

# Pengenalan Tutur Vokal Bahasa Indonesia Menggunakan Metode *Multi Layer Perceptron*

Risky Via Yuliantari<sup>1</sup>, Risanuri Hidayat<sup>1</sup>, Oyas Wahyunggoro<sup>1</sup>, Anan Nugroho<sup>1</sup>

<sup>1</sup>Departemen Teknik Elektro dan Teknik Informasi Fakultas Teknik

Universitas Gadjah Mada

Jalan Grafika 2 Yogyakarta 55281

<sup>1</sup>rviayuliantari@gmail.com

## ABSTRACT

*Indonesian language is a national language of Indonesia which can be interpreted as sound symbol that can be used as a communication device. In that communication process, information in the form of signal can be obtained. There are many researches on voice signal of Indonesian language. In this research, Indonesian vocal signal recognition automatization is developed by using Multilayer Perceptron to increase accuracy, sensitivity, specificity and precision. The recognition is done by using Discrete Wavelet Transform to 100 speakers that result in 500 characteristic data for training and testing. The classification process results inaccuracy as much as 95,9%, sensitivity 90,2%, specificity 97,5%, and precision 89,8%.*

**Keyword:** *Multi Layer Perceptron, Discrete Wavelet Transform, Vocal, Indonesian*

## INTISARI

Bahasa Indonesia merupakan bahasa nasional bangsa Indonesia yang dapat diartikan sebagai lambang bunyi yang digunakan sebagai alat komunikasi. Dalam proses komunikasi tersebut dapat diperoleh sebuah informasi berupa isyarat. Banyak penelitian tentang isyarat tutur Indonesia. Pada penelitian ini dikembangkan otomatisasi pengenalan isyarat vokal Indonesia dengan *Multi Layer Perceptron* untuk meningkatkan akurasi, sensitifitas, spesifitas dan presisi. Pengenalan dilakukan dengan *Discrete Wavelet Transform* pada 100 orang penutur menghasilkan 500 data ciri untuk *training* dan *testing*. Proses klasifikasi menghasilkan akurasi sebesar 95,9%, sensitifitas 90,2%, spesifitas 97,5%, dan presisi 89,8%.

**Kata Kunci:** *Multi Layer Perceptron, Discrete Wavelet Transform, Vokal, Bahasa Indonesia*

## PENDAHULUAN

Bahasa Indonesia merupakan bahasa nasional negara Indonesia. Bahasa dapat diartikan sebagai lambang bunyi yang digunakan oleh suatu anggota masyarakat untuk bekerja sama, berkomunikasi, dan mengidentifikasi diri (Dr. Alek 2012). Secara umum, bunyi bahasa dikelompokkan menjadi dua kelas bunyi utama, yaitu konsonan (*consonant*) dan vokal (*vowel*). Konsonan adalah bunyi bahasa yang dihasilkan dengan keadaan rongga mulut atau hidung tertutup kemudian dihambat oleh alat ucap yang dialirkan melewati celah sempit.. Sedangkan vokal merupakan bunyi bahasa yang dihasilkan alat ucap melalui rongga mulut dengan udara tanpa hambatan dan memiliki kualitas yang bergantung pada posisi, tinggi rendah dan maju mundur lidah. Bahasa Indonesia mengandung 5 vokal yaitu /a/, /i/, /u/, /e/, dan /o/ (Yendra 2016). Bunyi vokal mengandung ciri-ciri sebagai penyampai informasi pada telinga manusia berbentuk isyarat stasioner yang berderau dengan beda

kecepatan (Sakoe & Chiba 1978).

Penelitian tentang pengenalan isyarat tutur telah banyak dilakukan dalam berbagai bahasa dunia. Sementara penelitian tentang pengenalan isyarat tutur menggunakan bahasa Indonesia masih terbatas. Beberapa penelitian pengenalan isyarat tutur menggunakan bahasa Indonesia dilakukan oleh (Sutisna 2013) menggunakan metode *Neuro Fuzzy Inference System* (ANFIS) sebagai pengenalan isyarat tutur kata yang terisolasi dengan tingkat akurasi sebesar 89,33%, (H et al. 2015) menggunakan *Hidden Markov Models* (HMM) sebagai pengenalan isyarat tutur berbasis suku kata bahasa Indonesia dengan tingkat dengan tingkat akurasi 75%, (Ivana n.d.) melakukan pengenalan vokal Bahasa Indonesia menggunakan metode *Linear Predictive Coding* (LPC) dengan tingkat akurasi 86,9%.

