



Tentang Penulis



Andriana, ST., MT.

Lahir di Bandung 15 Maret 1969, Pendidikan Tinggi Unjani Teknik Elektro (ST) Cimahi Bandung, Pascasarjana (MT) ITB Teknik Elektro Bandung. Mengajar di UNLA, ITENAS dan UNNUR. Publikasi di jurnal nasional dan internasional. Memiliki paten tentang Speech Recognition untuk Kendali Lift bagi Penyandang Difabel. Mendapatkan dana hibah BP3ipteks Pemprov Jabar, hibah ristekdikti hibah Fundamental (Dasar) dan hibah Bersaing (Terapan) tentang Speech Recognition sejak tahun 2012 sampai sekarang.

Andriana, ST., MT.

PENGENALAN POLA

2019

PENGENALAN POLA

Andriana, ST., MT.

PENGANTAR

Pengenalan pola merupakan bidang di dalam pembelajaran mesin yang bertujuan untuk mengklasifikasikan objek berdasarkan ciri-ciri yang dimilikinya. Seperti image, berat, atau parameter-parameter lain yang ditentukan kedalam kategori atau kelas.

Seiring dengan perkembangan zaman, teknologi yang menggunakan komputer berkembang dengan pesat. Hampir setiap individu di dunia memerlukan komputer sebagai alat bantu untuk menyelesaikan masalahnya. Hampir semua sistem analog digantikan dengan sistem komputerisasi. Keunggulannya adalah sistem komputerisasi lebih mudah dalam pengontrolan. Dalam hal ini misalnya pengontrolan di dalam mengidentifikasi suatu objek atau citra. Komputer diusahakan untuk dapat bekerja mendekati proses kerja otak manusia.

Kemampuan inilah yang dikembangkan dengan menggunakan mesin (komputer). Dengan pengenalan pola kita mampu mengimplementasikan kemampuan cerdas komputer agar dapat mendekati proses otak manusia.

Penulis

DAFTAR ISI
BUKU : PENGENALAN POLA
Penulis : Andriana, ST., MT.

**BAB 1 *SPEECH RECOGNITION* DAN SINYAL UCAPAN
(*SPEECH RECOGNITION AND ACKNOWLEDGMENT SIGNALS*)**

- A. Proses *Speech Production* (*Speech Production Process*)**
- B. Tampilan Sinyal Ucapan (*Display of Speech Signals*)**
- C. Suara Dan Fitur-Fitur Ucapan (*Voice and Speech Features*)**
- D. Ucapan Otomatis Oleh Mesin/Computer (*Automatic Speech By Machine / Computer*)**

**BAB 2 METODA PEMROSESAN DAN ANALISIS SINYAL
(*METHOD OF PROCESSING AND ANALYSIS OF SIGNALS*)**

- A. Model-model Analisis Spektral (*Spectral Analysis Models*)**
- B. Model *Linear Predictive Coding* (LPC)**
- C. *Vector Quantization* (*Vector Quantization*)**
- D. *Auditory* Yang Berdasarkan Model Analisis Spektral
(*Auditory Based On Spectral Analysis Models*)**

**BAB 3 TEKNIK PERBANDINGAN POLA UCAPAN
(*COMPARISON OF ENVIRONMENTAL PATTERN TECHNIQUES*)**

- A. Deteksi Sinyal Ucapan (*Detection Of Speech Signals*)
- B. Pengukuran Distorsi (*Measurement Of Distortion*)
- C. Pengukuran Spektral-Distorsi (*spectral-distortion measurements*)
- D. *Time Alignment* Dan Normalisasi (*Time Alignment And Normalization*)

BAB 4 TEORI HIDDEN MARKOV MODELS (HMM) (*THEORY OF HIDDEN MARKOV MODELS (HMM)*)

- A. Discrete-time Markov Processes (*Discrete-Time Markov Processes*)
- B. Ekstensi dan Jenis pada HMM (*Extensions and Types on HMM*)
- C. Tiga masalah dasar pada HMM (*Three basic problems with HMM*)
- D. Variasi pada struktur HMM (*Variation in HMM structure*)

BAB 5 *SPEECH RECOGNITION*

- A. Model Kata Terhubung (*Connected word*)
- B. Model Kata Terpisah (*Isolated word*)
- C. Aplikasi Perekaman Suara (*Voice Recording Application*)

BAB 6 *SPEECH DATABASE* (BASIS DATA UCAPAN)

- A. Perancangan *Speech Database* (*Designing A Speech Database*)
- B. Sumber-sumber *Speech Database* (*Sources Of Speech Databases*)

C. Implementasi *Speech Database (Implementation Of Speech Databases)*

BAB 7 *TEXT TO SPEECH (TTS), SERTA MASALAH TEXT TO SPEECH (TTS) (TEXT TO SPEECH (TTS), AND THE PROBLEM TEXT TO SPEECH (TTS))*

A. Definisi TTS (*Definition of TTS*)

B. Akustik, Fonetik, Fonologi, Morphologi, Sintaks, Semantik, dan Pragmatik (*Acoustics, Phonetics, Phonology, Morphology, Syntax, Semantics, and Pragmatics*)

C. Ucapan Dan Tulisan, Organisasi Sistem TTS, Sistem Tts Yang Sederhana, Konsep Ucapan (*Speeches And Writings, The Organization Of The Tts System, The Tts System Is A Simple, Conceptual Greeting*)

