar 3.6 terlihat nilai *error* yang paling kecil terjadi ketika nilai *epoch* berada pada nilai 83 dengan nilai *error* sebesar 1,152809. Niai yang cukup besar untuk sebuah nilai *error*, karena nilai *error* yang paling baik adalah yang mendekati 0. Dengan hasil yang muncul pada gambar, maka telah ditentukan nilai *epoch* yang akan digunakan pada saat proses *training* adalah sebesar 83. Plot data asli dan data *training* dengan nilai *epoch* 83 dapat diamati pada Gambar 13.



Gambar 13. Plot data asli dan data original dengan nilai epoch 83

Dari proses *training* ANFIS ini telah didapatkan sebuah sistem ANFIS yang akan digunakan sebagai alat klasifikasi dialek. Dimana pada keluaran dari sistem merupakan nilai yang bersifat *crisp*. Nilai *crisp* ini memiliki range antara 0 – 0,3 dan angka inilah yang akan digunakan sebagai penentu dari setiap dialek. Penentuan ini dilakukan dengan perintah :

```
anf_output = ANFIS.predict(anf, MFCC_mean)
if anf_output < 0:
  print("Others")
elif anf_output < 0.1;
  print("Betawi")
elif anf_output < 0.2;
  print("Sunda")
elif anf_output < 0.3;
  print("Banyumasan")
elif anf_output <= 0.4;
  print("Suroboyoan")
else
  print("Others"
```

maka bentuk kaidah dari perintah diatas adalah sebagai berikut:

1) ∞ - 0 output adalah other
2) 0 < 0,1 output adalah Betawi
3) 0,1 < 0,2 output adalah Sunda
4) 0,2 < 0,3 output adalah Banyumasan
5) 0,3 < 0,4 output adalah Soroboyoan
6) 0,4 - ∞ ouput adalah other

3.4. Pengujian Sistem ANFIS

Proses pengujian sistem ANFIS dilakukan dengan cara menjalankan program atau GUI pada *Python*. Selanjutnya dilakukan pengujian tingkat akurasi dari sistem. Pengujian dilakukan untuk mengetahui dari tingkat ketepatan sistem dalam mengklasifikasi dialek Betawi, Sunda, Banyumasan, dan Suroboyoan. Data *input* dari pengujian ini berupa 40 data suara yang terdiri dari 10 suara dari setiap dialek. Hasi dari nilai akurasi didapat dengan cara membandingkan jumlah nilai klasifikasi yang benar dengan jumlah seluruh data yang diuji. Nilai klasifikasi bernilai benar apabila hasil klasifikasi dari program sesuai dengan dialek yang sebenarnya, sedangkan nilai bernilai salah apabila hasil klasifikasi dari program tidak sama dengan dialek yang sebenarnya atau ketika program tidak dapat memasukan jenis dialek yang telah diklasifikasi ke salah satu dari empat dialek yang digunakan pada penelitian ini.

Setiap hasil data uji dialek menghasilkan nilai akurasi yang kurang dari 50%, dan total keakuratan kurang dari 50% hal ini bisa terjadi karena beberapa faktor yaitu :

1) Sedikitnya jumlah sampel yang digunakan

Pada penelitian ini 30 data digunakan untuk *training* dan 10 data digunakan untuk data uji. Semakin banyak sampel yang digunakan pada saat proses *training* maka akan semakin baik pula sistem dalam mengklasifikasikan suatu dialek.

2) Rendahnya jumlah variabel dari ekstraksi ciri

Default jumlah keluaran dari MFCC adalah 13 ciri yang dimana pada jumlah tersebut diatur pada parameter *numcep*. Penulis mencoba mengurangi jumlah keluaran MFCC menjadi 7 ciri dan mencobanya menjadi *database training*, tetapi muncul *error* pada sistem yang memberitahukan bahwasannya *memory* perangkat tidak mampu bekerja dengan jumlah variabel yang banyak. Oleh karena itu penulis mengurangi hanya 5 ciri saja yang digunakan.

3) Program ekstraksi ciri yang masi terdapat kekurangan

Penentuan nilai parameter pada modul MFCC dilakukan berdasarkan 4 percobaan dengan masing-masing percobaan berdasar pada peneliti sebelumnya yang telah mampu menggunakan MFCC sebagai model ekstraksi ciri. Tetapi pada saat parameter digunakan masih banyak nilai yang belum mencirikan dari dialek tersebut. Pengujian untuk penentuan parameter sebaiknya dilakukan dengan banyak percobaan dengan membandingkan perubahan dari masing-masing parameter.