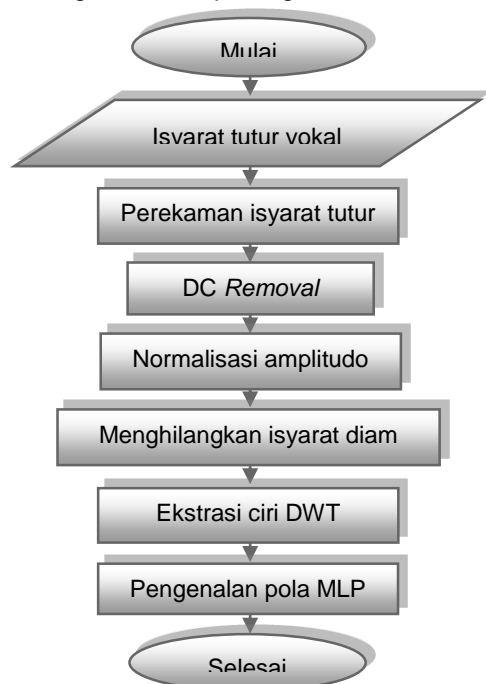
Dalam penelitian ini *Multi Layer Perceptron* (MLP) digunakan sebagai metode klasifikasi pola isyarat tutur vokal bahasa Indonesia. Dimana metode *Multi Layer*

*Perceptron* (MLP) telah banyak diaplikasikan dalam berbagai bidang seperti pengenalan pola tulisan tangan, pengolahan citra medika, dan pengenalan pola suara (David 2013). *Multi Layer Perceptron* (MLP) merupakan salah satu metode pengenalan pola yang efektif. Selain itu, *Multi Layer Perceptron* (MLP) memiliki nilai toleransi terhadap derau, ketidakinieran, dan memiliki akurasi yang tinggi (Gaafar et al. 2014). Akurasi pengenalan juga dipengaruhi oleh proses ekstraksi ciri. Menurut (Asni 2014) *Discrete Wavelet Transform* (DWT) lebih efektif digunakan sebagai metode ekstraksi karena dapat menghemat waktu komputasi. Metode ekstraksi ciri *Discrete Wavelet Transform* (DWT) juga mampu mengatasi isyarat tutur yang memiliki sifat tidak stasioner, mengandung derau, dan memiliki perbedaan kecepatan (Wu et al. 2008).

Proses pengenalan isyarat tutur dilakukan dengan memasukkan nilai ekstraksi ciri yang diberi nilai bobot, kemudian dilakukan penjumlahan antara perkalian nilai ekstraksi ciri dengan nilai bobot. Hasil penjumlahan kedua nilai tersebut menjadi nilai masukan bagi fungsi sigmoid biner untuk mendapatkan nilai keluaran (David 2013) (Hidayatno et al. 2008)

## METODOLOGI

Tiap tahap proses pengenalan isyarat tutur digambarkan pada gambar 1.



Gambar 1 Proses pengenalan isyarat tutur

## PEREKAMAN ISYARAT TUTUR

Proses perekaman dilakukan dengan cara merekam isyarat tutur vokal /a/, /i/, /u/, /e/, dan /o/ selama tiga detik dengan frekuensi pencuplikan (fs) sebesar 8000 Hz. Perekaman dilakukan pada 100 orang penutur pria dan wanita. Frekuensi pencuplikan (fs) pada pengolahan isyarat tutur dipengaruhi oleh resolusi isyarat dan waktu komputasi. Frekuensi pencuplikan (fs) harus memenuhi kriteria Nyquist pada persamaan (1).

$$f_s \geq 2f \quad (1)$$

## DC REMOVAL

*DC removal* merupakan tahapan untuk menghilangkan komponen DC atau *DC offset*. Penghilangan unsur DC dilakukan menggunakan persamaan (2) dengan menghitung nilai rata-rata amplitudo isyarat tutur yang dilanjutkan dengan mengurangi isyarat tutur asli dengan nilai rata-rata tersebut.