D. Kendala dalam TTS (*Constraints in TTS*)

BAB 8 *TEKNIK MENGUBAH DARI TEKS KE FONEM (TECHNIQUES TO CHANGE FROM TEXT TO FONEM)*

A. Grammars, Inference, Parsing, and Transduction (*Grammars, Inference, Parsing, and Transduction*)

B. Arsitektur pemrosesan bahasa alamiah untuk TTS (*Synthesis (Architecture processes natural languages for TTS Synthesis)*)

C. Fonetik otomatis (*Automatic Phonetics*)

BAB 9 *TEKNIK MENGUBAH DARI FONEM KE UCAPAN (TECHNIQUE CHANGES FROM FONEM TO RECREATION)*

- A. Strategi sintesa, Linear Prediction Synthesis, Algoritma Time domain (*Synthesis strategy, Linear Prediction Synthesis, Time domain algorithm*)
- B. Pemrosesan bahasa alamiah dan TTS Synthesis (*Processing natural languages and TTS Synthesis*)
- C. Pemrosesan sinyal digital dan TTS Synthesis (*Digital signal processing and TTS Synthesis*)

BAB 10 DIPHONE STUDIO (*DIPHONE STUDIO*)

- A. Diphone Studio (*Diphone Studio*)
- B. Perekaman Suara Dalam Bentuk Kata-Kata Yang Diucapkan (*Voice Recording In The Form Of Spoken Words*)
- C. Penentuan Diphone Dalam Suatu Kata (*Determination of Diphone In A Word*)

BAB 11 PROYEK AKHIR PENGENALAN UCAPAN DAN TEKS KE UCAPAN (*FINAL RECOGNITION AND TEXT PROJECT TO RECEIVE*)

- A. Proyek Akhir (*Final Project*)
- B. Pengenalan Ucapan dan Teks (*Introduction to Speech and Text*)
- C. Ketepatan Dalam Mempresentasikan Proyek Akhir (*Accuracy in Presenting Final Projects*)

DAFTAR PUSTAKA (*BIBLIOGRAPHY*)

BAB 1

SPEECH RECOGNITION DAN SINYAL UCAPAN (SPEECH RECOGNITION AND ACKNOWLEDGMENT SIGNALS)

A. Tampilan Sinyal Ucapan (*Display of Speech Signals*)

Pengenalan ucapan (*speech recognition*), merupakan salah satu bentuk dari kecerdasan buatan. Menurut Wikipedia, yang disebut kecerdasan buatan ialah kecerdasan yang ditambahkan kepada suatu sistem yang bisa diatur dalam konteks ilmiah. Jadi, dengan fitur pengenalan ucapan, sebuah sistem komputer dapat menerjemahkan perkataan yang diucapkan oleh manusia menjadi perintah program.



Ilustrasi pengenalaan ucapan. Image taken from www.walkthechat.com

Pengenalan ucapan atau *Speech Recognition (SR)*, adalah suatu pengembangan teknik dan sistem yang memungkinkan komputer untuk menerima masukan berupa kata yang diucapkan.

Teknologi ini memungkinkan suatu perangkat untuk mengenali dan memahami kata-kata yang diucapkan dengan cara digitalisasi kata dan mencocokkan sinyal digital tersebut dengan suatu pola tertentu yang tersimpan dalam suatu perangkat. Kata-kata yang diucapkan diubah bentuknya menjadi sinyal digital dengan cara mengubah gelombang suara menjadi sekumpulan angka yang kemudian disesuaikan dengan kode-kode tertentu untuk mengidentifikasi kata-kata tersebut. Hasil dari identifikasi kata yang diucapkan dapat ditampilkan dalam bentuk tulisan atau dapat dibaca oleh perangkat teknologi sebagai sebuah komando untuk melakukan suatu pekerjaan, misalnya penekanan tombol pada telepon genggam yang dilakukan secara otomatis dengan komando suara.

Alat pengenalan ucapan, yang sering disebut dengan *speech recognizer*, membutuhkan sampel kata sebenarnya yang diucapkan dari pengguna. Sampel kata akan didigitalisasi, disimpan dalam komputer, dan kemudian digunakan sebagai basis data dalam mencocokkan kata yang diucapkan selanjutnya.

Pengenalan ucapan dalam perkembangan teknologinya merupakan bagian dari pengenalan suara (proses identifikasi seseorang berdasarkan suaranya). Pengenalan suara sendiri terbagi menjadi dua, yaitu pengenalan pembicara (identifikasi

suara berdasarkan orang yang berbicara) dan pengenalan ucapan (identifikasi suara berdasarkan kata yang diucapkan).

