

LAPORAN PENGENALAN POLA

TERHADAP SUARA



Disusun oleh :

1. Ari Hilda Mawaddah (A11.2017.10396)
2. Laksita Maulisa Liztio (A11.2017.10405)
3. Vannya Maheswari (A11.2017.10407)

FAKULTAS ILMU KOMPUTER

UNIVERSITAS DIAN NUSWANTORO

2019

1. Latar belakang

Pengenalan suara merupakan suatu bidang yang masih terus diteliti dan dikembangkan untuk berbagai keperluan seperti suara pembeda suara perempuan dan laki - laki, emosi dan lain - lain. Suara yaitu sebuah bentuk untuk mengenali sesuatu yang bisa dikenali oleh manusia. Seiring dengan perkembangan pemrosesan tersebut muncullah berbagai jenis pemrosesan suara yang dihasilkan, yang bisa digunakan sebagai pembeda antara suara (klasifikasi suara) hewan, manusia, musik, atau yang lain. Klasifikasi sendiri adalah salah satu proses pengolahan sinyal digital yang didalamnya terdapat speech recognition. Suatu sistem yang menggunakan suara dapat dijadikan jalan alternatif dalam mempermudah aktivitas atau pekerjaan manusia terutama untuk para pengguna yang memiliki keterbatasan fisik. Namun, pemrosesan menggunakan suara tidaklah mudah dilakukan oleh sebuah mesin. Diperlukan suatu pembelajaran kepada mesin untuk mendapatkan hasil ekstraksi dan mendapatkan sebuah ciri suara. Metode kecerdasan buatan khususnya jaringan syaraf tiruan (JST), backpropagation, dan algoritma HMM adalah metode yang banyak digunakan dalam klasifikasi suara. Di dalam review ini, menjelaskan bagaimana dalam perancangan suatu program dari sebuah sample suara yang direpresentasikan secara numerik. Nantinya akan diambil contoh sample percobaan, setelah itu hasilnya akan diekstrak untuk mendapatkan cirinya. Hasil ciri tersebut yang digunakan sebagai klasifikasi suara apakah itu. Dalam percobaan ini hasil ekstraksi digunakan sebagai jaringan syaraf buatan. Tujuan dari review ini untuk merancang sebuah program yang dapat mengklasifikasi file record dengan menggunakan backpropagation.

2. Penelitian yang Terkait

Dalam laporan ini, menggunakan beberapa jurnal penelitian sebagai acuan untuk menyelesaikannya, penelitian tersebut dapat mempermudah kita dalam melakukan hasil penelitian yang baik dalam segi apapun baik konsep maupun teori yang digunakan. Hasil penelitian lain yang relevan dalam pendekatan permasalahan penelitian : teori, metode, masalah, solusi, hasil, kelemahan dan keunggulannya.

Berikut rangkuman dari beberapa penelitian yang digunakan dalam laporan ini:

No	Nama Peneliti dan Tahun	Masalah	Metode	Solusi	Hasil
1.	Faradiba Ruslan, 2014 -> yang dibahas 1	Pengenalan pola sinyal suara manusia menggunakan metode backpropagation neural network	Backpropagation Neural Network	<ul style="list-style-type: none"> -Mengumpulkan pada 5 orang berbeda, perempuan , laki – laki dan anak-anak untuk mengucapkan kata “HALO”. - Melakukan filtering suara untuk menghilangkan noise. - Melakukan sampling untuk mendapatkan sinyal diskrit agar dapat mengetahui ciri akurat dari sebuah sinyal - Melakukan desimasi untuk mengurangi sample yang terlalu banyak namun tetap menjadi perwakilan dari 	<ul style="list-style-type: none"> - Sistem dapat menyimpan suara yang sudah pernah masuk kemudian sistem dapat mengenali suara yang sudah pernah terlatih atau tersimpan sebelumnya tersebut. - Sistem dibuat untuk mendeteksi suara orang yang sama meskipun dalam keadaan random. - Data terlatih tersebut dapat diuji dengan hasil ujinya mencapai 74%