$$DC_{offset}(n) = S(n) - \frac{\sum_{n=1}^N S(n)}{N} \quad (2)$$

## NORMALISASI AMPLITUDO

Normalisasi amplitudo dilakukan untuk mengatasi tingkat energi yang tidak konsisten antara tiap isyarat. Sehingga dapat meningkatkan kualitas ciri dan memiliki standar pengukuran yang sama untuk semua data. Proses normalisasi amplitudo diperoleh dengan cara membagi setiap nilai  $S(n)$  pada runtun ke- $n$  dengan nilai absolut amplitudo tertinggi yang terdapat pada isyarat dengan nilai batasan maksimal antara -1 dan 1, dirumuskan dengan persamaan (3).

$$S_{nor}(n) = \frac{DC_{offset}(n)}{\max(\text{abs}(DC_{offset}(n)))} \quad (3)$$

## MENGHILANGKAN ISYARAT DIAM

Proses penghilangan isyarat diam bertujuan untuk mengefektifkan komputasi karena adanya derau dan isyarat diam yang tidak memiliki informasi dalam pengolahan isyarat tutur. Langkah penghilangan isyarat diam dilakukan dengan membagi frame berukuran 20 ms pada proses sebelumnya. Kemudian menentukan nilai ambang bawah amplitudo sebesar 0,05. Nilai ambang digunakan untuk menyeleksi dan mengurutkan kembali masing-masing frame

yang berisi informasi isyarat tutur. Jika nilai maksimum frame sama dengan atau dibawah dari nilai ambang, maka frame tersebut tidak digunakan.

Proses menghilangkan suara diam dirumuskan dengan persamaan (4).

$$F_i = (s(n))_{n=(i-1)*N+1}^{i*N} \quad (4)$$

### DISCRETE WAVELET TRANSFORM (DWT)

*Wavelet* merupakan gelombang singkat dengan durasi terbatas yang memiliki nilai rata-rata nol. *Wavelet* mengkonsentrasikan energinya terhadap ruang dan waktu sehingga cocok untuk menganalisis isyarat yang bersifat sementara. Ada dua jenis yaitu *Continuous Wavelet Transform* dan *Discrete Wavelet Transform*.

*Discrete Wavelet Transform* (DWT) digunakan untuk mentransformasikan isyarat dari domain waktu ke domain frekuensi yang dapat diaplikasikan pada data diskrit untuk menghasilkan keluaran diskrit (Asni 2014). *Discrete Wavelet Transform* (DWT) dikatakan sebagai *Low Pass Filter* (LPF) dan *High Pass Filter* (HPF). Frekuensi tinggi dan rendah dipisahkan dari sinyal asli dengan menggunakan transformasi dekomposisi, Semakin rendah pendekatan sinyal frekuensi maka semakin tinggi sinyal frekuensi yang dihasilkan (Ghule & Deshmukh n.d.).

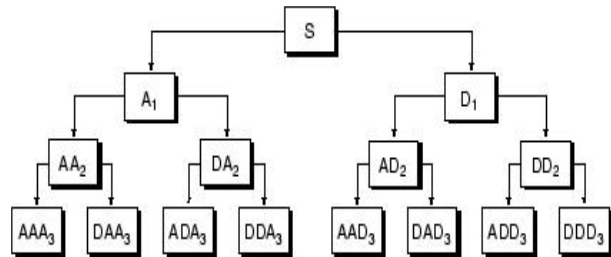
*Low Pass Filter* (LPF) maupun *High Pass Filter* (HPF) merupakan salah satu fungsi yang paling banyak digunakan pada pemrosesan sinyal. Perwujudan *wavelet* dapat berupa penskalaan ulang dengan iterasi. Resolusi sinyal diukur dari jumlah informasi sinyal ditentukan oleh operasi *filtering* dan menggunakan skala operasi *upsampling* dan *downsampling* (R & P 2009)(Ali et al. 2014). Perhitungan *Discrete Wavelet Transform* (DWT) dapat dilakukan dengan menkonvolusi koefisien LPF (h) dan HPF (g) yang ditunjukkan pada persamaan (5) dan (6) (Asni 2014).