Jenis-jenis pengenalan ucapan

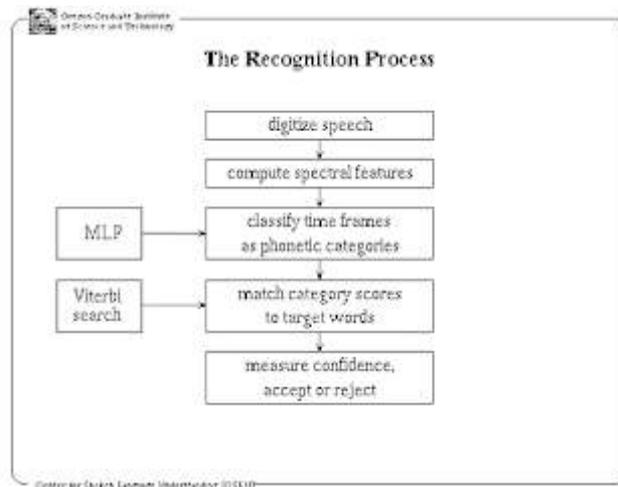
Berdasarkan kemampuan dalam mengenal kata yang diucapkan, terdapat 5 jenis pengenalan kata, yaitu:

- Kata-kata yang terisolasi
Proses pengidentifikasian kata yang hanya dapat mengenal kata yang diucapkan jika kata tersebut memiliki jeda waktu pengucapan antar kata.
- Kata-kata yang berhubungan
Proses pengidentifikasian kata yang mirip dengan kata-kata terisolasi, namun membutuhkan jeda waktu pengucapan antar kata yang lebih singkat.
- Kata-kata yang berkelanjutan
Proses pengidentifikasian kata yang sudah lebih maju karena dapat mengenal kata-kata yang diucapkan secara berkesinambungan dengan jeda waktu yang sangat sedikit, sehingga pengguna perangkat ini dapat mengucapkan kata-kata secara natural.
- Kata-kata spontan
Proses pengidentifikasian kata yang dapat mengenal kata/kalimat yang diucapkan secara spontan.
- Verifikasi atau identifikasi suara
Proses pengidentifikasian kata yang tidak hanya mampu mengenal kata, namun juga mengidentifikasi siapa yang berbicara.

Perangkat keras yang dibutuhkan

1. Sound Card
2. Microphone
3. Computer/processor

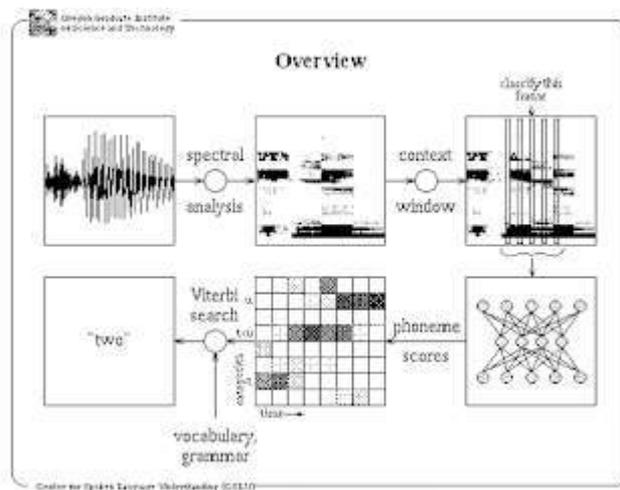
Proses kerja alat pengenalan ucapan



Alat pengenal ucapan memiliki enam tahapan dalam prosesnya, yaitu:

1. **Tahap penerimaan masukan:** Masukan berupa kata-kata yang diucapkan lewat mikrofon.
2. **Tahap *Pre-filtering*:** *pre-emphasis, normalization, banding,* dsb.
3. ***Framing and Windowing*:** Merubah data ke format yang dapat diproses, yaitu dari gelombang kontinu spektrum suara ke dalam bentuk diskrit.

4. **Filtering:** memfilter sinyal masukan dari setiap window/ frame/ frequency band
5. **Comparison and Matching :** Proses ini dilakukan berdasarkan metode Model Markov Tersembunyi atau Hidden Markov Model (HMM). Setiap elemen dari data yang terurut dikonversi ke dalam bentuk bilangan biner. Data biner tersebut nantinya akan dibandingkan dengan pola data suara dan kemudian diterjemahkan sebagai keluaran yang dapat berbentuk tulisan ataupun perintah pada perangkat.
6. **Action:** Alat pengenalan ucapan yang sudah memiliki sistem verifikasi/identifikasi suara akan menerjemahkan suara tersebut menjadi tulisan atau komando.



Tahapan ekstraksi pengenalan ucapan berdasarkan *Hidden Markov Model* (HMM) :

- a. Tahap ekstraksi tampilan: Penyaringan sinyal suara dan pengubahan sinyal suara analog ke digital
- b. Tahap pemodelan: Pembuatan suatu model HMM dari data-data yang berupa sampel ucapan sebuah kata yang sudah berupa data digital
- c. Tahap sistem pengenalan HMM: Penemuan parameter-parameter yang dapat merepresentasikan sinyal suara untuk analisis lebih lanjut.
- d. Tahap perbandingan: Tahap ini merupakan tahap pencocokan data baru dengan data suara (pencocokan tata bahasa) pada pola. Tahap ini dimulai dengan proses konversi sinyal suara digital hasil dari proses ekstraksi ke dalam bentuk spektrum suara yang akan dianalisa dengan membandingkannya dengan pola suara pada basis data.

Aplikasi alat pengenalan ucapan

- **Bidang komunikasi**

- a. Komando Suara

Komando Suara adalah suatu program pada komputer yang melakukan perintah berdasarkan komando suara dari pengguna. Contohnya pada aplikasi Microsoft Voice yang berbasis bahasa Inggris.

