**SISTEM PENGENALAN PENGUCAP MANUSIA DENGAN EKSTRAKSI CIRI MFCC DAN ALGORITMA JARINGAN SARAF TIRUAN PERAMBATAN BALIK SEBAGAI PENGENALANNYA**

**PADA SISTEM KEHADIRAN**

**Tesis**

**untuk memenuhi sebagian persyaratan**

**mencapai derajat Sarjana S-2 Program Studi**

**Magister Sistem Informasi**



**Oleh :**

**Eko Riyanto**

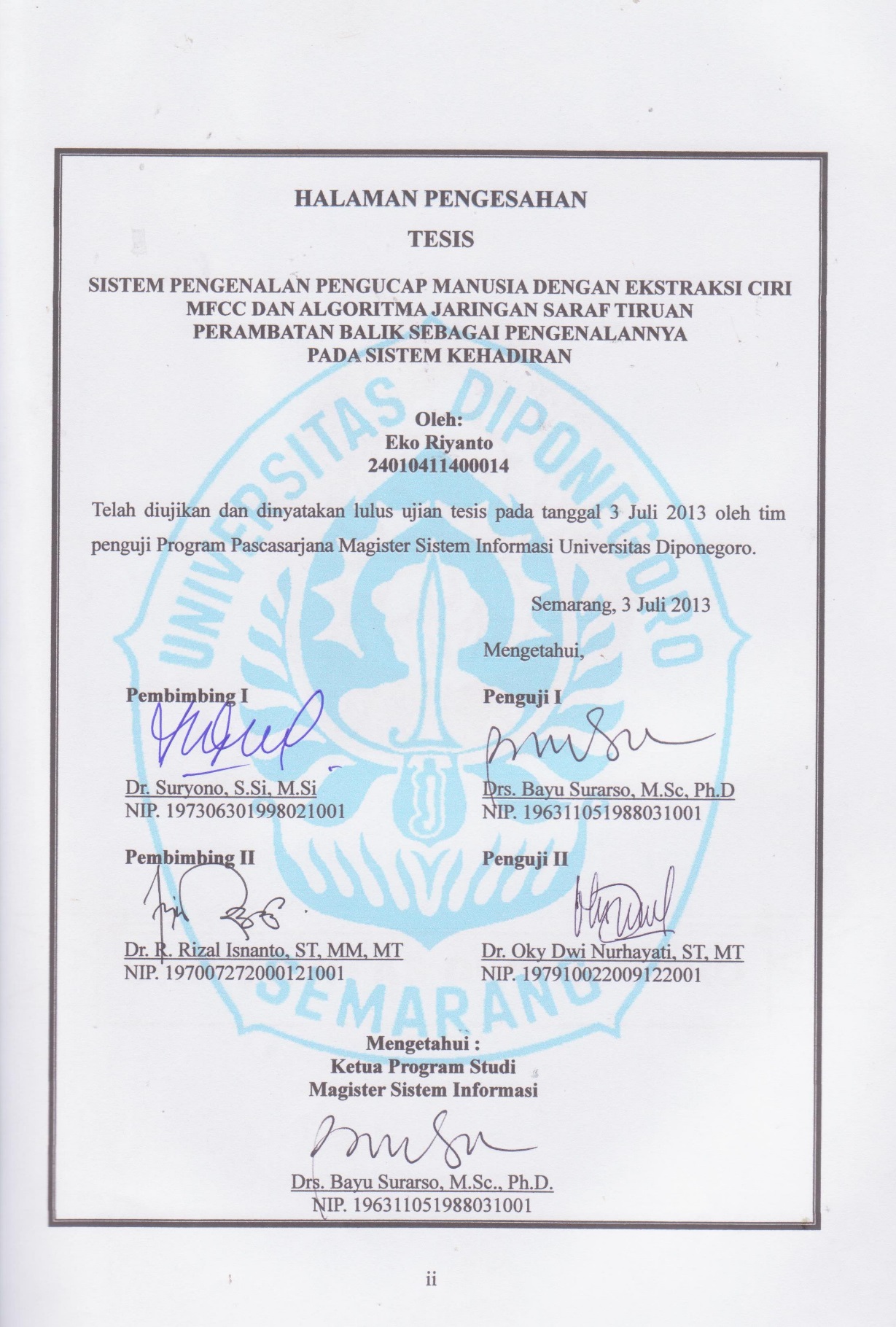
**24010411400014**

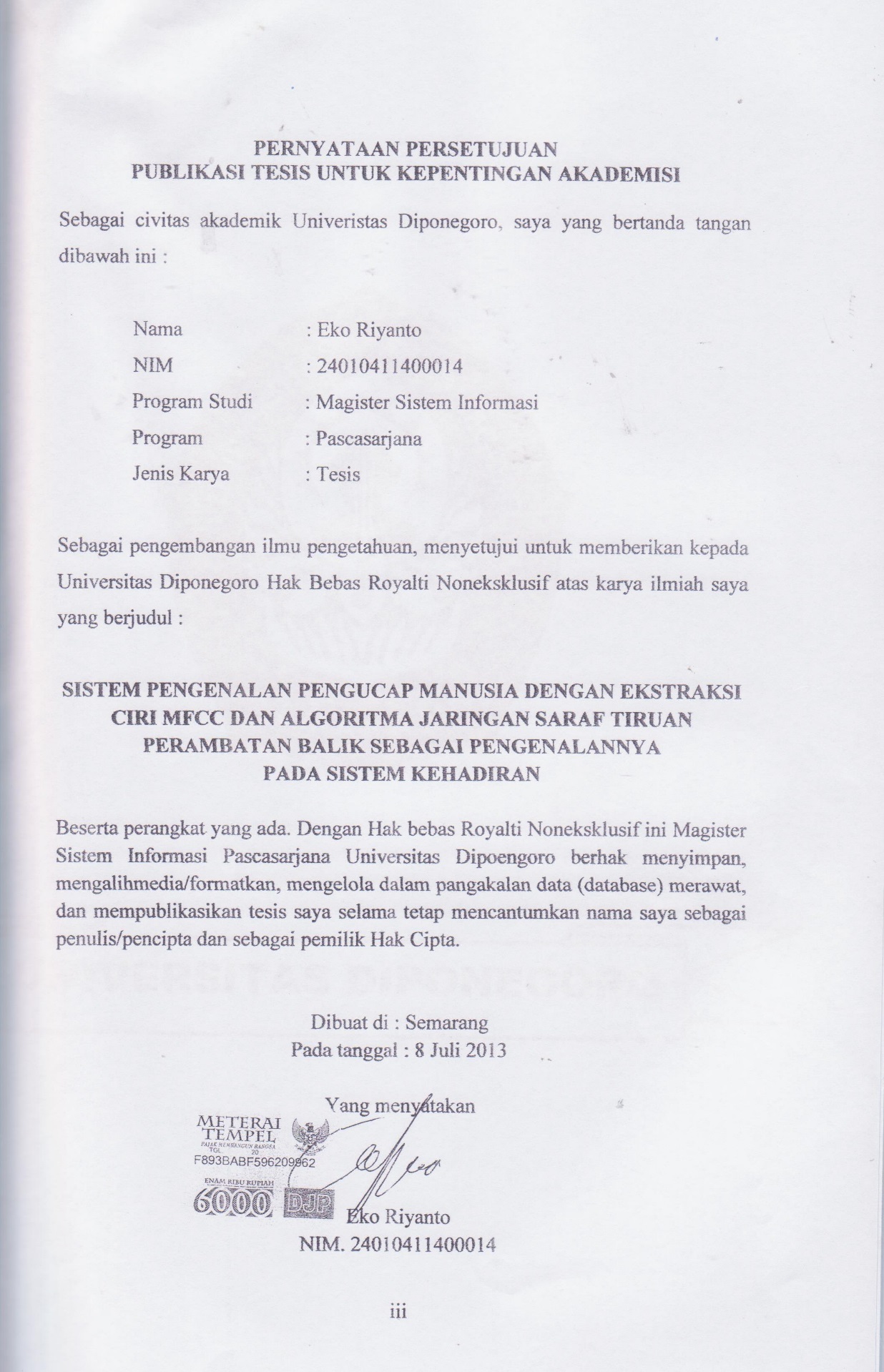
**PROGRAM PASCASARJANA**

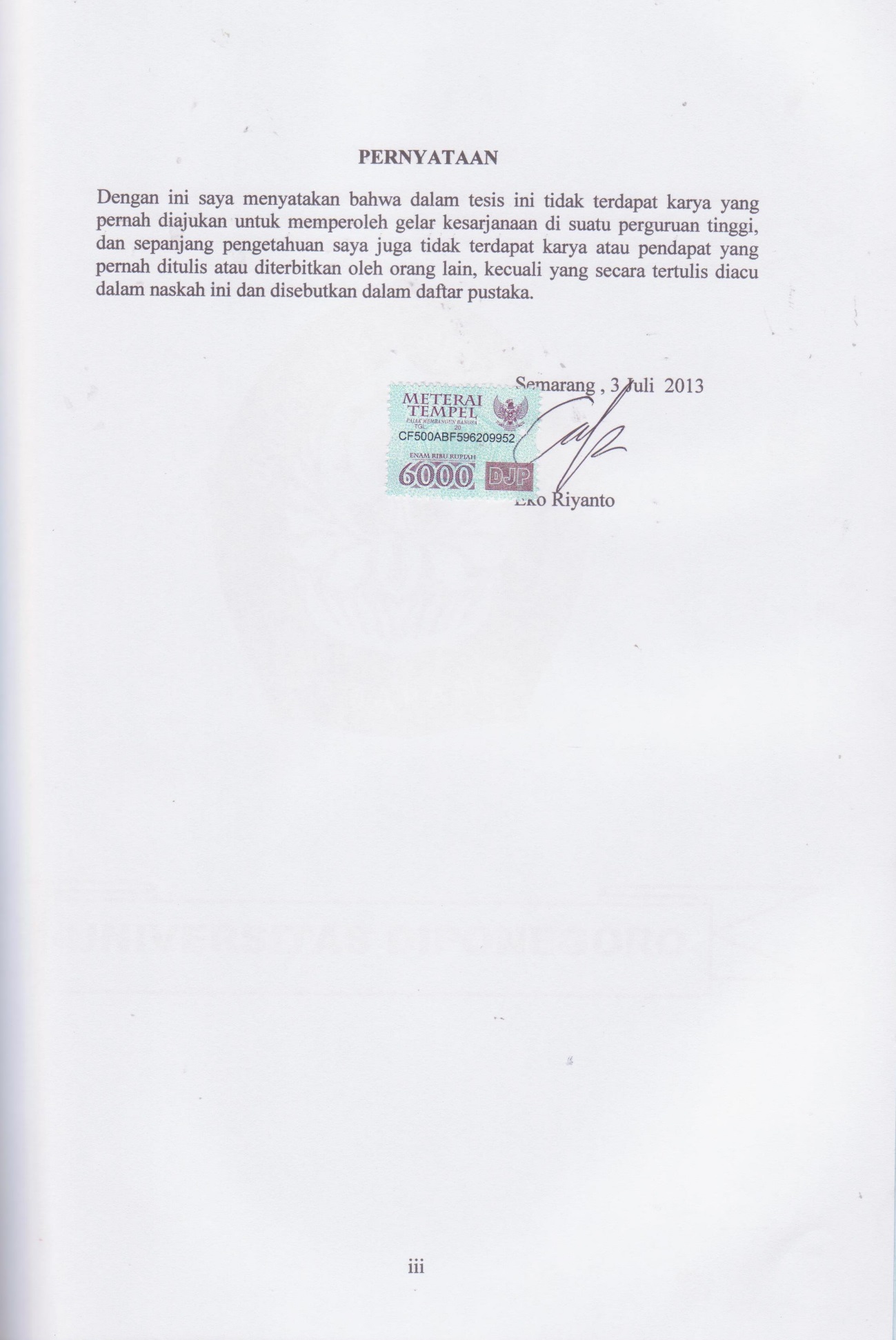
**UNIVERSITAS DIPONEGORO**

**SEMARANG**

**2013**

****





**KATA PENGANTAR**

Puji syukur kepada Tuhan Yang Maha Esa oleh karena rahmat, taufik dan hidayah-Nya sehingga penulis dapat menyelesaikan tesis dengan judul“Sistem Pengenalan Pengucap Manusia Dengan Ekstraksi Ciri *MFCC* dan Algoritma Jaringan Saraf Tiruan Perambatan Balik Sebagai Pengenalannya Pada Sistem Kehadiran”.

Keberhasilan dalam menyusun tesis ini tidak lepas dari bantuan berbagai pihak yang dengan tulus ikhlas memberikan masukan demi sempurnanya tesis ini. Oleh karena itu, dalam kesempatan ini dengan kerendahan hati penulis mengucapkan terima kasih kepada :

1. Drs. Bayu Surarso, M.Sc., Ph.D. sebagai Ketua Program Studi Magister Sistem Informasi, Program Pascasarjana, Universitas Diponegoro Semarang.
2. Dr. Suryono, S.Si, M.Si selaku dosen pembimbing I yang senantiasa memberikan arahan-arahan yang sangat membantu penulis dalam mengerjakan tesis ini.
3. Dr. R. Rizal Isnanto, ST, MM, MT selaku dosen pembimbing II yang senantiasa memberikan arahan-arahan yang sangat membantu penulis dalam mengerjakan tesis ini.
4. Semua pihak yang tidak bisa disebutkan namanya satu persatu, untuk bantuan, doa dan dukungannya.

Penulis menyadari sepenuhnya bahwa tesis ini jauh dari kesempurnaan, namun penulis berharap semoga kerja ini bernilai karya yang dapat memberikan sumbangan bagi kemajuan ilmu pengetahuan dan bermanfaat bagi pembacanya.

Semarang, 3 Juli 2013

Eko Riyanto

**DAFTAR ISI**

HALAMAN JUDUL i

HALAMAN PENGESAHAN ii