$$a_k^{(j+1)} = \sum_{n=-\infty}^{\infty} h_{n-2k} a_n^{(j)} = (a^{(j)} * h^{(0)})(2k) \quad (5)$$

$$d_k^{(j+1)} = \sum_{n=-\infty}^{\infty} g_{n-2k} a_n^{(j)} = (a^{(j)} * g^{(1)})(2k) \quad (6)$$

Ekstraksi ciri merupakan tahapan terpenting dalam sistem pengenalan isyarat tutur. Metode ekstraksi ciri *Discrete Wavelet Transform* (DWT) level 3 digunakan untuk

menghasilkan *sub-band* sebanyak  $2^i$  dimana  $j$  adalah level dekomposisi dari metode *Discrete Wavelet Transform* (DWT). Proses dekomposisi *wavelet* level 3 ditunjukkan pada gambar 2.



Gambar 2. Proses dekomposisi *wavelet* level 3

Vektor ciri didapatkan dengan cara menghitung energi tiap frekuensi *sub-band* dan energi total menggunakan persamaan (7) dan (8).

$$E_i = \sqrt{\sum_{k=1}^N |X_i(k)|^2} \quad (7)$$

$$E_{total} = \sqrt{\sum_{i=1}^I E_i^2} \quad (8)$$

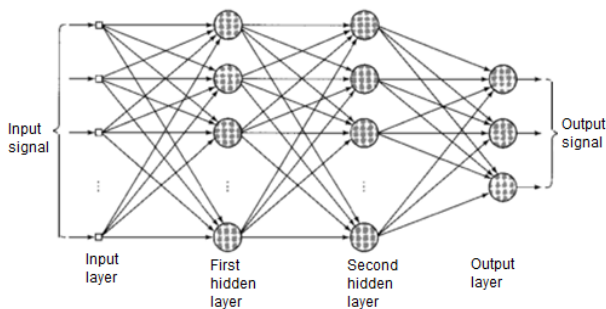
Agar vektor ciri memiliki nilai standar yang sama perlu dilakukan normalisasi energi dengan cara membagi setiap total energi *sub-band* dengan total energi *sub-band* yang digunakan menggunakan persamaan (9).

$$V_{energi} = \frac{E_i}{E_{total}} \quad (9)$$

Setelah dilakukan normalisasi energi maka dihasilkan delapan nilai vektor ciri yang selanjutnya akan digunakan dalam proses pengenalan pola.

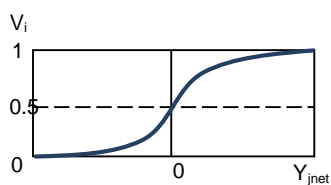
### MULTI LAYER PERCEPTRON (MLP)

Sebuah *Neural Networks* terdiri dari kumpulan satu set unit sensor yang merupakan input, satu atau lebih lapisan tersembunyi (*hidden layers*) dan lapisan output. Sinyal input menyebar maju secara lapis demi lapis melalui lapisan tersembunyi dengan nilai bobot yang selalu berubah, sehingga disebut sebagai *Multi Layer Perceptron* (MLP). Arsitektur MLP dapat dilihat pada gambar 2 (Haylin 1999).



Gambar 2. Arsitektur *Multi Layer Perceptron*

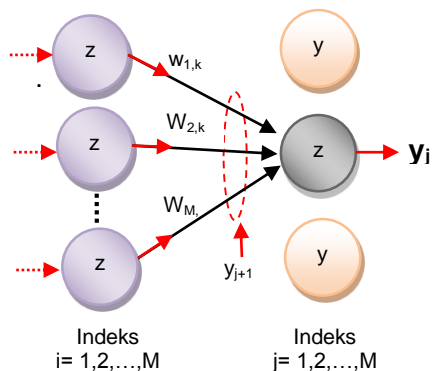
Pada lapisan input terkandung nilai-nilai kuantitatif yang memiliki karakter unik digunakan sebagai obyek klasifikasi pengenalan pola. Nilai kuantitatif berasal dari nilai masukan proses ekstraksi ciri akan menghasilkan nilai keluaran berupa fungsi sigmoid biner sebagaimana ditunjukkan oleh gambar 3 dan persamaan (10).