- b. Pendiktean

Pendiktean adalah sebuah proses mendikte yang sekarang ini banyak dimanfaatkan dalam pembuatan laporan atau penelitian. Contohnya pada aplikasi Microsoft

Dictation yang merupakan aplikasi yang dapat menuliskan apa yang diucapkan oleh pengguna secara otomatis.

c. Telepon

Pada telepon, teknologi pengenalan ucapan digunakan pada proses penekanan tombol otomatis yang dapat menelepon nomor tujuan dengan komando suara.

- **Bidang kesehatan**

Alat pengenalan ucapan banyak digunakan dalam bidang kesehatan untuk membantu para penyandang cacat dalam beraktivitas. Contohnya pada aplikasi Antarmuka Suara Pengguna atau *Voice User Interface* (VUI) yang menggunakan teknologi pengenalan ucapan.

- **Bidang militer**

a. Pelatihan Penerbangan

Aplikasi alat pengenalan ucapan dalam bidang militer adalah pada pengatur lalu-lintas udara atau yang dikenal dengan *Air Traffic Controllers* (ATC) yang dipakai oleh para pilot untuk mendapatkan keterangan mengenai keadaan lalu-lintas udara seperti radar, cuaca, dan navigasi. Alat pengenalan ucapan digunakan sebagai pengganti operator yang memberikan informasi kepada pilot.

b. Helikopter

Aplikasi alat pengenalan ucapan pada helikopter digunakan untuk berkomunikasi lewat radio dan menyesuaikan sistem navigasi.

Kelebihan dan kekurangan

Kelebihan:

- Cepat: Teknologi ini mempercepat transmisi informasi dan umpan balik dari transmisi tersebut.
- Mudah digunakan: Komando dapat dilakukan tanpa perangkat keras.

Kekurangan:

- Rawan terhadap gangguan: Karena proses ini masih berbasis frekuensi, sulit untuk memisahkan gangguan/noise dari sinyal suara
- Jumlah kata yang dapat dikenal terbatas: Hal ini disebabkan pengenalan ucapan bekerja dengan cara mencari kemiripan dengan basis data yang dimiliki.

B. Suara Dan Fitur-Fitur Ucapan (*Voice and Speech Features*)

Fitur bicara di Microsoft Office memungkinkan teks ke ucapan (TTS) di OneNote, Outlook, PowerPoint, dan Word. Dengan mengunduh mesin TTS tambahan untuk bahasa lain, Anda bisa mengubah pilihan suara untuk mendengar teks yang dibaca dalam bahasa tersebut. Mesin TTS juga perlu mendengar terjemahan dengan penerjemah Mini.

Untuk mempelajari cara menggunakan fitur TTS di Excel, lihat konversi teks ke ucapan di Excel.

Menambahkan mesin TTS untuk bahasa lainnya

Untuk menambahkan mesin bahasa lainnya, secara gratis, memerlukan Platform Speech yang tepat, lalu harus mengunduh mesin TTS dalam bahasa yang diinginkan.

Mengunduh Platform Microsoft Speech

Untuk mengunduh dan menginstal Platform Microsoft Speech, ikuti langkah berikut ini:

1. Tentukan platform sistem operasi mana yang dimiliki.
2. Masuk ke Microsoft Download Center, lalu Unduh Microsoft Speech Platform fornya.
3. Ikuti instruksi untuk menginstal platform ucapan di komputer.

Mengunduh mesin bahasa TTS tambahan

Untuk mengunduh dan menginstal mesin TTS dalam bahasa tambahan, ikuti langkah berikut ini:

1. Memilih bahasa TTS Anda dari daftar berikut ini.

Bahasa yang tersedia

Mesin TTS bahasa Cina (Hong Kong) - Unduh

Mesin TTS bahasa Cina (RRC) - Unduh

Mesin TTS bahasa Cina (Taiwan) - Unduh

Mesin TTS bahasa (Denmark Denmark) - Unduh

Mesin TTS bahasa (Belanda Belanda) - Unduh

Mesin TTS bahasa Inggris (Australia) - Unduh

Mesin TTS bahasa Inggris (Kanada) - Unduh

Mesin TTS bahasa Inggris (India) - Unduh

Mesin TTS bahasa Cina (Hong Kong) - Unduh
Mesin TTS bahasa Inggris (United Kingdom) - Unduh
Mesin TTS bahasa Inggris (Amerika Serikat) - Unduh
Mesin TTS bahasa (Finlandia Finlandia) - Unduh
Mesin TTS bahasa Prancis (Kanada) - Unduh
Mesin TTS bahasa (Prancis Prancis) - Unduh
Mesin TTS bahasa Jerman (Jerman) - Unduh
Mesin TTS bahasa (Italia Italia) - Unduh
Mesin TTS bahasa (Jepang Jepang) - Unduh
Mesin TTS bahasa (Korea Korea) - Unduh
Mesin TTS Bahasa (Norwegia Norwegia) - Unduh
Mesin TTS bahasa (Polandia Polandia) - Unduh
Mesin TTS bahasa Portugis (Brasil) - Unduh
Mesin TTS bahasa Portugis (Portugal) - Unduh
Mesin TTS bahasa (Rusia Rusia) - Unduh
Mesin TTS bahasa Spanyol (Catalan) - Unduh
Mesin TTS bahasa Spanyol (Meksiko) - Unduh
Mesin TTS bahasa Spanyol (Spanyol) - Unduh
Mesin TTS bahasa (Swedia Swedia) - Unduh

2. Ikuti instruksi untuk menginstal mesin TTS untuk bahasa yang dipilih.
3. Ulangi langkah 1 dan 2 untuk bahasa tambahan apapun.
4. Mulai ulang program Office agar perubahan bisa diterapkan.