HALAMAN PERNYATAAN iii

KATA PENGANTAR iv

DAFTAR ISI v

DAFTAR GAMBAR vii

DAFTAR TABEL viii

DAFTAR LAMPIRAN ix

DAFTAR ARTI LAMBANG DAN SINGKATAN x

ABSTRAK xi

ABSTRACK xii

BAB I PENDAHULUAN 1

1.1 Latar Belakang 1

1.2 Tujuan Penelitian 3

1.3 Manfaat Penelitian 3

BAB II TINJAUAN PUSTAKA 4

2.1 Tinjauan Pustaka 4

2.2 Dasar Teori 5

2.2.1 Pengenalan Pola 5

2.2.2 Autentikasi 6

2.2.3 Sinyal Percakapan 7

2.2.4 Teknologi SAPI 8

2.2.5 Ekstraksi Ciri MFCC 13

2.2.6 PCA (*Principal Components Analysis*) 17

2.2.6 Jaringan Saraf Tiruan 18

BAB III METODE PENELITIAN 30

3.1 Bahan Penelitian 30

3.2 Alat Penelitian 30

3.3 Prosedur Penelitian 30

3.3.1 Metode Pengumpulan Data 31

3.3.2 Metode Pengembangan Perangkat Lunak 31

3.4 Analisis Kebutuhan 33

3.5 Pembuatan Rancangan Tampilan 34

3.5.1 Desain Utama 34

3.5.2 Desain Pelatihan 34

3.5.3 Desain Pengujian 36

3.6 Perancangan 37

3.6.1 Tahap Akuisisi Data 39

3.6.2 Tahap Ektraksi Ciri 39

3.6.3 Tahap Pelatihan 40

3.6.4 Tahap Pengujian 41

3.6.5 Tahap Keputusan 42

3.7 Algoritma Program 42

3.8 Tahap Implementasi dan Pengujian 44

BAB IV HASIL PENELITIAN DAN PEMBAHASAN 45

4.1 Hasil Penelitian 45

4.1.1 Akuisisi Data 46

4.2 Pelatihan 47

4.1.3 Pengujian Presensi 48

4.1.4 Ekstraksi Ciri dan Pelatihan Jaringan Saraf Tiruan 49

4.2 Pembahasan 52

4.2.1 Pengujian Sistem 52

BAB V KESIMPULAN 71

5.1 Kesimpulan 71

5.2 Saran 71

DAFTAR PUSTAKA 72

LAMPIRAN 74

**DAFTAR GAMBAR**

Gambar 2.1 Arsitektur SAPI 8

Gambar 2.2 Proses ektraksi ciri *MFCC* 13

Gambar 2.3 Konsep jaringan saraf tiruan perambatan balik 26  
Gambar 3.1 Desain penelitian 31

Gambar 3.2 Siklus hidup pengembangan sistem (*O’Brien*, 1990) 32  
Gambar 3.3 Model pengembangan perangkat lunak 32

Gambar 3.4 Desain utama sistem pengenalan suara 34

Gambar 3.5 Desain pelatihan 35

Gambar 3.6 Desan pengujian 36

Gambar 3.7 Percancangan sistem pengenalan 37

Gambar 3.8 Diagram air tahap pemasukan dan pencocokan data 38  
Gambar 3.9 Ekstraksi ciri salah satu suara dengan *MFCC* 40