				<p>sample yang dihilangkan.</p> <ul style="list-style-type: none"> - Melalui proses Linear Predictive Coding yaitu Preemphasis, reshapes, windowing, dan analisis autokorelasi - Fast Fourier Transform digunakan untuk menunjukkan lebih jelas dari masing - masing pola sinyal. - Melakukan pelatihan sistem. - Melakukan pengujian sistem. 	
2.	Mada Sanjaya W.S, Zabidin Salleh, 2014	<ul style="list-style-type: none"> - Pengujian sistem ekstraksi ciri sinyal menggunakan Mel-Frekuensi Cepstrum Coefficients (MFCC) serta klasifikasi pola 	<ul style="list-style-type: none"> - Mel-Frequency Cepstrum Coefficients - Adaptive Neuro-Fuzzy Inference System - Microcontroller 	<ul style="list-style-type: none"> - Menggunakan metode MFCC, Preemphasis, frame blocking, windowing, Fast Fourier Transform, Mel-Frequency Wrapping, 	<ul style="list-style-type: none"> - Sistem mampu mendeteksi kata “nyala” dan “mati” dengan target “nyala” dengan “1” dan “mati” dengan “0”. - Sistem mampu

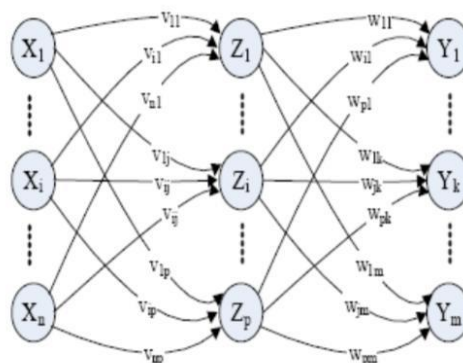
		suara menggunakan metode Adaptive Neuro-Fuzzy Inference System (ANFIS) - Implementasi sistem untuk otomatisasi kontrol lampu AC menggunakan mikrocontroller Arduino.	Arduino	cepstrum, dan adaptive neuro fuzzy inference system dalam pembuatan sistem tersebut. - Melakukan perancangan elektronik - Melakukan perancangan antarmuka menggunakan GUI Matlab.	menyimpan perintah dan memprosesnya ke microcontroller untuk menyalakan atau mematikan lampu.
3.	Barlian Henryranu P., Wijaya Kurniawan, Mochammad Hannats H. I., 2017 -> yang dibahas 2	Menganalisis tingkat kemarahan seseorang melalui suara dengan algoritma HMM.	Algoritma HMM	Memasukan sample suara yang kemudian dilakukannya ekstraksi ciri berdasarkan pitch, energi dan frekuensi. Dan dari hasil diklarifikasikan menggunakan HMM.	Sistem dapat mengenali emosi marah, bahagia dan netra, dan dengan rata - rata sebesar 86,66%. Rata waktu sebesar 21,6 ms.
4.	Ersa Triansyah, Youllia Indrawaty, 2017 -> yang dibahas 3	Menguji kecocokkan ucapan huruf hijaiyyah antara guru dengan muridnya.	Pattern recognition	- Menggunakan Pattern Recognition untuk pengenalan pola. - Tahap	- Sistem mampu melakukan pengenalan ucapan huruf hijaiyyah - Sistem mampu

				<p>pembandingan sinyal menggunakan algoritma Dynamic Time Warping (DTW)</p> <p>- Mendapat normalisasi sinyal suara berdasarkan waktu maupun pola menggunakan Mel-Frequency Cepstral Coefficient(MFCC).</p>	<p>mencocokkan huruf dan tanda baca dengan jarak yang ditentukan (1.3) kepada 3 pria dan 3 wanita dengan kecocokan hingga 94%.</p>
5.	Ina Agustina, Fauziah, Aris Gunaryati, 2016	Pengenalan identitas berdasarkan biometrik pola suara.	Biometrik jaringan saraf tiruan	<p>Sampling suara akan diproses dengan Frame Blocking, suara akan dibagi menjadi beberapa frame. Kemudian dilakukannya proses Windowing lalu Wavelet SCB MFCC.</p>	<p>Keakuratan sistem pengenalan data training set mencapai 87,5 % dan untuk blind data set mencapai 90 %.</p> <p>Adanya kesalahan pengenalan yang terjadi diakibatkan adanya perbedaan yang</p>

					terlalu besar antara sinyal suara yang hendak dikenali dengan suara yang dilatihkan.
--	--	--	--	--	--