Gambar 3. Fungsi Sigmoid biner

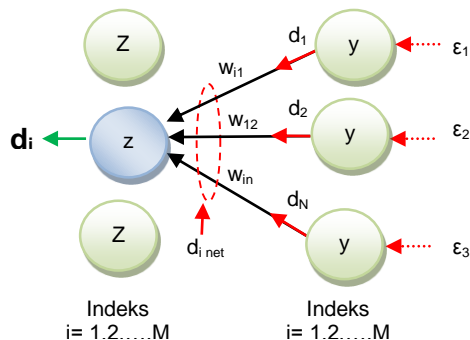
$$Y_j = \frac{1}{1 + e^{-ay_{jnet}}} \text{ dengan } 0 < a < 1 \text{ (10)}$$

*Multi Layer Perceptron* (MLP) merupakan pembelajaran terbimbing yang menggunakan algoritma *error back-propagation* (EBP). *Error back-propagation* (EBP) memungkinkan penyesuaian nilai bobot tersembunyi dengan cara merambat balikkan error output. Proses pembelajaran EBP ada dua tahapan yaitu arah maju dan arah balik. *Error back-propagation* (EBP) arah maju sebagaimana ditunjukkan pada gambar 4.



Gambar 4. *Error back-propagation* (EBP) arah maju

Sinyal masukan pada *error back-propagation* (EBP) arah maju dirambatkan ke lapisan tersembunyi 1, lapisan tersembunyi 2, hingga ke lapisan keluaran. Nilai keluaran pada setiap lapisan dihitung menggunakan fungsi aktivasi. Sedangkan pada *error back-propagation* (EBP) arah balik perambatan dimulai dari lapisan keluaran ke lapisan tersembunyi 2, lapisan tersembunyi 1, hingga ke lapisan masukan seperti yang ditunjukkan pada gambar 5.



Gambar 5. *Error back-propagation* (EBP) arah balik

Pada sambungan percabangan setiap lapisan bobot dihitung dan diperbarui terus menerus tiap eksemplar dan tiap *epoch*. Satu eksemplar merupakan satu kali pelatihan pada satu data dengan sejumlah ciri input, dimana satu epoch adalah satu kali pelatihan dengan melibatkan keseluruhan data. Untuk menghitung jumlah nilai input pada tiap suatu lapisan neuron yang dimisalkan dengan  $y_j$  menggunakan persamaan (11).

$$y_{jnet} = \sum_{i=1}^M w_{ij} z_i \quad (11)$$

Sedangkan nilai output suatu lapisan neuron dihitung menggunakan persamaan (10), yang dapat dituliskan kembali dengan menggunakan persamaan (12).

$$y_j = \frac{1}{1 + e^{-y_{jnet}}} \quad (12)$$

Dimana persamaan (13) adalah nilai turunan pertama dari persamaan (12). Proses ini dilakukan pada keseluruhan lapisan error back propagation arah maju.

$$y'_j = y_j(1 - y_j) \quad (13)$$

Nilai  $y_j$  dan  $y'_j$  disimpan untuk kemudian dilakukan perhitungan pada *error back propagation* arah balik. Proses selanjutnya adalah menghitung nilai bobot  $w_i$  yang diperbaharui tiap percabangan dengan persamaan (14), (15), dan (16).

$$\Delta w_i = \eta \sum_{j=1}^M (t_i - y_{jk})(y_{jk}(1 - y_j))z_i \quad (14)$$

$$\Delta w_i = \eta \sum_{j=1}^M (t_i - y_{jk})y'_j z_i \quad (15)$$

$$w_i \text{ baru} = w_i \text{ lama} + \Delta w_i \quad (16)$$

Dengan  $\eta$  adalah pesat belajar (learning rate) dengan nilai antara 0 sampai dengan 1 dan  $t_i$  adalah target pelatihan yang dijadikan acuan dalam pembelajaran.