Gambar 3.10 Grafik *performance* salah satu suara 41

Gambar 3.11 Algortima pemrograman 43

Gambar 4.1 Tampilan halaman utama 45

Gambar 4.2 Aplikasi free sound recoder 46

Gambar 4.3 Sisem file suara pada folder *sample\_suara* 46

Gambar 4.4 Tampilan halaman pelatihan 47

Gambar 4.5 Tampilan presensi dosen 48

Gambar 4.6 Konsep ADC 49

Gambar 4.7 Diagram blok *MFCC* 50

Gambar 4.8 Identifikasi suara yang dikenali di basis data 53

**DAFTAR TABEL**

Tabel 4.1 Pelatihan jaringan saraf tiruan perambatan balik 51

Tabel 4.2 Proses identifikasi suara Ali Wedo Sarjono 54

Tabel 4.3 Proses identifikasi suara Eko Riyanto 54

Tabel 4.4 Proses identifikasi suara Bedy Suprapti 55

Tabel 4.5 Proses identifikasi suara Mariza Devega 55

Tabel 4.6 Proses identifikasi suara Muhammad Masrur 56

Tabel 4.7 Proses identifikasi suara Rachmawati Suwono 56

Tabel 4.8 Proses identifikasi suara Toni Wijanarko 57

Tabel 4.9 Proses identifikasi suara Agus Sifaunnajah 57

Tabel 4.10 Proses identifikasi suara Akhmad Qashlim 58

Tabel 4.11 Proses identifikasi suara Amaludin Arifia 58

Tabel 4.12 Proses identifikasi suara Anindita Kusuma Dewi 59

Tabel 4.13 Proses identifikasi suara Asrul Azhari Muin 59

Tabel 4.14 Proses identifikasi suara Budi Prasetyo 60

Tabel 4.15 Proses identifikasi suara Daniel Alfa Puryono 60

Tabel 4.16 Proses identifikasi suara Daniel Wahyu S.p 61

Tabel 4.17 Proses identifikasi suara Herlina 61

Tabel 4.18 Proses identifikasi suara Hevi Herlina Ullu 62

Tabel 4.19 Proses identifikasi suara Mukrodin 62

Tabel 4.20 Proses identifikasi suara T. Indriati Wardani 63

Tabel 4.21 Proses identifikasi suara Kasmawi 63

Tabel 4.22 Proses identifikasi suara Mahmudi 64

Tabel 4.23 Proses identifikasi suara Muhammad Yusuf 64

Tabel 4.24 Proses identifikasi suara Priyo Nugroho Adi 65

Tabel 4.25Proses identifikasi suara Nur Azizah 65

Tabel 4.26 Proses identifikasi suara Purwanto 66

Tabel 4.27 Proses identifikasi suara Rima Maulini 66

Tabel 4.28 Proses identifikasi suara Sugeng Murdowo 67

Tabel 4.29 Proses identifikasi suara Teguh 67

Tabel 4.30 Proses identifikasi suara Sutejo 68

Tabel 4.31 Proses identifikasi suara Yulius Sumantri 68

Tabel 4.32 Identifikasi Keseluruhan 69

**DAFTAR LAMPIRAN**

Lampiran 1. Tampilan Program 74

**DAFTAR ARTI LAMBANG**

DAFTAR ARTI LAMBANG

|  |  |
| --- | --- |
| Lambang | Arti Lambang |
|  | Cuplikan frame |
|  | Koefisien mel |
|  | Windowing |
|  | Mel filter |
| *Fs* | Frekuensi pencuplikan |
|  | Matrik ekstrasi ciri |

DAFTAR SINGKATAN

|  |  |
| --- | --- |
| ASR | Automatic Speech Recognition |
| NLP | Natural Languange Processing |
| VUI | Voice User Interface |
| LPCC | Linear Predictive Cepstral Coefficients |
| MFCC | Mel-Frequency Cepstrum Coefficient |
| JST | Jaringan Saraf Tiruan |
| API | Application Programming Interface |
| FCRN | (FullyConnected Recurrent Neural Network |
| BPTT | Backpropagation through Time-BPTT |
| BPNN | Back-propagation Neural Network |
| SR | Speech Recognition |
| TTS | Text to Speech |
| DCT | Discrets Cosine Transform |
| ART | Adaptive Resonance Theory |



**ABSTRACT**

Voice is a unique identity which is owned and attached to human being. It is a part of our body. Voice acknowledgment system which utilizes human voice will be extracted by MFCC, to produce a matrix that will be stored in database. Voice acknowledgment process will be performed by matching the tested voice and matrix in database.

Prototype model will be used in voice acknowledgment system through analyzes, design, and implementation. Voice acknowledgment system needs data of human voice that is acquired with microphone and recorded by using Free Sound Recorder with wav file format and stored in a certain named folder. Data of suara latih needs 10 voices from each person. Voice data that has been read and extracted with MFCC characteristics will produce 10x503 matrix. Since the data is too large, it will be simplified through PCA that produces 10x4 matrix, and in the next phase it will be reshaped to produce 1x40 matrix. As a result there will be 300x40 matrix as a whole, because data of suara latih comes from 300 voices of 30 persons. Simplified matrix will become the input of imitation nerve tissue and it will be organized as Max, Epoch, Goal, Learning Rate, PCA, NumOf Hidden Neuron, activated function from input to hidden layer and activated function from hidden layer to output. Having trained the tissue, it will be stored in database. Testing process will be performed by matching the testing data with thw sample data that has been trained with imitation nerve tissue.

Voice acknowledgment system with imitation nerve tissue has accuracy rate as 81.67% .Further improvement of voice acknowledgment system will be performed by enhancing accuracy rate in voice acknowledment, as a result we need to add acquisition instrument which has high resolution. In addition, we need data of suara latih in various environment, so there will be got a high accuracy rate in each testing.

Key words: pattern speaker recognition, MFCC, artificial neural network



merupakan proses pelatihan yang bersifat mengatur dirinya sendiri dengan algoritma perambatan balik. Dengan memanfaatkan teknologi informasi, suara bisa dijadikan sebagai autentikasi seseorang dalam identifikasi suatu permasalahan.

Pengenalan suara mengusulkan hubungan penuh pada lapisan tersembunyi antara input dan output. Selain itu, juga menyelidiki dan menunjukkan bahwa lapisan tersembunyi membuat pembelajaran yang lebih efisien untuk mengklasifikasikan secara kompleks dan menyelidiki perbedaan antara LPCC dan MFCC dalam proses ekstraksi fitur. Jaringan Saraf terkenal sebagai teknik yang memiliki kemampuan untuk memecahkan masalah *non-linear* (Manan dkk., 2004)*.*

Jaringan saraf tiruan digunakan untuk mengenali suara yang terisolasi. Tahapan mengenali suara terisolasi terdiri dari dua langkah, pra-pengolahan dengan *Digital Signal Processing* (DSP) teknik dan bagian pos-pengolahan dengan Saraf Tiruan Jaringan (JST). Kedua bagian ini secara singkat dijelaskan dan *recognizers* pidato menggunakan berbeda arsitektur JST yang diimplementasikan pada Matlab. Tiga desain model jaringan saraf yang berbeda yakni *multi layer back propagation*, *Elman* dan *probabilistic neural networks.* Perbandingan kinerja dengan studi serupa yang ditemukan dalam literatur terkait yang ditunjukkan yang dengan mengusulkan struktur JST dengan hasil yang memuaskan (Dede dan Sazlı ,2009).

Penggunaan *Speech API* dalam proses ekstraksi dan penyeleksian pengucap sangat dibutuhkan untuk pengenalan, dan diintegrasikan dengan sistem pengenalan pengucap yang nantinya akan dibuat sistem pakar untuk mengetahui tingkat keakuratan dalam mengenali suara masing-masing orang untuk presensi.

Berdasarkan latar belakang bahwa suara manusia mempunyai frekuensi suara yang berbeda-beda sehingga bersifat unik oleh karena itu pada penelitian ini dibuat suatu sistem presensi suara dengan jaringan saraf tiruan perambatan balik.

* 1. **Tujuan Penelitian**

Adapun tujuan yang hendak dicapai antara lain :

* 1. Merancang sistem pengenalan suara manusia yang dikonversikan kedalam bobot numerik dengan ektraksi ciri MFCC
  2. Melakukan proses pencocokan dan pencarian di basis data menggunakan jaringan saraf tiruanperambatan balik.

* 1. **Manfaat Penelitian**

Diharapkan dari penelitian ini dapat memberikan manfaat antara lain :

1. Dapat mengetahui tingkat keakuratan pengenalan pengucapan.
2. Dapat mengetahui tingkat presensi.
3. Dapat mengetahui karakteristik suara yang berbeda dari setiap orang

**BAB II**

**TINJAUAN PUSTAKA DAN DASAR TEORI**

* 1. **Tinjauan Pustaka**

Jaringan saraf tiruan (JST) tersusun atas elemen-elemen sederhana yang beroperasi secara paralel. Elemen-elemen ini diilhami oleh sistem saraf biologis sehingga dapat menirukan kerja otak manusia. JST yang telah dilatih dapat melakukan fungsi-fungsi kompleks di berbagai bidang, antara lain identifikasi, pengenalan pola, klasifikasi, restorasi citra, termasuk juga untuk sistem kendali (Demuth dan Beale, 1994).

Sistem pengenalan mengusulkan karakter invarian jaringan saraf tiruan perambatan balikganda. Model ini terdiri dari dua bagian, yang pertama adalah preprosesor yang dimaksudkan untuk menghasilkan terjemahan, rotasi dan skala pola input representasi invarian. Kedua adalah klasifikasi jaringan saraf tiruan yang dilatih oleh sebuah pembelajaran algoritma perambatan balik ganda. Setelah dilakukan pengetesan meliputi diputar skalanya dan diterjemahkan polanya. Pengenalan sederhana dengan klasifikasi perambatan balik ganda berhasil hampir 97% dari pola tersebut (Kamruzzaman dan Aziz, 1998).