1. Metode yang Digunakan

Pada laporan ini kami akan menggunakan metode Jaringan Syaraf Tiruan (JST) Backpropagation. JST Backpropagation sering digunakan sebagai pengenalan pola, mengklasifikasi dalam pengambilan keputusan agar dapat menghasilkan hasil yang memuaskan. Metode ini merupakan Jaringan Syaraf Tiruan (JST) adalah data yang terdiri dari input dan target (output). JST Backpropagation memiliki dua tahap yaitu *forward pass* (perambatan maju) dan *backward pass* (perambatan mundur). *Forward pass* adalah mencari nilai hidden layer dan output sedangkan *backward pass* adalah membandingkan output dengan target. perambatan mundur. Selama *forward pass*, tiap Input (X_i) menerima sebuah tanggung jawab yang bisa disebut dengan hidden layer (Z_i) untuk menghasilkan sebuah output (Y_i).



Gambar 2 Arsitektur Jaringan Lapis Jamak dengan Satu Lapis Unit Tersembunyi (Laurene F, 1994)

Proses identifikasi suara dengan JST Backpropagation melalui 3 tahap yaitu *forward pass*, *backward pass*, dan perubahan bobot. Output dari sistem adalah suara yang terdeteksi baik dikenali sebagai target maupun tidak dikenali. Proses pelatihan dilakukan sampai jaringan memperoleh nilai *error* target sekecil mungkin. Setelah syarat tersebut tercapai, maka dilakukan penyimpanan bobot dari setiap ruas pada jaringan. Bobot - bobot tersebut nantinya akan digunakan pada proses pengujian.

Tabel 1. Hasil Pelatihan Jaringan untuk 5 data latih dengan Fungsi Aktivasi Sigmoid Bipolar

Parameter		Percobaan	Pelatihan		Mean Square Error (MSE)
Hidden Layer	Neuron		Iterasi	waktu (detik)	
2	15,10	1	137	45,844	$9,127 \times 10^{-6}$
		2	844	405,58	$9,453 \times 10^{-6}$
		3	112	37,27	$1,785 \times 10^{-6}$

Dapat dilihat pada gambar diatas hasil review jurnal dengan metode backpropagation bahwa MSE yang dihasilkan rata – ratanya sebesar $1,785 \times 10^{-6}$ waktu yang dibutuhkan 37,27 dan dengan iterasi sebesar 112

Tabel 2. Pengenalan pola suara latih Latih untuk 5 data latih dengan Fungsi Aktivasi Sigmoid Bipolar

No	Input Latih	Dikenali sebagai				
		Darman	Imel	Farah	Yusuf	Januar
1	Darman1	√				
	Darman2	√				
	Darman3	√				
	Darman4	√				
	Darman5	√				
2	Imel1		√			
	Imel2		√			
	Imel3		√			
	Imel4		√			
	Imel5		√			
3	Farah1			√		
	Farah2			√		
	Farah3			√		
	Farah4			√		
	Farah5			√		
4	Yusuf1				√	
	Yusuf2				√	
	Yusuf3				√	
	Yusuf4				√	
	Yusuf5				√	
5	Januar1					√
	Januar2					√
	Januar3					√
	Januar4					√
	Januar5					√

MSE = $1,785 \times 10^{-6}$, Iterasi = 112, Waktu = 37,266 detik, Jumlah sample = 25, sukses = 25, gagal = 0, tingkat keberhasilan = 100%