Menambahkan Speak ke pita atau Toolbar Akses Cepat

Sebelum menggunakan Speak, harus menambahkannya ke pita atau Toolbar Akses Cepat. Untuk mempelajari cara menambahkan bicara ke pita Anda, baca mengkustomisasi pita di Office.

Menggunakan Bicara

Setelah menginstal platform dan mesin TTS bahasa yang benar, gunakan bicara untuk mendengar pemutaran teks dalam bahasa yang dipilih.

Untuk mendengar pemutaran teks, ikuti langkah berikut:

1. Pilih string kata atau teks yang ingin didengar diucapkan.
2. Klik Bicara di pita atau Toolbar Akses Cepat.
3. Klik Bicara lagi untuk menghentikan pemutaran.

C. Ucapan Otomatis Oleh Mesin/Komputer (*Automatic Speech By Machine / Computer*)

Perkembangan komputer sampai saat ini sangat pesat, sebelum mengenal komputer seperti saat ini, 5000 tahun yang lalu di Asia kecil orang menemukan alat yang disebut Abacus dan dianggap sebagai awal mula komputer. Pada tahun 1642, Blaise Pascal menemukan kalkulator roda nomerik untuk membantu ayahnya melakukan perhitungan pajak. Tetapi alat ini memiliki kelemahan, yaitu hanya sebatas melakukan penjumlahan. Komputer sendiri di artikan Hamacher sebagai mesin penghitung elektronik yang cepat dan dapat menerima input digital kemudian

memrosesnya sesuai dengan program yang tersimpan dimemorinya dan menghasilkan output berupa informasi. Menurut Nasotion(2001), komputer dibagi menjadi beberapa generasi. Yaitu generasi pertama (1953-1958), generasi kedua(1958-1966), generasi ketiga (1966-74), generasi keempat (1974-1982), dan generasi kelima (1982-sekarang). Dengan perkembangannya yang semakin canggih, maka sampai saat ini banyak dirasakan manfaatnya dalam berbagai bidang kehidupan.

Salah satu manfaat komputer adalah dalam bisnis dan perbankan, *bidang pendidikan*, industri dan manufaktur, teknik dan ilmu pengetahuan, penerbangan dan kemiliteran. Misalnya dalam bidang pendidikan dimana dengan pemanfaatan multimedia komputer, proses pembelajaran lebih bermakna, karena mampu menampilkan teks, warna, suara, video, gerak, gambar serta mampu menampilkan kepintaran yang dapat menyajikan proses interaktif.

Teknologi jaringan komputer merupakan solusi yang dapat dimanfaatkan untuk mengatasi fix-point to point dan biaya, serta menjadi model fleksibel multi point to multi point. Pesatnya perkembangan dunia jaringan komputer akhir-akhir ini, memicu berkembangnya teknologi baru yang memanfaatkan teknologi jaringan komputer sebagai media untuk mewujudkan impian manusia akan sebuah aplikasi.

Mesin sekuensial

Mesin sekuensial terdiri atas struktur masukan-transisi-keluaran. Struktur transisi meliputi: state (keadaan), masukan dan fungsi transisi. Struktur keluaran meliputi: keadaan, keluaran dan fungsi keluaran. Karena struktur keluaran adalah bagian eksternal dari mesin, maka lebih tepat untuk aplikasi praktis daripada struktur transisi yang merupakan bagian internal. Automata (automaton) ialah suatu mesin yang bekerja dengan sendirinya (otomatis). Sebuah automaton merupakan struktur transisi dari mesin sekuensial. Sedangkan sekuensial bermakna urutan, maka mesin sekuensial adalah mesin yang keluarannya tergantung kepada struktur rangkaiannya. Atau dengan kata lain, keluaran mesin sekuensial tergantung kepada keadaan sebelumnya, yaitu urutan dari masukan terakhir yang pernah diberikan kepada mesin dan keadaan dari mesin sebelum urutan tersebut dikenakan. Dua kombinasi inilah yang menyatakan keadaan sekarang (present state) dari mesin sekuensial. Dengan demikian, mesin diarahkan ke suatu keadaan tertentu, sehingga kita dapat memprediksi keluarannya. Prosedur mengarahkan mesin ke suatu keadaan yang diketahui ini disebut mereset mesin (resetting machine).

Paling tidak ada 2 hal penting yang harus diperhatikan. Pertama, karena adanya masukan, perubahan apakah yang terjadi pada mesin sekuensial. Kedua, perubahan yang terjadi menjadi informasi masukan bagi mesin sekuensial berikutnya, apabila perubahannya jelek akan mempebgaruhi hasil.