Pada penelitian sebelumnya, jaringan saraf tiruan digunakan untuk menyelesaikan pengenalan suara yang terisolasi. Topik ini diselidiki dengan dua langkah, yang terdiri dari bagian pra-pengolahan dengan teknik pengolahan sinyal digital dan klasifikasi dengan jaringan saraf tiruan. Kedua bagian ini dijelaskan secara singkat dan pengenalan ucapan menggunakan berbeda arsitektur jaringan saraf tiruan diimplementasikan pada Matlab. Tiga desain model jaringan saraf yang berbeda ; *multi layer back propagation, Elman and probabilistic neural networks* (Gaulin dan Sazli, 2010).

Studi tentang pengenalan suara sudah dilakukan bertahun-tahun dengan mengusulkan *fullyconnected* di lapisan hidden antara node input dan output. Selain itu, dengan penelitian tentang lapisan tersembunyi membuat pembelajran klasifikas kompleks yang lebih efisien, serta menyelidiki perbedaan Prediksi Koefisien Linear Cepstrum (*Linear Prediction Cepstrum Coefficients-* LPCC) dan Koefisien Mel-Frekuensi cepstrum (*Mel-Frequency Cepstrum Coefficient-*MFCC) dalam proses fitur ekstraksi. Dalam hal ini sudah dilakukan penelitian perbedaan huruf arab dari “alif’ sampai “ya” menggunakan pembelajaran Sepenuhnya Terhubung Jaringan Saraf Berulang (*FullyConnected Recurrent Neural Network* -FCRN) dan Peramabatan Balik Melalui Waktu (*Backpropagation through Time*-BPTT) (Ahmad dkk., 2004).

Pengenalan alfabet lisan sebagai salah satu dari *speech recognition* dan pengenalan pola yang memiliki banyak aplikasi. Pengenalan alfabet yang diucapkan bukan tugas sederhana karena kumpulan yang sangat rapi seperti alfabet bahasa inggris. Kesamaan yang sangat akustik yang berkontribusi terhadap *confustability* yang dapat menghambat keakuratan sistem pengenalan suara, salah satu kumpulan “E” *confustable* yaitu huruf B, C, D, E, G, P, T, V, Z. Pengenalan alfabet yang terisolasi menggunakan MFCC dan Jaringan Saraf Perambatan Balik (*Back-propagation Neural Network-*BPNN) untuk kumpulan “E” dan semua 26 huruf. Tingkat pembelajaran dan tingkat momentum BPNN yang diseuaikan dan bervariiasi untuk mencapai pengenalan terbaik dari semua huruf dengan tingkat keberhasilan mencapai 62,28% dan 70,49% tingkat pengenalan untuk kumpulan “E” dalam suara independen dan tergantung kondisi masing-masing (Adam dan Salam, 2012).

* 1. **Dasar Teori**
     1. **Pengenalan Pola**

Pengenalan pola (*pattern recognition*) dapat diartikan sebagai proses klasifikasi dari objek atau pola menjadi beberapa kateori atau kelas, dan bertujuan untuk pengambilan keputusan (Theodori dan Koutroumbas, 2006).

Pola adalah bentuk atau model (abstrak, suatu set peraturan) yang dapat dipakai untuk membuat atau mengahasilan suatu atau bagian dari sesuatu, khususnya jika sesuatu yang di timbulkan mempunyai sejenis pola dasar yang dapat ditunjukan atau terlihat, yang mana dapat dikatan mempertunjukan pola. Deteksi pola dasar disebut pengenalan pola. Untuk mengenali suara dan pemilik suara, diperlukan suatu pengenalan pola yang dapat mengenal dan mencocokan pola yang masukan dan pola yang telash disimpan di dalam basisdata,

Pendekatan pengenalan pola ada tiga, yaitu secara sintaks, statistik, serta melalui jaringan saraf tiruan. Pendekatan dengan pola jaringan saraf tiruan adalah pendekatan dengan menggabungkan pendekatan sintaks dan statistik. Pendekatan melalui pola-pola ini meniru cara kerja otak manusia, pada pola ini sistem membuat aturan-aturantertentu disertai dengan menggunakan data statistik sebagai dasar untuk pengambilan keputusan.

* + 1. **Autentikasi**

Autentifikasi tindakan untuk menentukan atau mengkonfirmasi sesuatu atau seseorang sebagai autentik, yaitu klaim oleh atau mengenai objek atau individu tersebut adalah benar. Autentifikasi terhadap seseorang pada umumnya mencakup verifikasi identitas orang tersebut (Kadim,2005).

Dalam keamanan komputer, autentikasi adalah proses memverifikasi identitas dari satu pihak yang auntetifikasi dapat seseorang yang menggunakan komputer, komputer itu sendiri, atau sebagai program aplikasi.

Metode-metode autentifikasi yang digunakan manusia dibagi menjadi tiga:

1. Pengguna itu sendiri/biometrik (sidik jari, retina, DNA, iris, suara)
2. Sesuatu yang dimiliki (kartu pengenal)
3. Sesuatu yang diingat (*passwaord,* PIN)

Suatu sistem pengenalan pola pada dasarnya terdiri atas tiga tahap, yaitu penerimaan data, pengenalan data dan pengenalan objek atau pembuatan keputusan. Adapun pengenalan utama dalam pengenalan pola adalah pendekatan geometrik dan pendekatan struktural. Pendekatan struktural dilakukan dengan penentuan dasar yang mendeskripsikan objek yang akan dikenali.

Pengenalan pola dengan pendekatan Jaringan Saraf Tiruan, seolah-olah membuat sebuah sistem yang kinerjanya sama dengan otak manusia. Agar sistem tersebut bisa menjadi cerdas, kita harus memberikan pelatihan terhadap sistem tersebut selama rentang waktu yang kita tentukan. Karena dengan melatih sistem tersebut maka akan menambah *rule*-*rule* serta data statistik yang di gunakan oleh sistem untuk mengambil keputusan (Kusumadi, 2004).

* + 1. **Sinyal Percakapan**

Sinyal percakapan adalah sinyal yang dihasilkan dari suara manusia sewaktu melakukan percakapan. Sinyal percakapan merupakan kombinasi kompleks dari variasi tekanan udara yang melewati pita suara dan vokal *tract*, yaitu mulut, lidah, gigi, bibir, dan langit-langit mulut. *Speech* (wicara) dihasilkan dari sebuah kerjasama antara *lungs* (paru-paru), *glottis* (dengan vokal cords) dan *articulation tract* (mulut dan rongga hidung). Sinyal suara terdiri dari serangkaian suara yang masing–masing menyimpan sepotong informasi. Berdasarkan cara menghasilkannya, suara dapat dibagi menjadi *voiced* dan *unvoiced*. *Voiced sounds* atau suara ucapan dihasilkan dari getaran pita suara, sedangkan *unvoiced sounds* dihasilkan dari gesekan antara udara dengan *vocal tract*. Sinyal percakapan memiliki beberapa karakteristik, seperti *pitch* dan intensitas suara yang berguna dalam melakukan analisis sinyal suara. *Pitch* adalah frekuensi dari sinyal atau yang sering disebut intonasi. Intensitas suara adalah tingkat kekuatan suara.

Impuls tekanan pada umumnya disebut sebagai pitch impulses dan frekuensi sinyal tekanan adalah *pitch frequency* atau *fundamental frequency*. Sederet impuls (fungsi tekanan suara) dihasilkan oleh pita suara untuk sebuah suara. Hal ini merupakan bagian dari sinyal suara yang mendefinisikan *speech melody* (melodi wicara). Ketika berbicara dengan pitch yang stabil, suara sinyal wicara monotonous tetapi dalam kasus normal sebuah perubahan permanen pada frekuensi terjadi. Impuls pitch merangsang udara dalam mulut dan untuk suara tertentu (nasals) juga merangsang *nasal cavity* (rongga hidung). Ketika rongga beresonasi, timbul radiasi sebuah gelombang suara yang merupakan sinyal percakapan. Kedua rongga beraksi sebagai resonators dengan karakteristik frekuensi resonansi masing–masing yang disebut *formant frequencies*, sehingga *formant* merupakan variasi resonasi yang dihasilkan oleh *vocal tract*. Pada saat rongga mulut mengalami perubahan besar, dihasilkan beragam pola ucapan suara yang berbeda. Di dalam kasus *unvoiced sounds*, keluaran pada vocal tract lebih menyerupai noise atau derau (Rabiner dkk, 1989).

* + 1. **Teknologi SAPI**

Diperkenalkan oleh Microsoft pada tahun 1995, *Speech Application Programming Interface* (SAPI) memungkinkan sistem untuk mengenali suara manusia sebagai masukan dan keluaran berupa suara manusia dari text. Fasilitas ini menyediakan hubungan interaksi baru antara manusia dengan komputer, SAPI merupakan bagian dari *Windows Open Services Architecture* model. Pada *Speech Recognition* (SR) dan *Text to Speech* (TTS) terdapat beberapa modules yang disebut engines, yang telah tersedia pada semua sistem operasi Windows XP. SAPI terdiri dari 2 antar muka yaitu *Application Programming Interface* (API) dan *Device Driver Interface* (DDI). Gambar 2.1 merupakan arsitektur SAPI.