HASIL DAN PEMBAHASAN

- **Implementasi Hasil Jaringan Syaraf Tiruan Melalui Proses Pelatihan**

Contoh implementasi yang dihasilkan menggunakan metode backpropagation untuk memprediksi sebuah suara atau pengenalan suara. Disini kami memprediksi lebih dari 5 data suara untuk hasil pengujiannya atau trainingnya terdiri dari suara perempuan dan laki - laki, sinyal suara formatnya dalam bentuk wave. Pada proses selanjutnya implementasi sistem dengan memanfaatkan toolbox dalam pemrograman Matlab. Pengenalan pola suara dengan *backpropagation* adalah suatu proses agar jaringan dapat mengenali suara masukan dengan benar. Untuk menghasilkan sebuah hasil dalam mendeteksi suara maka harus dilakukan pelatihan data agar dapat dikenali secara baik dan benar. Sebelum mengenali pola suatu masukan tentunya harus dilakukan pelatihan dengan nilai parameter yang benar sehingga jaringan dapat mengenali suara dengan baik.

- **Fungsi Aktivasi Sigmoid Biner**

Fungsi sigmoid biner sebenarnya tidak pernah mencapai angka 0 maupun 1. Oleh sebab itu, data suara perlu dinormalisasi terlebih dahulu salah satu contohnya ke dalam range 0,1 sampai dengan 0,9 menggunakan persamaan berikut ini:

$$X' = \frac{0,8 (X - b)}{(a - b)} + 0,1$$

dimana:

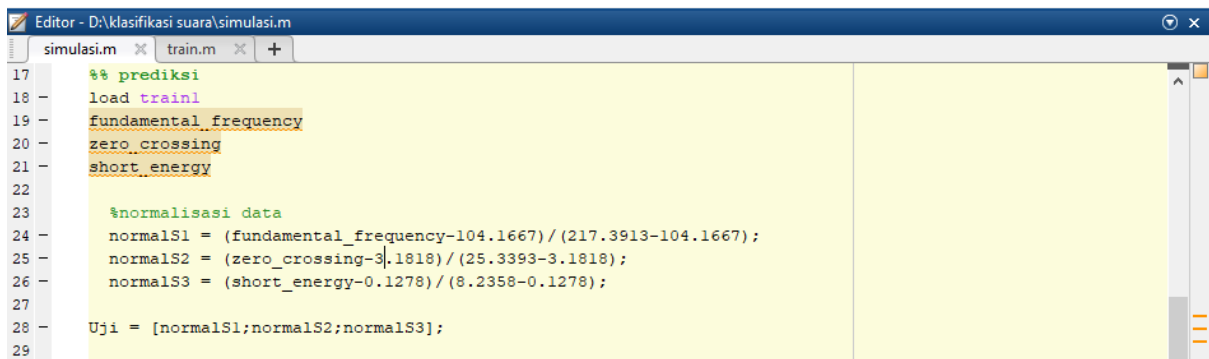
X' = data hasil normalisasi

X = data awal

a = nilai maksimum data awal

b = nilai minimum data awal

berikut source code pada tahap normalisasi



```
Editor - D:\klasifikasi suara\simulasi.m
simulasi.m  train.m  +
17  %% prediksi
18  load train1
19  fundamental_frequency
20  zero_crossing
21  short_energy
22
23  $normalisasi data
24  normalS1 = (fundamental_frequency-104.1667) / (217.3913-104.1667);
25  normalS2 = (zero_crossing-3.1818) / (25.3393-3.1818);
26  normalS3 = (short_energy-0.1278) / (8.2358-0.1278);
27
28  Uji = [normalS1;normalS2;normalS3];
29
```

	X1	X2	X3	TARGET	
1					
2	112.6126	18.4706	3.2174	0	
3	174.8252	18.6698	5.8299	0	
4	194.0994	21.6923	3.5125	0	
5	147.0588	14.0259	3.5271	0	
6	163.3987	23.875	8.2358	0	
7	185.4599	21.1466	5.6266	0	
8	217.3913	25.3393	6.8601	0	
9	186.0119	18.8421	4.0542	0	
10	104.1667	5	0.1278	1	
11	104.1667	3.1818	0.2552	1	
12	104.1667	6.037	0.1643	1	
13	104.1667	3.1818	0.1278	data minimum	
14	217.3913	25.3393	8.2358	data maksimum	