Banyak untai sekuensial yang merupakan realisasi secara elektronis dari mesin sekuensial. Beberapa diantaranya seperti sistem switching telepon, elevator otomatis, sistem kontrol lampu lalu lintas dan untai-untai elektronis dalam kalkulator. Sebuah mesin sekuensial elektronis adalah sebuah peralatan yang mampu menampilkan sekumpulan karakteristik berurutan. Untai ini merealisasikan urutan sekuensial yang diperlukan dan diberikan. Karena keadaan berikutnya dan keluaran dari sebuah mesin sekuensial tergantung pada keberadaan masa lalu dari mesin, maka mesin sekuensial elektronis harus menggunakan peralatan yang mempunyai kemampuan untuk “mengingat” tingkah laku sebelumnya dari mesin. Sebuah untai kombinasi tidak dapat menyatakan keberadaan masa lalunya, karena keluarannya hanya tergantung pada masukan saat ini.

Pemodelan Mesin Sekuensial

Untuk memudahkan mempelajari dan mengaplikasikan mesin sekuensial, digunakanlah sistem pemodelan. Dua model yang umum dipergunakan yaitu model Mealy dan model Moore. Perbedaan keduanya hanya terletak pada bagaimana menempatkan keluaran.

Dalam model atau mesin Moore, keluaran dianggap seperti menempel pada keadaan saat ini (present state), bukan pada masa transisinya dan keluaran didapatkan secara otomatis ketika mesin berada atau memasuki keadaan tersebut. Pada mesin Mealy keluaran didapatkan pada masa transisinya. Keduanya sebenarnya dapat digunakan untuk mewakili tugas

yang sama dengan hasil yang tidak berbeda. Pemilihan model kemudian lebih dikarenakan kemudahan dalam penggunaannya.

Teori pemrograman Pengenal Ucapan

Berkata ataupun berbicara merupakan cara yang paling alami dan efisien dalam bertukar informasi. Perkembangan yang pesat dalam ilmu komputer mengakibatkan banyak ahli mencoba menerapkan teknologi pengenal percakapan sehingga memudahkan interaksi dengan komputer. Teknologi pengenal perkataan (speech recognition technology) bertujuan menciptakan mesin yang mampu menerima informasi lewat perkataan serta beraksi dengan cepat dan tepat sesuai dengan informasi tersebut.

Penelitian tentang pengenal perkataan adalah bagian dari pengembangan teknologi komputer dengan kecerdasan buatan (artificial intelligent) yang dapat “mendengar”, “mengerti” dan “bereaksi” sesuai informasi perkataan yang diberikan. Disisi lain, banyak hal yang menyebabkan perkembangan teknologi pengenal ucapan ini berjalan begitu lambat. Sebuah mesin komputer yang diciptakan memiliki kemampuan pengenal ucapan terbagi dalam dua bagian besar. Pertama, bagian yang merubah sinyal suara (akustik) ke bentuk sinyal-sinyal digital. Kedua, bagian yang mengolah sinyal-sinyal ini, menganalisisnya, mengetahui bentuk perintah yang dimaksud, kemudian memberikan respon dengan melakukan perintah tersebut. Bagian ini terdiri dari bagian yang menganalisa struktur leksikal, sintaksis, semantik dan terakhir analisa pragmatis

sehingga menghasilkan sebuah pesan yang jelas. Dalam konteks tersebut, analisa grammar melibatkan analisa leksikal dan sintaksis.

Untuk sebuah mesin sekuensial, bagian akustik lebih banyak melibatkan masalah matematika, bagaimana mengubah satu sinyal ke bentuk diskret (yang diolah komputer). Sementara itu bagian linguistik mensyaratkan pemahaman-pemahaman terhadap beberapa teori yang sifatnya abstrak seperti teori tentang linguistik dan teori bahasa formal yang dapat diterima oleh mesin sekuensial. Di bawah dibahas sekilas tentang bahasa formal.

Bahasa Formal

Perkataan formal mengacu pada kenyataan bahwa semua peraturan untuk bahasa dinyatakan secara eksplisit yaitu string-string apa yang dapat terjadi. Di sini perkataan formal ditekankan pada bentuk simbol string bukan kepada arti. Teori bahasa formal diperkenalkan pertamakali oleh ahli bahasa Noam Chomsky (1950), ia memberikan karakteristik pada struktur bahasa-bahasa seperti Bahasa Inggris ataupun Bahasa Jerman dengan aturan matematika formal. Chomsky berusaha mencari suatu penggambaran syntax dari bahasa-bahasa melalui beberapa aturan sederhana dan transformasi, ia menganjurkan beberapa aturan-aturan pembatasan, salah satunya tentang grammar atau context-free grammar.

Grammar ialah sekumpulan aturan dengan simbol-simbol (dalam konteks pengenalan ucapan adalah fonem-fonem)

dikombinasikan dalam bahasa alami. Sedangkan bahasa (language) merupakan kumpulan dari seluruh kemungkinan kombinasi simbol-simbol.

Konsep bahasa dan grammar dinyatakan sebagai fenomena yang dilihat sebagai kesatuan struktur dihasilkan dengan membuatnya dari pola-pola primitif sesuai dengan aturan-aturan tertentu. Ucapan atau perkataan adalah fenomena yang terpola seperti tersebut, tapi kemudian dapat membentuk grammar yang menyatakan bentuk dari gambar-gambar, kumpulan karakter (string-string) dari bilangan biner/kode komputer. Pola primitif yang berasosiasi dengan bahasa formal disebut terminal atau batas dari bahasa. Bahasa itu sendiri didefinisikan sebagai kumpulan dari seluruh string-string kalimat yang dihasilkan dengan aturan-aturan dari grammar formal. Untuk itu perlu definisikan sebuah mesin abstrak sekuensial (automaton) yang mempunyai kemampuan mengetahui aturan-aturan dari sebuah grammar untuk menghasilkan setiap elemen dari bahasa. Model yang paling banyak dipakai adalah Hidden Markov Model (HMM), pada dasarnya merupakan sebuah finite state automata menghasilkan bahasa yang tersusun dengan grammar umum. Grammar tersebut digunakan untuk memodelkan kode yang dihasilkan oleh ucapan pada setiap tingkatan proses linguistik dan akustik.