Gambar 2.1 Arsitektur SAPI

SAPI (*Speech Application Program Intervace*) 5.1 adalah salah satu API (*Application Programming Interface*) yang di sediakan sebagai pengantar untuk pemrograman dengan menggunakan metode *speech recognition* dan *Text To Speech.*

Dengan adanya SAPI memungkinkan pembuat untuk mengimplementasikan sistem pengenal suara dengan menggunakan *engine* yang berbeda tanpa merubah aplikasi yang telah dibuat (Adipranata, 2003).

Terdapat dua macam mode pada pengenalan suara atau wicara, yaitu (Adipranata, 2003) :

1. **Mode Diktasi**

Pada metode ini pengguna komputer dapat mengucapkan kata atau kalimat, yang selanjutnya akan dikenali oleh komputer dan diubah menjadi data text. Kemungkinan jumlah kata yang dikenali dibatasi oleh jumlah kata yang telah terdapat pada basis data. Pengenalan Mode Diktasi merupakan sistem bergantung pengucap. Keakuratan pengenalan mode ini bergantung pada pola suara dan aksen pembicara serta pelatihan yang telah dilakukan.

1. ***Mode Command and Control***

Pada metode ini, pengguna komputer mengucapkan kata atau kalimat yang sudah didifinisikan terlebih dahulu pada basis data dan selanjutnya akan digunakan untuk menjalankan perintah tertentu pada aplikasi komputer. Jumlah perintah yang dapat jenis-jenis perintah yang dapat dieksekusikan. Mode ini merupakan *Speaker Independent* karena jumlah kata yang dikenali biasanya terbatas sekali dan ada kemungkinan pembicara tidak perlu melakukan pelatihan pada sistem sebelumnya. Terdapat empat proses utama pada sistem pengenalan suara, baik pada mode Diktasi ataupun *Command and Control*, yaitu (Adipranata, 2003) :

* + 1. Pemisahan kata

Pemisahan kata adalah proses untuk memisahakan suara yang diucapkan oleh pengguna menjadi beberapa bagian. Masing-masing bagian bisa berupa kalimat ataupun hanya sebuah kata. Terdapat tiga macam metode yang dapat digunakan pada proses pemisahan kata, yaitu (Adipranata, 2003) :

* + - 1. Pengucap Diskret

Pada metode ini, pengguna diharuskan mengucapkan kalimat secara terpenggal dengan adanya jeda sejenak diantara kata. Jeda tersebut digunakan oleh sistem untuk mendeteksi awal dan akhir sebuah kata. *Discrete Speech* mempunyai kelebihan yaitu sedikit *resource* (memori komputer waktu proses) yang digunakan oleh sistem untuk mendeteksi suara, tetapi mempunyai kelemahan yaitu ketidaknyamanan pengguna dalam mengucapkan sebuah kalimat.

* + - 1. Pengejaan Kata

Dalam sebuah kalimat yang diucapkan pengguna, sistem hanya mendeteksi kata yang terdapat didalam perbendaharaan yang dimilikinya, dan mengabaikan kata-kata lain yang tidak dimilikinya. Sehingga walau pengguna mengucapkan kalimat yang berbeda tetapi didalam kalimat tersebut terdapat sebuah kata yang sama dan terdapat diperbendaharaan sistem, maka hasil pengenalan akan sama. Kelemahan metode ini adalah besar kemungkinan sistem akan melakukan kesalahan arti pengenalan dalam bentuk kalimat. Tetapi metode ini mempunyai kelebihan yaitu pengguna dapat mengucapkan kalimat secara normal tanpa harus berhenti diantara kata.

* + - 1. Ucapan Menerus

Sistem akan mengenali dan memproses setiap kata yang diucapkan. Metode ini akan menghasilkan keakuratan dalam mengenali ucapan pengguna. Tetapi disamping itu, metode ini memerlukan resource (memori komputer waktu proses) yang besar dalam prosesnya. Pada metode ini sistem harus bisa mendeteksi awal dan akhir setiap kata dalam kalimat tanpa adanya jeda diantara kata-kata tersebut, dan setelah berhasil memisahkan kata, langkah selanjutnya adalah mencocokan dengan perbendaharaan kata yang dipunyai.

* + 1. Ketergantungan terhadap pengguna

Sistem pengenalan suara mempunyai beberapa sifat, dilihat dari ketergantungan terhadap pengguna, yaitu (Adipranata, 2003) :

* + - 1. *Speaker Dependent*

Sistem membutuhkan pelatihan untuk setiap pengguna yang akan menggunakan sistem tersebut. Sistem tidak akan bisa mengenali pengguna yang belum pernah melakukan pelatihan.

* + - 1. Pengucap bebas

Pengguna tidak perlu melakukan pelatihan sebelum dapat menggunakan sistem, karena sistem mampu mengenali suara pengguna tidak tergantung warna suara dan dialek yang digunakan.

* + - 1. *Speaker Adaptive*

*Speaker Adaptive* merupakan perpaduan dari *Speaker Dependent* dan *Speaker Independent*, dimana pengguna tidak perlu melakukan pelatihan dan keakuratan pengenalan sistem akan makin meningkat jika pengguna yang sama bekerja terus menerus beberapa waktu tertentu.

* + 1. Pencocokan kata

Pencocokan kata adalah proses untuk mencocokan kata ucapan yang berhasil diidentifikasikan dengan basis data yang dipunyai oleh sistem. Terdapat dua metode yang dapat dipakai pada proses pencocokan kata ini, yaitu (Adipranata, 2003) :,

* + - 1. *Whole-Word Matching* (Pencocokan Seluruh Kata)

Sistem akan mencari dibasis data kata yang sama persis dengan kata hasil ucapan pengguna.

* + - 1. *Phoneme Matching* (Pencocokan Fonem)

Sistem memiliki kamus fonem. Fonem ialah bagian terkecil dan unik dari suara yang

berbentuk sebuah kata.

* + 1. Perbendaharaan kata

Perbendaharaan kata adalah bagian terakhir dalam sebuah sistem pengenalan suara. Terdapat dua hal yang perlu diperhatikan pada perbendaharaan kata, yaitu ukuran dan keakuratan. Jika perbendaharaan kata berjumlah banyak maka sebuah sistem akan mudah dalam melakukan pencocokan kata, tetapi dengan meningkatnya jumlah perbendaharaan kata, maka jumlah kata yang mempunyai ucapan hampir sama juga mengingkat, dimana hal ini menurunkan keakuratan pengenalan. Dan begitu juga sebaliknya, jika sebuah sistem mempunyai perbendaharaan kata yang sedikit, maka keakuratan akan tinggi karena edikitnya kata yang hampir sama, tetapi akan semakin banyak kata yang tidak dikenali (Adipranata, 2003).

Untuk sistem pengenalan pembicaraan dengan *mode Command and Control*, akan lebih baik jika menggunakan jumlah perbendaharaan kata sedikit (kurang dari 100 kata), tetapi untuk mode Diktansi akan lebih membutuhkan jumlah perbendaharaan kata yang banyak.

Pada penelitian ini, aplikasi yang dikembangkan adalah aplikasi berbasis teks untuk menerima diktansi bahasa Indonesia. Sehubungan dengan aplikasi tersebut, maka mode yang digunakan adalah perintah dan pengaturan dengan proses mengucap penerus *dan speaker adaptive* pada penggunanya.

SAPI (*Speech Application Program Interface*) 5.1 secara garis besar dapat

dibagi menjadi dua struktur utama yang berfungsi untuk proses ucapan ke teks ataupunTeks ke Ucapan yaitu (Adipranata, 2003) :

Pada aplikasi pengenalan suara pada Microsoft Word, API berfungsi menghubungkan antara aplikasi Microsoft Word dengan SAPI Runtime.

1. *Device Drive Interface* (DDI)

DDI menyediakan fungsi untuk menerima data suara dari SAPI Runtime dengan mengembalikan pengenalan frase pada level SAPI paling dasar. Terdapat dua engine yang digunakan oleh DDI yaitu mesin pengenal yang berfungsi sebagai pengubah suara menjadi teks atau tulisan dan TTS *engine* (*Text to Speech Engine*). Dalam perancangan aplikasi ini hanya menggunakan mesin pengenal saja, karena TTS hanya berfungsi untuk penerapan aplikasi yang berbasis teks ke ucapan.

1. *Device Drive Interface* (DDI)

DDI menyediakan fungsi untuk menerima data suara dari SAPI Runtime dengan mengembalikan pengenalan frase pada level SAPI paling dasar. Terdapat dua engine yang digunakan oleh DDI yaitu mesin pengenal yang berfungsi sebagai pengubah suara menjadi text atau tulisan dan TTS engine (*Text to Speech Engine*). Dalam perancangan aplikasi ini hanya menggunakan mesin pengenal saja, karena TTS hanya berfungsi untuk penerapan aplikasi yang berbasis teks ke ucapan.