4) Rendahnya jumlah *membership function* yang digunakan

Pada proses training ANFIS, penelitian ini hanya menggunakan 2 *membership function*. Hal ini dilakukan karena keterbatasan dari perangkat. Penulis telah mencoba dengan menambah jumlah *membership function* menjadi 4, tetapi muncul *error* pada saat sistem dijalankan. *Error* tersebut terjadi karena *memory* perangkat tidak bisa menampung tugas yang diberikan. 2 *membership function* terlalu sedikit untuk digunakan sebagai dasar klasifikasi.

5) Pengaturan nilai *membership function*

Pada skrip program, terdapat penentuan jenis *membership function* yang digunakan, pada penelitian ini menggunakan *membership function* tipe Gaussian. *Membership function* tipe Gaussian ini meminta 2 parameter untuk digunakan yaitu *mean* dan sigma. Nilai tersebut mempengaruhi *membership function* dalam mengelompokan data.

4. KESIMPULAN

Berdasarkan analisis yang telah dilakukan, maka diperoleh kesimpulan hasil penelitian sebagai berikut :

1. Metode *Mel Frequency Cepstrums Coeffiecients* (MFCC) dapat memberikan nilai ciri yang berbeda pada setiap suara.
2. Nilai parameter yang digunakan pada metode MFCC mempengaruhi keluaran dari ekstraksi ciri MFCC.
3. Patameter MFCC yang digunakan pada penelitian ini adalah *Preemph* = 0,99, *Nfilt* = 30, *Nfft* = 512, *Winlen* = 20ms, *Winstep* = 10ms, *Numcep* = 5, dan *Lowfreq* = 100 mampu memberikan ciri yang berbeda pada setiap dialek.
4. Hasil ciri yang digunakan sebagai *database* adalah nilai rata-rata (*mean*) dari hasil matriks MFCC.
5. Jumlah sampel yang digunakan untuk *training*, jumlah variabel, serta jumlah *membership function* yang digunakan dalam pembuatan ANFIS berbanding lurus dengan nilai akurasi dari klasifikasi.
6. Sistem ANFIS yang dibuat pada penelitian ini menghasilkan nilai akurasi rata-rata sebesar 32,5% (30% untuk Betawi, 50% untuk Sunda, 30% untuk Banyumasan, dan 20% untuk Suroboyoan).
7. Pada sistem ANFIS belum tercapai target yang diharapkan yaitu nilai akurasi rata-rata diatas 80%. Sehingga sistem ANFIS yang dibuat pada penelitian ini dikatakan belum ideal.

UCAPAN TERIMAKASIH

Peneliti secara khusus mengucapkan terima kasih yang sebesar-besarnya kepada semua pihak yang telah membantu. Peneliti banyak menerima bimbingan, petunjuk dan bantuan serta dorongan dari berbagai pihak baik yang bersifat moral maupun material

DAFTAR PUSTAKA

[1] BPS, 2017. *Jumlah Pulau di Indonesia Menurut Provinsi*, Jakarta: Badan Pusat Statistika.
[2] Dharma, A., 2011. *Pembinaan dan Pengembangan Bahasa Daerah*. Semarang, Pusat Program Bahasa Universitas Diponegoro.
[3] Ashari, A., 2020. BOBO.id. [Online] Available at: https://bobo.grid.id/read/082008972/pulau-jawa-pulau-paling-banyak-penduduknya-di-dunia-pulau-apalagi-yang-banyak-penduduknya?page=all [Accessed 30 April 2020].
[4] Adi, M. R., Osmond, A. B. & Prasasti, A. L., 2019. Penentuan Dialek Jawa Menggunakan Metode Deep Neural Network. *e-Proceeding of Engineering*, Volume 6, pp. 5637-5647.
[5] Helmiyah, S., Fadil, A. & Yudhana, A., 2018. Pengenalan Pola Emosi Manusia Berdasarkan Ucapan Menggunakan Ekstraksi Fitur Mel-Frequency Cepstral Coefficients (MFCC). *Cogito Smart*, 4(2), pp. 372-381.
[6] Putra, R. M, 2021. Pengenalan Dialek Di Sumatera Selatan Menggunakan Algoritma Deep Neural Network. Prosiding Seminar Nasional Penelitian Dan Pengabdian Masyarakat Avoer 13. Universitas Sriwijaya
[7] Brunet, Kevin et. all.. 2013. Speaker Recognition for Mobile User Authentication: An Android Solution. 8ème Conférence sur la Sécurité des Architectures Réseaux et Systèmes d'Information (SAR SSI).