BAB 2

METODA PEMROSESAN DAN ANALISIS SINYAL (*METHOD OF PROCESSING AND ANALYSIS OF SIGNALS*)

A. Model-model Analisis Spektral (*Spectral Analysis Models*)

Kebutuhan prakiraan cuaca merupakan kebutuhan utama untuk mendukung kegiatan di berbagai sektor, sehingga upaya pengembangan metode prakiraan menuju ketepatan dan keakuratan informasi yang tinggi sangat diperlukan. Berbagai model prakiraan cuaca dengan pendekatan teknik/metode stokastik telah dikembangkan meskipun setiap metode memiliki kelemahan dan kelebihan, namun upaya pengembangan teknik/metode untuk mendapatkan model terbaik harus tetap dilakukan. Dalam makalah ini akan dibahas mengenai identifikasi model taksiran dalam domain frekuensi (frequency domain) yang tepat dengan menggunakan metode analisis spektral untuk mengetahui periodisitas cuaca di kota Surabaya. Analisis spektral adalah analisis deret waktu yang dapat menguraikan data ke dalam gelombang sinus dan atau cosinus pada berbagai frekuensi yang dapat digunakan untuk mencari periodisitas tersembunyi. Data yang digunakan adalah data pemantauan kondisi cuaca di kota Surabaya periode 1 Oktober 2004 sampai dengan 31 Desember 2011. Variabel yang digunakan Suhu/Temperature (oC) dan Kelembaban/Humidity (%). Hasil analisis diperoleh model peramalan terbaik untuk variabel Suhu/Temperature (oC) dengan frekuensi sebanyak 13 peak dan model peramalan terbaik untuk variabel Kelembaban/Humidity (%) dengan frekuensi

sebanyak 14 peak. Model terbaik ditentukan dengan besaran nilai koefisien determinasi dari model yang sudah divalidasi (R^2 Pred) yaitu masing-masing sebesar 95% dan 75%.

B. Model *Linear Predictive Coding* (LPC)

Suara merupakan alat komunikasi paling mendasar bagi manusia. Saat ini, suara tidak hanya digunakan untuk komunikasi antar manusia tetapi juga digunakan untuk berkomunikasi dengan perangkat-perangkat teknologi, oleh sebab itu topik pengenalan suara semakin berkembang.

Hal yang dilakukan untuk mengenali konsonan dengan cara merekam suara vokal terlebih dahulu dan setelah itu merekam konsonan dari beberapa orang dan hasil rekamannya dicatat dan dilatih mengetahui seberapa besar konsonan terhadap vokal dapat dikenali.

C. *Vector Quantization* (*Vector Quantization*)

Kuantisasi vektor belajar (KVB), (bahasa Inggris: *learning vector quantization* (LVQ)) adalah ilmu komputer yang merupakan algoritme klasifikasi untuk melihat pola prototipe. LVQ merupakan sistem *kuantisasi vektor* pasangan pengawas.

Ikhtisar

LVQ dapat dimengerti dengan kasus khusus dari jaringan saraf buatan, lebih tepat, KVB menerapkan pencapaian *pembelajaran Hebbian pemenang-mendapatkan-segala*. KVC

merupakan pendahulu dari *peta-yang-mengatur-dirinya-sendiri* dan berhubungan dengan salah satu materi JST Gas Alam (Neutral Gas), dan algoritme tentang k-terdekat (k-NN). LVQ ditemukan oleh **Teuvo Kohonen**.

Sistem LVQ diwakili oleh prototipe $W=(w(i),\dots,w(n))$ yang didefinisikan pada ruang fitur pada data yang diobservasi. Pada algoritme latihan *pemenang-mendapatkan-segala* menentukan, pada setiap titik data, prototipe yang paling dekat dengan input menurut pengukuran jarak yang diberikan. Posisi ini kemudian diadaptasi, seperti contoh pemenang (prototipe) bergerak mendekat jika secara benar mengklasifikasikan titik data atau berpindah jika prototipe mengklasifikasikan titik data secara tidak benar.

Keuntungan dari LVQ ialah ia membuat prototipe yang dapat secara mudah diterapkan pada domain aplikasi masing-masing yang sulit. Sistem LVQ diterapkan pada klasifikasi masalah multi-kelas secara alami. LVQ digunakan secara ragam pada penerapan praktis.

Masalah utama dari LVQ merupakan pilihan pada pengukuran yang sesuai dari jarak atau kesamaan pada setiap latihan dan klasifikasi. Akhir-akhir ini, teknik LVW telah dikembangkan untuk beradaptasi dengan pengukuran jarak yang telah memiliki parameter pada arah dari sistem training, lihat lebih lanjut pada buku (Schneider, Biehl, and Hammer, 2008) dan referensi therein.