* + 1. **Ektraksi Ciri MFCC**

Cir-ciri khusus yang akan diekstrak dari masukan sinyal suara pembicara yang nantinya akan dikenali. Ekstraksi merupakan representasi parametric terbaik dari sinyal akustik untuk menghasilkan kinerja pengenalan yang lebih baik. *Mel Frequency cepstral coefficients* (MFCC) adalah salah satu ekstraksi fitur yang paling berhasil dalam pengenalan suara, dan koefisien yang diperoleh melalui *filter bank analysis.* Langkah-langkah yang dilakukan dalam ekstraksi adalah *pre-emphais, frame blocking, windowing, filter bank analysis, logarithmic compression* dan *discrete cosine transformation.* Secara keseluruhan proses dapat dilihat pada Gambar 2.2 (Rojathai dkk, 2012).



Gambar 2.2 Porese ekstrasi ciri *MFCC*

1. ***Pre-Emphasis***

Prapemrosesan sinyal suara dari kedua basisdata yang diberikan kepada tahap ekstraksi ciri MFCC prapenakan. Prapenakan adalah proses yang dirancang untuk meningkatkan beberapa besarnya frekuensi sehubungan dengan besarnya frekuensi lainya.

Disini sinyal ucapan diproses diteruskan melalui filter untuk menekankan frekuensi yang lebih tinggi. Proses ini akan meningktakan energy sinyal pembicara pada frekuensi yang lebih tinggi. Sinyal suara pertama kali di prapenekanan oleh *filter FIR* pada .urutan pertama dengan koefisien δ pra-penekanan. Urutan filter pertama FIR dengan fungsi transfer dalam domain z adalah

(2.1)

Keterangan :

Koefisien prapenekanan δ terletak antara

(2.2)

(2.3)

Keterangan :

i

j

1. ***Frame Blocking***

Ciri statistik dari sinyal suara yang invariant hanya dalam interval waktuyang singkat. Disini sinyal prapenakanan diblokir menjadi *sampling frame* , dengan bingkai yang berdekatan menjadi dipisahkan oleh *sample FM (Frame Shift)*. Jika kerangka pengucap adalah dan dalam sinyal suara kesuluruhan, maka

(4) (2.4)

(2.5)

Keterangan :

1. ***Windowing***

Proses selanjutnya adalah *windowing,*  dimana setiap *frame windowed* untuk mengurangi sinyal ketidaksinambungan pada awal dan akhir frame. *Windowing* yang dipilih untuk rekaman sinya di tepi setiap frame. Jika *windowing*nya didefinisikan sebagai berikut : (2.6)

(2.7)

Maka hasilnya,

(2.8)

(2.9)

*Hamming Window* adalah pilihan terbaik dalam pengenalan suara, yang mengintegrasikan semua frekuensi garis terdekat. Persamaan *Hamming Window* sebagai berikut :

(2.10)

(2.11)

1. ***Filter Bank Analysis***

*Filter Bank Analysisi* dilakukan untuk mengkonversi setiap contoh kerangka waktu dari ke kewasan frekuensi. Transformasi Fourierditerapkan untuk mengkonversi konvolusi dari sinyal *glottal* dan respon saluran sinyal vocal dalam domain waktu ke domain frekuensi. Rentang frekuensi di spectrum FFT sangat luas dan sinyal suara tidak mengikuti skala linear. Satu himpunan tapis segitiga yang digunakan untuk menghitung jumlah tertimbang komponen spectral tapisr sehingga keluaran dari proses mendekati ke skala Mel. Besarnya respon frekuensi setiap tapis adalah berbentuk segitiga dan sama dengan keutuhan dipusat frekuensi dan mengurangi linear ke nol pada dua frekuensi tengah yang dekat tapis. Selanjutnya, masing-masing keluaran tapis adalah jumlah dari komponen yang ditapis spectral. Skala *mel* didefinisikan sebagai berikut

(2.12)

Keterangan :

= Skala Mel

Filter yang sama disebut sebagai bank skala *mel filter* dan respon frekuensi dari filter bank mensimulasikan proses persepsi yang dilakukan didalam telinga.

1. ***Logarithmic Compression***

Fungsi algoritma mengkompresi keluaran filter yang diperolah dari analisis *filter bank.* Keluaran kompresi filter sebagai berikut :

(2.13)

(2.14)

Keterangan :

= Analisis *filter bank*

1. ***Discrete Cosine Trnasformation***

*Discrets Cosine Transform* (DCT) yang diterapkan pada filter keluaran dan koefisien pertama dikelompokan bersama sebagai *feature vekctor* dari kerangka pembicara yang spesifik. Koefisien MFCC antara 1 dapat diberikan sebagai berilkut :

(2.15)

(2.16)

Dimana C adalah urutan skala *mel cepstrum.*

Keterangan :

= Mel Frekuensi

K = Koefisien

* + 1. **PCA (*Principal Components Analysis*)**

*Principal Components Analysis* (PCA) merupakan teknik statistik multivariate, dalam rangka eksplorasi untuk menyederhanakan data set yang kompleks (Basilevsky 1994, Everitt & Dunn 1992, Pearson 1901). Mengingat banyaknya pengamatan (sebesar p) pada n variabel, oleh karena itu tujuan PCA adalah untuk mengurangi dimensi dari matriks data dengan mencari variabel r baru, di mana r kurang dari n. Disebut komponen utama, variabel-variabel r baru yang terbentuk sebanyak varians dalam n variabel asli, tetap saling berkorelasi dan ortogonal. Setiap komponen utama adalah kombinasi linear dari variabel asli, dan sehingga dapat menangkap makna komponen yang mewakili-nya.

Untuk menghitung komponen utama, nilai eigen dan vektor eigen dari n variabel yang berhubungan dihitung dari matriks kovarians dengan kondisi n x n. Setiap *eigenvector* mendefinisikan komponen utama. Sebuah komponen dapat dilihat sebagai jumlah tertimbang dari kondisi, dimana koefisien dari vektor eigen merupakan bobot. Proyeksi gen i sepanjang sumbu didefinisikan oleh komponen jth  pokok:

(2.17)

Keterangan :

n = variabel

* + 1. **Jaringan Saraf Tiruan**

Jaringan Saraf Tiruan (JST) terinspirasi dari penelitian kecerdasan buatan, terutama percobaan untuk menirukan toleransi kesalahandan kemampuan untuk belajar dari sistem saraf biologi dengan model struktur level rendahotak.

Otak manusia memiliki sekitar 10.000.000.000 sel saraf yang saling berhubungan. Sel saraf mempunyai cabang struktur *input* (dendrites), yaitu sebuah inti sel dan percabangan struktur *output* (axon). Axon dari sebuah sel terhubung dengan dendrites yang lain melalui sebuah *synapse*. Ketika sebuah sel saraf aktif, kemudian menimbulkan suatu signal *electrochemical* pada axon. Signal ini melewati *synapses* menuju ke sel saraf yang lain. Sebuah sel saraf yang lain akan mendapatkan signal jika memenuhi biasan tertentu yang sering disebut dengan nilai ambang (*threshold*) (Kusumadewi, 2003).

Bila dibandingkan dengan komputer modern, kecepatan dari *neuron* pada otak manusia jauh lebih lambat, namun karena jumlah *neuron* yang sangat banyak pada otak manusia membuat otak manusia dapat melakukan banyak pekerjaan yang sangat cepat serta menggunakan energi atau tenaga yang sangat sedikit dan efisien. Menurut Faggin (1991) sebagai contoh, pengenalan obyek seperti wajah seseorang dapat dilakukan hanyadalam waktu sepersepuluh detik, sementara dengan *basis data* yang sama dengan komputer modern saat ini dibutuhkan waktu yang jauh lebih lama setidaknya beberapa menit lebih lama dari otak manusia.

* + - 1. **Pengertian Jaringan Syaraf Tiruan**

Jaringan Saraf Tiruan merupakan suatu metode *Artificial Intelligence* yang konsepnya meniru sistem jaringan saraf yang ada pada tubuh manusia, dimana dibangun node-node yang saling berhubungan satu dengan yang lainnya. Node-node tersebut terhubung melalui suatu link yang biasa disebut dengan istilah weight. Ide dasarnya adalah mengadopsi cara kerja otak manusia yang memiliki ciri-ciri *paralel processing*, *processing element* dalam jumlah besar dan toleransi kesalahan.

Suatu jaringan saraf tiruan memproses sejumlah besar informasi secara parallel dan terdistribusi, hal ini terinspirasi oleh model kerja otak biologis. Beberapa definisi tentang jaringan saraf tiruan adalah sebagai berikut di bawah ini.