data asli sebelum dinormalisasi

	A	B	C	D	E
1	X1	X2	X3	TARGET	
2	0.0746	0.6900	0.3811	0	
3	0.6241	0.6990	0.7033	0	
4	0.7943	0.8354	0.4175	0	
5	0.3788	0.4894	0.4193	0	
6	0.5231	0.9339	1.0000	0	
7	0.7180	0.8108	0.6782	0	
8	1.0000	1.0000	0.8303	0	
9	0.7229	0.7068	0.4843	0	
10	0.0001	0.0821	0.0001	1	
11	0.0001	0.0001	0.0157	1	
12	0.0001	0.1289	0.0045	1	

data setelah dinormalisasi

- **Proses Pengujian Jaringan**

Data uji disiapkan dalam format xlsx (excel), dilakukan pemrograman untuk melakukan pelatihan jaringan. source code pelatihan jaringan untuk prediksi seperti berikut:

```

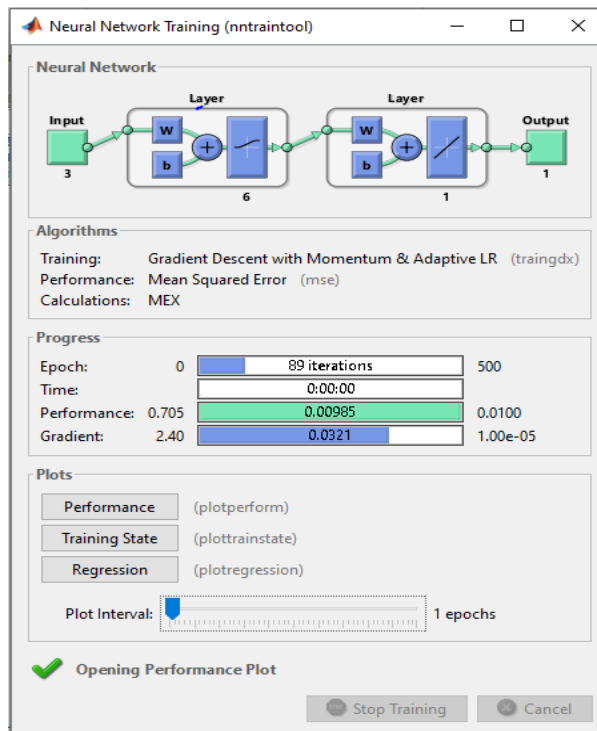
Editor - D:\klasifikasi suara\train.m
simulasi.m train.m +
2 - xlsdata=xlsread('data.xlsx',1,'a2:d12');
3 - urut = 1:11;
4 - data = xlsdata(urut(1:end),:);
5 - P=data(:,1:3)';
6 - T=data(:,4)';
7
8 - a1=500; %epoch
9 - a2=0.01; %learning rate
10 - a3=0.001; %goal
11 - a4=6; %neuron hidden layer
12
13 %% membangun jaringan saraf tiruan
14 - net = newff(minmax(P),[a4 1],{'logsig','purelin'},'traingdx');
15
16 %% paramater jaringan
17 - net.trainParam.epochs=a1;
18 - net.trainParam.goal=a3;
19 - net.trainParam.lr=a2;
20 - net.trainParam.show=1;
21 - net.trainParam.mc = 0.5;
22 - net.performFcn='msereg';
23 - net.performParam.ratio=0.5;
24
25 % pelatihan jaringan
26 - [train1,tr]=train(net,P,T);
27 - save train1 train1;

Editor - D:\klasifikasi suara\train.m
simulasi.m train.m +
25 % pelatihan jaringan
26 - [train1,tr]=train(net,P,T);
27 - save train1 train1;
28 %=====
29 %% simulasi
30 %read data
31 - load train1;
32 - xlsdata2=xlsread('data.xlsx',1,'a2:d12');
33 - urut = 1:11;
34 - datatesting = xlsdata2(urut(1:end),:);
35 - P2=datatesting(:,1:3)';
36 - T2=datatesting(:,4)';
37 - [m,n] = size(P2);
38 - an= round(sim(train1,P2));
39 - cek=[an' T2']
40
41 %hitung error
42 - compareA = T2 == an;
43 - hitung = sum(compareA == 0);
44
45 - akurasi_error=(hitung./n)*100;
46 - akurasi_berhasil = 100-akurasi_error
47
48 - mat = T2-an;
49 - m1 = mat.^2;
50 - MSE = mean(m1)

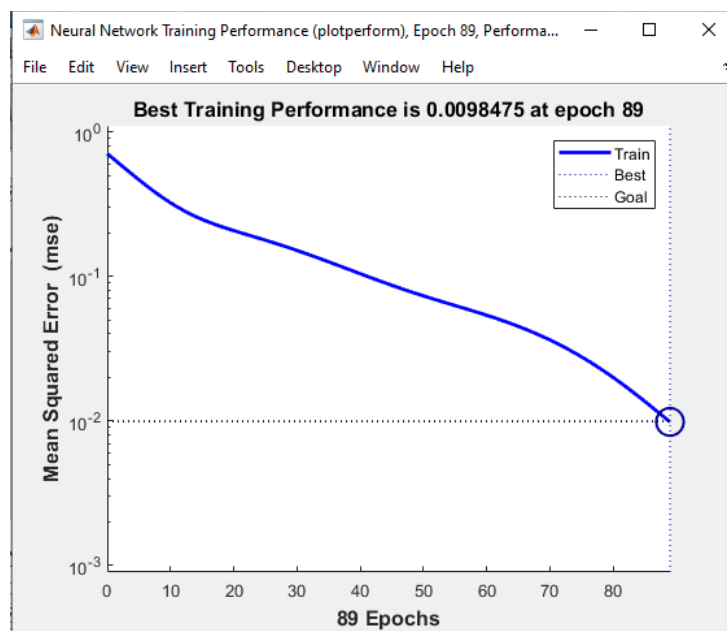
```