Perhitungan *error factor*  $d$  pada *error back propagation* arah balik suatu lapisan digunakan persamaan (17) dan (18).

$$d_j = (t_j - y_j)y'_j \quad (17)$$

$$d_j = \varepsilon_j y'_j \quad (18)$$

Sedangkan untuk perhitungan neuron-neuron pada lapisan lainnya menggunakan persamaan (19) dan (20).

$$d_{inet} = \sum_{j=1}^M w_{ij} d_j \quad (19)$$

$$d_i = d_{inet} y'_j \quad (20)$$

Persamaan (21) digunakan untuk menghitung nilai bobot yang terhubung ke neuron  $z_i$  pada *error back propagation* arah balik.

$$\Delta w_i = \eta d_j z_i \quad (21)$$

Dimana  $z_i$  merupakan hasil *perhitungan error back propagation* arah maju sebelumnya. Kemudian nilai  $z_i$  dirambatkan balik ke lapisan berikutnya hingga sampai kelapisan input dalam satu jaring MLP. Proses dilanjutkan untuk eksemplar input berikutnya sampai selesai satu epoch. Jika nilai error masih besar maka proses epoch dilanjutkan hingga mendapatkan nilai error yang sangat kecil.

Proses pembelajaran menggunakan Multi Layer Perceptron (MLP) tidak sepenuhnya sesuai dengan target klasifikasi

dalam proses pengujian. Maka untuk mencapai keberhasilan pengenalan indeks pengukuran ditentukan berdasarkan keakuratan, sensitifitas, spesifitas, dan presisi.

Keakuratan hasil klasifikasi dapat diukur dengan cara membagi antara jumlah klasifikasi yang benar sesuai target dengan jumlah klasifikasi yang berbeda dengan target dari semua kelas menggunakan rumus yang ditunjukkan oleh persamaan (22) (Carvalho et al. 2014).

$$acc = \frac{\sum TP + \sum TN}{\sum TP + \sum TN + \sum FP + \sum FN} \times 100\% \quad (23)$$

Sensitivitas merupakan sebuah ukuran kemampuan prediksi untuk memilih contoh kelas tertentu dari serangkaian data set dan bersesuaian dengan *True Positive Rate* (TPR) yang dirumuskan pada persamaan (24).

$$TPR = \frac{\sum TP}{\sum TP + \sum FN} \times 100\% \quad (24)$$

Spesifitas atau bisa disebut dengan *true negative rate* (TNR) berkaitan dengan sensitivitas, dimana spesifitas merupakan sebuah ukuran yang biasanya digunakan dalam permasalahan dua kelas yang dapat dirumuskan oleh persamaan (25).

$$TNR = \frac{\sum TN}{\sum TN + \sum FP} \times 100\% \quad (25)$$

Presisi disebut juga sebagai *positive predictive value* (PPV) yang merupakan kondisi dimana prediksi benar dihitung menggunakan persamaan (26).

$$PPV = \frac{\sum TP}{\sum TP + \sum FP} \times 100\% \quad (26)$$

Dimana TP (*true positive*) adalah jumlah data positif target yang terklasifikasi positif pada sistem, TN (*true negative*) adalah jumlah data negative pada target yang terklasifikasi pada sistem. FP (*false positive*) representasi jumlah data negatif pada target yang terklasifikasi positif pada sistem dan FN (*false negatif*) merupakan jumlah data positif pada target yang terklasifikasi negatif pada sistem. Kesemua nilai ini terwujud dalam sebuah matrik confusion (Powers 2007)

## HASIL DAN PEMBAHASAN

Nilai vektor ciri diperoleh berdasarkan nilai normalisasi energi frekuensi *sub-band* yang berada pada level 3 dengan 8 ciri yang memiliki panjang 4000 sampel. Dimana tiap *sub-band* memiliki lebar yang sama sebesar 500 Hz.

Hasil ekstraksi ciri tiap vokal ditunjukkan pada tabel 1 sampai tabel 5.

Tabel 1 Vektor ciri vokal /a/

Frekuensi <i>sub-band</i> (Hz)	Vektor ciri			
	a1	a2	...	a100
500	0,1248	0,1008	...	0,0859
1000	0,3256	0,3431	...	0,4816
1500	0,4275	0,4539	...	0,2304
2000	0,5696	0,5576	...	0,6992
2500	0,3249	0,1313	...	0,1742
3000	0,2702	0,2841	...	0,2735
3500	0,2667	0,3177	...	0,1425
4000	0,3485	0,3954	...	0,3054