1. Definisi sistem saraf buatan sebagai berikut: "Suatu Jaringan Saraf Tiruan adalah suatu struktur pemroses informasi yang terdistribusi dan bekerja secara paralel, yang terdiri atas elemen pemroses (yang memiliki memori lokal dan beroperasi dengan informasi lokal) yang diinterkoneksi bersama dengan alur sinyal searah yang disebut koneksi. Setiap elemen pemroses memiliki koneksi keluaran tunggal yang bercabang ke sejumlah koneksi kolateral yang diinginkan (setiap koneksi membawa sinyal yang sama dari keluaran elemen pemroses tersebut). Keluaran dari elemen pemroses tersebut dapat merupakan sebarang jenis persamaan matematis yang diinginkan. Seluruh proses yang berlangsung pada setiap elemen pemroses harus benar-benar dilakukan secara lokal, yaitu keluaran hanya bergantung pada nilai masukan pada saat itu yang diperoleh melalui koneksi dan nilai yang tersimpan dalam memori lokal" (Hecht-Nielsend,1988).
2. Jaringan Saraf Tiruan: A Comprehensive Foundation, NY, Macmillan, mendefinisikan jaringan saraf sebagai berikut: “Sebuah jaringan saraf adalah sebuah prosesor yang terdistribusi paralel dan mempunyai kecenderungan untuk menyimpan pengetahuan yang didapatkannya dari pengalaman dan membuatnya tetap tersedia untuk digunakan. Hal ini menyerupai kerja otak dalam dua hal yaitu: 1. Pengetahuan diperoleh oleh jaringan melalui suatu proses belajar. 2. Kekuatan hubungan antar sel saraf yang dikenal dengan bobot sinapsis digunakan untuk menyimpan pengetahuan” (Haykin, S,1994).
3. Introduction To Artificial Neural Sistems, Boston: PWS Publishing Company, mendefinisikan sebagai berikut: “Sistem saraf tiruan atau jaringan saraf tiruan adalah sistem selular fisik yang dapat memperoleh, menyimpan dan menggunakan pengetahuan yang didapatkan dari pengalaman” (Zurada, J.M,1992).

Jaringan saraf tiruan menyerupai otak manusia dengan dua cara :

1. Pengetahuan yang diperoleh jaringan dari lingkungannya melalui proses pembelajaran.
2. Kekuatan hubungan antar neuron, dikenal dengan istilah bobot sinaptik, dan digunakan untuk menyimpan pengetahuan yang diperoleh.

Jaringan Saraf Tiruan sendiri pun dibagi-bagi kembali menjadi beberapa bagian yang lebih kecil, dimana masing-masing metode mempunyai karakteristik sendiri-sendiri, serta memiliki keunggulan dan kelemahan dalam mengenali suatu pola.

Metode-metode tersebut diantaranya adalah *Bidirectional associate memory* atau lebih dikenal dengan istilah *BAM,* Jaringan Hopfield *, Counter Propagation Network, Backpropagation* dan masih banyak metode-metode lainnya yang sudah atau sedang dikembangkan dengan para ahli.

Pada umumnya Jaringan Saraf Tiruan dibagi berdasarkan layer-layer yaitu *input layer, hidden layer, output layer*. Setiap node pada masing-masing layer memiliki suatu error rate, yang akan digunakan untuk proses training.

Pada kenyataannya (kebiasaannya), kebanyakan neural sistem harus diajari (training) terlebih dahulu. Mereka akan mempelajari asosiasi, pattern, dan fungsi yang baru. Pemakai-pemakai Jaringan Saraf Tiruan tidak menspesifikasikan sebuah algoritma untuk dieksekusi dalam setiap perhitungan. Mereka akan memilih arsitektur tertentu dengan pandangan mereka, dengan karakteristik neuron*,* bobot, dan memilih model training sendiri. Sehingga dari hasil tersebut, informasi network dapat diubah oleh para pemakai. Sistem jaringan buatan juga dapat mengkalkulasi teknik matematik, seperti meminimalisasi kesalahan suatu perhitungan.

Jaringan Saraf Tiruan sangat berperan dalam teknologi dan beberapa disiplin ilmu, yang membantu dalam menentukan model-model Jaringan Saraf Tiruan dan *sistem non linear dynamic*. Salah satunya adalah matematika. Matematika adalah model neural yang paling berpotensi karena kekompleksannya. Elektronika dan ilmu komputer juga menggunakan metode ini, karena berperan dalam pengiriman sinyal data.

* + - 1. **Sejarah Perkembangan Jaringan Saraf Tiruan**

Sejarah Perkembangan Jaringan Saraf Tiruan dimulai pada tahun 1949, Donald Hebb melakukan algoritma *Learning* atau *training* yang mengatakan bahwa jika dua *neuron* aktif secara simultan, maka kekuatan hubungan antar *neuron* akan meningkat. Kemudian pada tahun 1950an dan 1960an, sebuah perkumpulan yang bernama Group of Researchers berhasil menemukan konsep *perceptrons.* Pada tahun berikutnya yaitu tahun 1969, Marvin Minsky secara matematis menunjukkan kelemahan *perceptrons*. Penelitian Jaringan Saraf Tiruan pada tahun 1970an. Tahun 1987, Group of Researchers menggelar *International Conference on* Jaringan Saraf Tiruan, San Diego, AS (Puspitaningrum, 2006).

* + - 1. **Sistem Neural**

1. Node

*Neuron* adalah unit pemprosesan informasi yang merupakan dasar dari operasi jaringan saraf tiruan. Sel-sel saraf tiruan ini dirancang berdasarkan sifat-sifat dari neuron biologis. Sel saraf tiruan ini biasa disebut sebagai *processing elements,* unit atau *node*.

Nodeadalah sebuah sel neuronyang disetiap *node-*nya memiliki kelauaran, *error*, dan *weight*. Jadi di setiap *node*, dimanapun itu pasti memiliki ketiga unsur tersebut. Hubungan antar *node* diasosiasikan dengan suatu nilai yang disebut dengan bobot atau *weight*. Setiap *node* pasti memiliki keluaran*, error* dan bobotnya masing-masing.

Keluaranmerupakan keluaran dari suatu *node*. *Error* merupakan tingkat kesalahan yang terdapat dalam suatu *node* dari proses yang dilakukan. *Weight* merupakan bobot dari *node* tersebut ke *node* yang lain pada lapisan yang berbeda. Nilai bobotberkisar antara -1 dan 1.

Bobot-bobot atau *weight* yang tersimpan di dalam jaringan saraf tiruan ini disebut sebagai bobot interkoneksi. Nilai bobot yang baik akan memberikan keluaran yang sesuai, dalam arti mendekati keluaran yang diharapkan target keluaran untuk suatu masukanyang diberikan.

Bobot awal dalam suatu jaringan saraf tiruan biasanya diperoleh secara acakdan sebaiknya diinisialisasi dengan nilai yang relative kecil, yaitu berkisar antara -0,1 sampai 0,1. Memasuki tahap pelatihan, bobot tersebut akan mengalami penyesuaian melalui suatu proses perhitungan matematik agar tercapai nilai bobot yang sesuai (Mitchell,1997).

1. Masukan, lapisan tersembunyi, dan lapisan keluaran.

Lapisan masukanmerupakan lapisantempat sebuah masukdimasukkan dan lapisanini dilakukan proses-proses selanjutnya. Menurut Michell (1997) lapisan tersembunyiberfungsi untuk membantu proses. Semakin banyak lapisan tersebunyiyang digunakan, makasemakin bagus dan semakin cepat pula keluaranyang diinginkan didapat. Akan tetapiwaktu pelatihanakan berlangsung semakin lama.Lapisan keluaranadalah lapisanyang menampung hasil proses dari suatu Jaringan Saraf Tiruan*. Forward propagation* bertujuan untuk menentukan keluarandari suatu *node.* Keluaranyang dimaksud disini adalah keluarandari *output layer.* Karena masing-masing *node* tersebut memiliki keluaran.

1. Pelatihan

Proses belajar suatu Jaringan Saraf Tiruanterdiri dari proses *Forward, Backward,* dan *Update Weight,* sekali melewati 3 tahap itu disebut dengan 1 kali pelatihan(1 cycle). Semakin banyak *training* yang dilakukan maka akan semakin kecil pula tingkat *error* yang dihasilkan di lapisan keluaranya. Dengan demikian semakin kecil juga *error* suatu sistem.

Dua metode Pembelajarandalam Jaringan Saraf Tiruan*,* yaitu (Rao,1995) :

* 1. *Supervised Learning* adalah suatu metode dimana *Jaringan Saraf Tiruan* belajar dari pasangan data *input* dan *target*, pasangan ini disebut *training pair*. Biasanya jaringan dilatih dengan sejumlah *training pair*, dimana suatu *input* vektor di aplikasikan, menghasilkan nilai di *output*, lalu hasil pada *output* tersebut akan dibandingkan dengan target *output*. Selisihnya akan dikembalikan ke jaringan, kemudian dihitung *error-*nya, melalui *error* ini akan didapatkan selisih yang terdapat di dalam *weight*. Oleh karena itu terdapat *weight* baru yang cenderung memiliki *error* yang lebihkecil, sehingga akan didapat *error* yang lebih minimum dari *error* yang pertama. Vektor-vektor dalam *training set* diaplikasikan seluruhnya secara berurutan. Pertama-tama *error* dihitung, kemudian *weight* disesuaikan sampai seluruh *training* *set* menghasilkan *error* yang sekecil-kecilnya. Pada dasarnya konsep ini berawal dari konsep otak manusia.