sehingga menghasilkan tampilan proses pelatihan jaringan seperti:

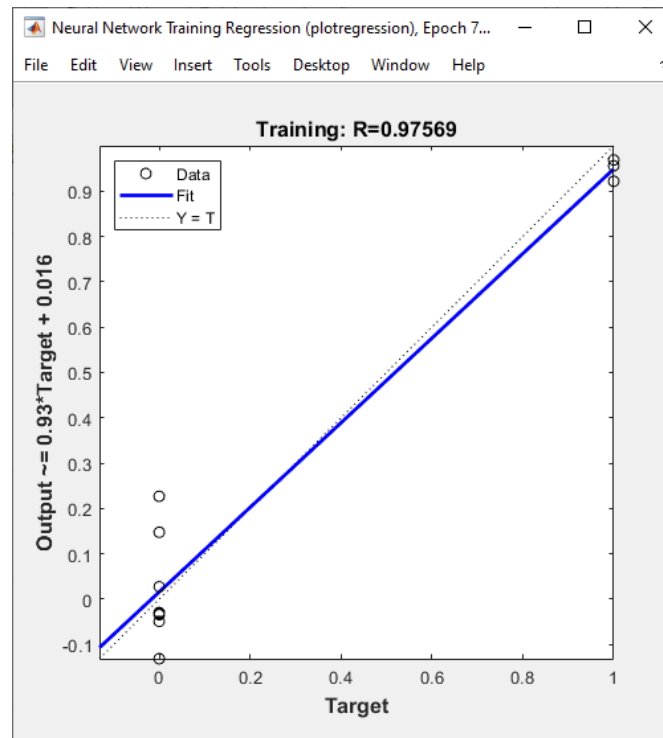
- Hasil run training



dengan iterasi 89 menggunakan pada data ke-11 (laki - laki)



Pada pelatihan ini error goal (MSE) sebesar 0,0098475 dicapai pada epoch yang ke-89 seperti ditunjukkan pada gambar di atas.



garis biru = hasil prediksi suara dan garis putus - putus = target
dapat dinyatakan bahwa hasilnya mendekati akurat