Tabel 2 Vektor ciri vokal /i/

Frekuensi <i>sub-band</i> (Hz)	Vektor ciri			
	i1	i2	...	i100
500	0,2447	0,1113	...	0,0643
1000	0,2254	0,1692	...	0,0696
1500	0,2255	0,0969	...	0,0983
2000	0,1939	0,1326	...	0,1298
2500	0,3918	0,4472	...	0,8099
3000	0,6403	0,7248	...	0,4428
3500	0,2262	0,1561	...	0,2010
4000	0,2447	0,1113	...	0,0643

Tabel 3 Vektor ciri vokal /u/

Frekuensi <i>sub-band</i> (Hz)	Vektor ciri			
	u1	u2	...	u100
500	0,2411	0,5752	...	0,3560
1000	0,2070	0,6783	...	0,3981
1500	0,5117	0,1779	...	0,1136
2000	0,4428	0,2420	...	0,2680
2500	0,3757	0,1744	...	0,5517
3000	0,4378	0,2609	...	0,4726
3500	0,2273	0,0808	...	0,1014
4000	0,2381	0,1176	...	0,3033

Tabel 4 Vektor ciri vokal /e/

Frekuensi <i>sub-band</i> (Hz)	Vektor ciri			
	e1	e2	...	e100
500	0,0963	0,0567	...	0,1411
1000	0,0964	0,2265	...	0,2723
1500	0,4130	0,3855	...	0,1576
2000	0,2242	0,3556	...	0,2043
2500	0,1723	0,1320	...	0,6565
3000	0,3968	0,3831	...	0,4725
3500	0,6233	0,5429	...	0,1666
4000	0,4300	0,4598	...	0,3966

Tabel 5 Vektor ciri vokal /o/

Frekuensi <i>sub-band</i> (Hz)	Vektor ciri			
	e1	e2	...	e100
500	0,3242	0,2290	...	0,1858
1000	0,4070	0,5483	...	0,6077
1500	0,2998	0,0679	...	0,1039
2000	0,2854	0,6432	...	0,6557
2500	0,2644	0,0624	...	0,2009
3000	0,4377	0,3187	...	0,2431
3500	0,3984	0,0659	...	0,0619
4000	0,3710	0,3445	...	0,2283

Proses selanjutnya adalah menentukan data set berupa ciri vokal bahasa Indonesia sebanyak 500 data yang diperoleh dari 100 orang penutur untuk dilakukan *trainingset* sebanyak 350 data dan 150 data sebagai *testingset* Hasil *trainingset* dan *testingset* ditunjukkan dengan *confusion matrix* seperti pada tabel 6.

Tabel 6 *Confusion matrix* pengenalan pola

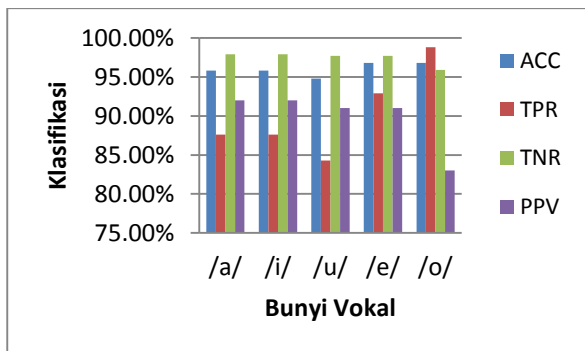
		Kelas Sebenarnya				
		A	I	U	E	O
Prediksi	A	92	1	5	0	7
	I	1	92	3	5	4
	U	6	2	91	4	5
	E	1	4	1	91	1
	O	0	1	0	0	83

Data set tiap vokal bahasa Indonesia pada tabel 7 dihitung berdasarkan nilai pada tabel 6.

Tabel 7 Data set vokal bahasa Indonesia

Vokal	TP	FP	FN	TN
/a/	92	8	13	387
/i/	92	8	13	387
/u/	91	9	17	383
/e/	91	9	7	393
/o/	83	17	1	399

Dari data set tabel 7, dapat dilakukan perhitungan akurasi, sensitifitas, spesifitas, dan persisi pengenalan pola berdasarkan rumus yang telah dijelaskan sebelumnya. Hasil akurasi (acc), sensitifitas (TPR), spesifitas (TNR), dan persisi (PPV) pengenalan pola ditunjukkan pada gambar 6.