Model *Jaringan Saraf Tiruan* yang menggunakan metode *Supervised Learning* diantaranya adalah sebagai berikut :

* + 1. Model *Backpropagation*
    2. Model *Biderectional Associative Memory*
    3. *Hopfield Network*
  1. *Supervised (Unsupervised) Learning*

*Unsupervised Learning* dianggap sebagai model dalam konsep sistem biologis. Teori ini dikembangkan oleh Kohonen (1984) dan beberapa ilmuan lainnya. Dalam *Unsupervised Learning* tidak diperlukan target keluaran. Training hanya terdiri dari vektor-vektor masukan tanpa memiliki pasangan target. Algoritma pelatihanmengubah bobot jaringan untuk menghasilkan keluaranyang konsisten. Aplikasi dari vektorvektor yang cukup serupa menghasilkan pola keluaranyang sama. Dengan demikian proses pelatihanakan menghasilkan sifat-sifat statistik dalam bentuk pengelompokan vektor-vektor dalam beberapa kelas. Dengan mengaplikasikan suatu vektor dari suatu kelas sebagai masukan, maka akan menghasilkan vektor keluaranyang spesifik.

Model Jaringan Saraf Tiruanyang menggunakan metode *Non-Supervised (Unsupervised)* *Learning* diantaranya adalah sebagai berikut :

* 1. *Adaptive Resonance Theory* (ART)
  2. *Compotitive Learning*
  3. *Kohonen’s self organizing maps.*
     + 1. **Arsitektur Jaringan**

Neuron-neuron dalam jaringan saraf tiruan dikelompokan dakam lapisan-lapisan. Neuron yang mempunyai letah yang sama akan memiliki keadaan yang sama. Faktor terpenting dalam menentukan kelakuan suatu neuron adalah fungsi aktivasi dan pola bobotnya. Pada setiap lapisan yang sama, neuron-neruron akan memiliki fungsi aktivasi yang sama. Apabila neuron-neuron dalam suatu lapisan (misalnya lapisan tersembunyi) akan dihubungkan dengan neuron-neuron pada lapisan yang lain (misalkan lapisan output) (Kusumadewi,2003).

* + - 1. ***Single Layer* Jaringan Saraf Tiruan**

Lapisan tunggal jaringan syaraf tiruanadalah Jaringan Saraf Tiruanyang hanya memiliki satu lapisan dengan bobot terhubung. Jaringan ini hanya menerima masukankemudian secara langsung mengolahnya menjadi kelaurantanpa harus melalui lapisan tersembunyi. Neuron-neuron pada lapisan ini saling berhubungan. Semua unit masukanakan dihubungkan dengan setiap unit keluaran.

* + - 1. ***Multi Layer* Jaringan *Saraf Tiruan***

Lapisan GandaJaringan Saraf Tiruanadalah Jaringan Saraf Tiruanyang memiliki karakteristik lapisan gandadimana setiap nodepada suatu lapisan terhubung dengan setiap nodepada *lapisan* didepannya. Berasitektur umpan maju atau lebih dikenal dengan *feed forward* *network* yang menggunakan metode pemebelajaran terpandu (*supervised Learning*).

Model ini merupakan model yang paling sering dipakai dalam pengembangan sistem neural dan memiliki kinerja yang sangat baik dalam sisi keakuratan. Model ini mempunyai dua fase dalam pelatihannya yaitu fase *forward* dan *backward*. Cara kerja jaringan ini adalah setelah *input* masuk ke lapisan masukanmaka data akan diolah dan diteruskan ke masing-masing bagian di depannya sampai pada lapisan keluaran. Nilai di *Output layer* akan dibandingkan dengan nilai target, lalu akan dihasilkan sinyal *error* bagi masing *node* di lapisan keluaran*.* Kemudian sinyal ini ditransmisikan kembali atau lebih dikenal dengan *Backpropagation* yang berasal dari lapisan keluaran ke masing-masing sel pada lapisan sebelumnya seperti yang terlihat pada Gambar 2.3.



Gambar 2.3. Konsep Jaringan Saraf Tiruan Perambatan Balik

Arsitektur jaringan saraf tiruan seperti Gambar 2.3 dengan keterangan sebagai berikut :

* + - * 1. Lapisan Input

X1, Xi, ..., Xn = unit input 1 sampai n

* + - * 1. Lapisan Tersembunyi

Z1, Zj, ..., Zp = unit tersembunyi 1 sampai p

* + - * 1. Lapisan Output

Y1, Yk, ..., Ym = unit tersembunyi 1 sampai m

* + - * 1. Bias

V0j dan W0k adalah bias untuk unit tersembunyi ke-j dan untuk unit output ke-k

Umumnya operasi model jaringan ini terdapat dua mekanisme kerja yaitu :

1. Mekanisme latihan atau belajar (*Training mode / Learning mod* ). Pada mekanisme ini, jaringan akan dilatih untuk dapat menghasilkan data seusai dengan target yangdiharapkan melalui satu atau lebih pasangan-pasangan data (data *input* dan data target). Semakin lama waktu latihan maka kinerja jaringan akan semakin baik. Demikian juga dengan semakin banyak pasangan data yang digunakan dalam pelatihan maka kinerja akan semakin baik.
2. Mekanisme produksi (*Production Mode*) atau biasa disebut dengan mekanisme pengujian (*Try Out Mode*), pada mekanisme ini, jaringan diuji apakah dapat mengenali sesuai dengan apa yang diharapkan, setelah melalui proses pelatihan terlebih dahulu.
   * + - 1. **Metode Jaringan Saraf Tiruan Perambatan Balik**

*Backpropagation* merupakan algoritma pembelajaran yang terawasi dan biasanya digunakan oleh *perceptron* dengan banyak lapisan untuk mengubah bobot-bobot yang terhubung dengan *neuron*-*neuron* yang ada pada lapisan tersembunyinya. Algoritma *backpropagation* menggunakan *error output* untuk mengubah nilai bobot-bobotnya dalam arah mundur (*backward*). Untuk mendapatkan *error* ini, tahap perambatan maju (*feedfoward*) harus dikerjakan terlebih dahulu.

Inti dari *Backpropagation* adalah untuk mencari *error* suatu *node*. Dari hasil *forward phase* akan dihasilkan suatu *output*, dari *output* tersebut, pastilah tidak sesuai dengan target yang diinginkan. Perbandingan kesalahan dari target yang diinginkan dengan *output* yang dihasilkan disebut dengan *error*.

*Backpropagation* merupakan suatu teknik untuk meminimalisasi *gradient* pada dimensi *weight* dalam jaringan saraf tiruan lapis banyak (*multi layer Jaringan Saraf Tiruan*), proses pelatihan akan dilakukan berulang-ulang sampai nilai *error* lebih kecil dari yang ditentukan (Haykin, 1999).

Dalam proses pelatihan jaringan *Backpropagation* ini, digunakan fungsi nilai ambang batas *binary sigmoid*. Pada saat melakukan proses pelatihan, terdapat beberapa parameter jaringan yang harus ditentukan dahulu, yaitu (Fausett, 1994) :

* 1. Tingkat pelatihan (*Learning rate*) yang dilambangkan dengan parameter, harus diberikan dan me mpunyai nilai positif kurang dari satu. Semakin tinggi nilainya, maka semakin cepat kemampuan jaringan untuk belajar. Akan tetapi hal ini kurang baik. Karena *error* yang dihasilkan tidak merata, tidak merata disini maksutnyua adalah error untuk pembelajarannya tidak tentu, bisa lebih banyak di sisi tertentu.
  2. Toleransi kesalahan (*error tolerance*), semakin kecil kesalahan, maka jaringan akan memiliki nilai bobot yang lebih akurat, tetapi akan memperpanjang waktu pelatihan.
  3. Jumlah maksimal proses pelatihan yang dilakukan (*maximum epoch*), biasanya bernilai besar dan diberikan untuk mencegah terjadi perulangan tanpa akhir.
  4. Nilai ambang batas atau bias (*threshold value*), dilambangkan θ. Parameter ini tidak harus diberikan (optional). Apabila tidak diberikan, maka nilainya sama dengan nol.

Inti dari *Backpropagation* adalah untuk mencari *error* suatu *node*. Dari hasil *forward phase* akan dihasilkan suatu *output*, dari *output* tersebut, pastilah tidak sesuai target yang diinginkan. Perbandingan kesalahan dari target yang diinginkan dengan *output* yang dihasilkan disebut dengan *error*.

Dalam *Backpropagation* juga dikenal istilah yang disebut inisialisasi *output*. Inisialisasi *output* pada dasarnya adalah menentukan *error* di suatu *node* dengan sebuah target yang diinginkan. Karakteristik *Backpropagation* dapat diuraikan sebagai berikut:

1. *Node* / *processing element* dan fungsi aktivasi
2. Kontinu.
3. Dapat dideferensiasikan / diteruskan.
4. Turunan fungsi mudah dihitung.
5. Fungsi aktivasi yang biasa digunakan adalah fungsi *sigmoid*.
6. *Topology*

Terdiri dari satu lapisan masukan (*input layer*), satu atau lebih lapisan tersembunyi (*hidden layer*), dan satu lapisan keluaran (*output layer*). Setiap *neuron* / *processing* *element* pada suatu lapisan mendapat sinyal masukkan dari semua *neuron* pada lapisan sebelumnya (beserta sinyal bias).

1. *Learning Rule*

Menggunakan *delta rule* atau *error* connection *Learning rule*.