• Source Code Simulasi

```

Editor - D:\klasifikasi suara\simulasi.m
simulasi.m x train.m x +
1  %% Digital Signal Processing Project- Gender Identification and Classification
2  clc;
3  %% Program:
4
5  %Reading a file and getting the fundamental frequency, zero crossing, short energy
6  %values
7  [my2,fs] = audioread('s11.wav');
8
9  %menampilkan grafik
10 figure;
11 plot(my2);title('Test Signal');
12 xlabel('Index');ylabel('Amplitude');
13
14 %mengambil nilai
15 [fundamental_frequency,zero_crossing,short_energy]=Charac_features(my2,fs);
16
17 %% prediksi
18 load train1
19 fundamental_frequency
20 zero_crossing
21 short_energy
22
23 %normalisasi data
24 normalS1 = (fundamental_frequency-104.1667)/(217.3913-104.1667);
25 normalS2 = (zero_crossing-3.1818)/(25.3393-3.1818);

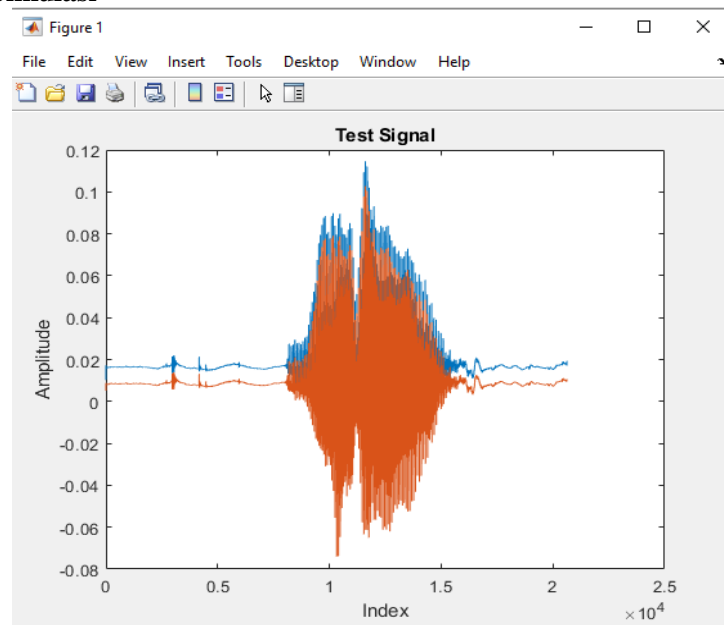
```

```

Editor - D:\klasifikasi suara\simulasi.m
simulasi.m  train.m  +
17  %% prediksi
18  load train1
19  fundamental_frequency
20  zero_crossing
21  short_energy
22
23  %normalisasi data
24  normalS1 = (fundamental_frequency-104.1667)/(217.3913-104.1667);
25  normalS2 = (zero_crossing-3.1818)/(25.3393-3.1818);
26  normalS3 = (short_energy-0.1278)/(8.2358-0.1278);
27
28  Uji = [normalS1;normalS2;normalS3];
29
30  %% simulasi
31  an=round(sim(train1,Uji));
32
33  switch an
34      case 0
35          hasil = 'Perempuan'
36      case 1
37          hasil = 'Laki - Laki'
38  end
39
40  %play sound
41  sound(my2,fs)

```

- Hasil Simulasi



histogram biru pekat = hasil keluaran JST dan histogram merah = hasil target

KESIMPULAN

Dapat disimpulkan bahwa analisis dari setiap jurnal yang kita kumpulkan yaitu mendeteksi metode backpropagation dapat digunakan sebagai sarana untuk mengetahui tingkat deteksi suara manusia dan dapat membedakan antara suara perempuan dan laki - laki. Dengan tingkat akurasi yang baik dan mendekati akurasi 100% dengan menggunakan lebih dari 5 data.

DAFTAR PUSTAKA

<http://ejournal.uki.ac.id/index.php/edumatsains/article/download/372/277>

<https://docplayer.info/42192173-Pengenalan-suara-manusia-dengan-menggunakan-jaringan-saraf-tiruan-model-propagasi-balik.html>

http://www.jurnal.stmikpontianak.ac.id/file/David_-_Buana_Informatika_Vol4_No1_2013.pdf

<http://faculty.petra.ac.id/resmana/basiclab/publications/voicerecog.doc>

http://eprints.undip.ac.id/25298/1/M_L2F004490.pdf