Gambar 6. Hasil akurasi (acc), sensitifitas (TPR), spesifitas (TNR), dan persisi (PPV) pengenalan pola

Berdasarkan tabel 8, maka diperoleh rata-rata akurasi pengenalan sebesar 95,9%, rata-rata sensitifitas pengenalan sebesar 90,2%, rata-rata spesifitas pengenalan sebesar 97,5%, dan rata-rata presisi pengenalan pola sebesar 89,8%.

## KESIMPULAN

Dari hasil dan pembahasan dapat disimpulkan bahwa proses pengenalan isyarat tutur dipengaruhi oleh nilai ciri yang dihasilkan. Semakin banyak data yang memiliki variasi nilai ciri maka semakin tinggi nilai akurasi yang didapat. Selain berpengaruh terhadap akurasi, nilai ciri yang bervariasi juga berpengaruh dalam menentukan persentase sensitifitas, spesifitas dan presisi suatu pengenalan.

## DAFTAR PUSTAKA

Ali, H. et al., 2014. DWT features performance analysis for automatic speech recognition of Urdu. *SpringerPlus*, 3, pp.1–10.

Asni, A., 2014. Ekstraksi Ciri Dan Pengenalan Tutur Vokal Bahasa Indonesia Menggunakan Metode Discrete Wavelet Transform (DWT) dan Dynamic Time Warping (DTW). In *Universitas Gadjah Mada*.

Carvalho, L.F. et al., 2014. Digital signature of network segment for healthcare environments support. *Irbm*, 35(6), pp.299–309.

David, S.K., 2013. Penerapan Algoritma Jaringan Saraf Tiruan Backpropagation Untuk Pengenalan Pola Tanda Tagan. *Jurnal Teknologi*, 6 Nomor 2, pp.139–146.

Dr. Alek, P.D.H.A.H., 2012. *Linguistik Umum* N. I. Sallama, ed., Jakarta: Erlangga.

Gaafar, T.S., Abo Bakr, H.M. & Abdalla, M.I., 2014. An improved method for speech/speaker recognition. 2014 *International Conference on Informatics, Electronics and Vision, ICIEV 2014*, pp.2–6.

Ghule, K.R. & Deshmukh, R.R., Feature Extraction Techniques for Speech Recognition: A Review. *International Journal of Scientific & Engineering Research*, 6(5), pp.2229–5518.

H, S., Hidayat, R. & Adji, T.B., 2015. Sistem Pengenal Tutur Bahasa Indonesia Berbasis Suku Kata Menggunakan MFCC , Sistem Pengenal Tutur Bahasa Indonesia Berbasis Suku Kata. *Conference on Information Technology and Electrical Engineering*, (September), pp.246–251.

Haylin, S., 1999. *Neural Networks A Comprehensive Foundation* 2nd ed., New Jersey: Prentice Hall, Inc.

Hidayatno, A. et al., 2008. Perambatan-Balik ( Backpropagation ). , pp.100–106.

Ivana, Pengenalan Ucapan Vokal Bahasa Indonesia Dengan Jaringan Sarag Tiruan Menggunakan Linear Predictive Coding. , pp.1–10.

Powers, D.M.W., 2007. Evaluation: From Precision , Recall and F-Factor to ROC , Informedness , Markedness & Correlation. *J. Mach. Learn. Technol*, 2, pp.37–63.

R, V.K. V & P, B.A., 2009. Features of Wavelet Packet Decomposition and Discrete Wavelet Transform for Malayalam Speech Recognition. *Aceee*, 1(2), pp.93–96.

Sakoe, H. & Chiba, S., 1978. Dynamic Programming Algorithm Optimization for Spoken Word Recognition. In *IEEE Transactions on Acoustic Speech and Signal Processing*. pp. 43–49.

Sutisna, U., 2013. Pengenalan Tutar Kata Terisolasi Menggunakan MFCC dan ANFIS. , p.136.

Wu, X., Tian, F. & Liu, J., 2008. An improved speech feature extraction algorithm using DWT. *ICALIP 2008 - 2008 International*

*Conference on Audio, Language and Image Processing, Proceedings*, pp.1086–1090.

Yendra, 2016. *Mengenal Ilmu Bahasa (Linguistik)* 1st ed., Yogyakarta.