LVQ dapat menjadi sumber dalam membantuk mengklasifikasi dokumen teks.

Basis data merupakan sesuatu yang selalu dipakai pada penyimpanan data-data baik dalam jumlah sedikit maupun banyak. Alasan penyimpanan dalam basisdata adalah keamanan data dan kemudahan dalam pencarian jika akan digunakan. Seiring berjalannya waktu, basisdata berkembang semakin kompleks. Kompleksitas tersebut dalam hal konsep maupun ragam data yang mampu disimpan di dalamnya. Jika dulu data yang dapat disimpan hanya berupa tulisan dan angka, maka sekarang basisdata tidak hanya mampu menyimpan tulisan dan angka, tapi juga gambar. Karena tidak semua data dapat dicari dengan metode yang sama, maka diperlukan metode pencarian data sesuai dengan tipe data yang dimaksud. Pertama dilakukan pembuatan codebook sebagai acuan untuk melakukan indexing basisdata. Codebook ini dihasilkan dari algoritma Fair Share Amount (FSA).

Tahapan selanjutnya adalah melakukan penyimpanan data dengan menggunakan acuan codebook yang sudah dibuat. Untuk mencari gambar, dilakukan dengan memasukkan parameter pencarian berupa gambar ke aplikasi dan mencocokkan parameter pencarian yang sudah terindeks dengan menggunakan codebook yang sama dengan yang digunakan pada penyimpanan. Ujicoba dan evaluasi dilakukan dengan menggunakan sejumlah gambar yang diambil secara acak dari yang sudah ada dalam basisdata untuk menguji akurasinya.

Pilihan yang digunakan untuk mengatur akurasi pencarian adalah tepat sama (kecocokan 100%) dan kesamaan (kecocokan 50%-90%). Selain itu juga digunakan gambar yang tidak ada dalam basisdata untuk menguji validitas pencarian. Data yang tersimpan dalam basisdata sejumlah 50 data. Hasil uji coba menunjukkan bahwa yang berpengaruh pada pencarian adalah ukuran gambar dan codebook yang digunakan, sedangkan warna dan pola gambar tidak berpengaruh pada pencarian

D. *Auditory Yang Berdasarkan Model Analisis Spektral (Auditory Based On Spectral Analysis Models)*

Persepsi pendengaran menyoroti berbahaya, sering diabaikan, dan dalam hal ini mematikan, masalah . Untuk meringkas, bahwa " kelelahan alarm " ditemukan menjadi faktor dalam kematian pasien jantung pada pejabat Rumah Sakit Umum Massachusetts . Keamanan percaya hal ini bukanlah insiden yang terisolasi , tetapi masalah nasional dalam perangkat yang alarm sering diabaikan , terdengar , sengaja dimatikan , atau sengaja dinonaktifkan karena efek mengganggu mereka memiliki staf .

Kelelahan pendengaran , atau desensitisasi pendengaran, terjadi di banyak lingkungan kerja di mana kebutuhan persepsi pendengaran pergi terpenuhi . Lingkungan pendengaran miskin menantang kemampuan kita memahami situasi , membuat keputusan yang tepat , dan menanggapi secara tepat waktu . Untungnya tantangan ini dapat dan harus ditangani melalui desain sistem pendengaran yang sesuai . Sebuah pengantar singkat ke

sistem pendengaran akan memungkinkan kita untuk lebih memahami kebutuhan persepsi.

Suara memainkan peran yang mendalam dan meresap dalam pengalaman kita tentang realitas , namun sebagian besar dari pengalaman kami dengan suara pergi sebagian besar tidak diketahui . Kami menanggapi suara dengan cepat , intuitif , dan secara naluriah, dan masih, beberapa ruang sepenuhnya mempertimbangkan implikasi dari lingkungan pendengaran kita .Jawabannya mungkin terletak pada fakta sederhana bahwa banyak dari pengalaman kami dengan suara pra - penuh perhatian, dan pemahaman kita tentang lingkungan pendengaran terutama sadar . Jelas kesempatan telah terjawab.

Meskipun ada banyak kesamaan antara persepsi kita tentang informasi visual dan pendengaran, ada beberapa perbedaan mendasar. Salah satu perbedaan utama adalah bahwa peristiwa pendengaran yang dinamis -mereka mengalami over time (Stevens , 1938) . Para peneliti merujuk pada fakta ini sebagai pendengaran streaming. Aliran pendengaran adalah unit perseptual suara , seperti objek adalah unit persepsi penglihatan (Bregman, 1990). Sebagai contoh, aliran pendengaran bisa menjadi hum rendah dari kipas pemanas atau sekelompok orang yang tenang berbicara di latar belakang . Kemampuan kita untuk memisahkan aliran pendengaran ke daerah atau kelompok yang tepat disebut sebagai analisis scene pendengaran (Bregman , 1990) .

Mengingat adegan pendengaran , benda pendengaran dapat diintegrasikan untuk membentuk aliran perseptual dalam salah satu dari dua cara : secara berurutan atau bersamaan . Integrasi Sequential berarti bahwa peristiwa pendengaran yang terhubung dari waktu ke waktu , sedangkan integrasi simultan berarti bahwa peristiwa pendengaran membagi kualitas spektral , seperti frekuensi , pada saat yang sama dalam waktu.

Para peneliti dalam ilmu persepsi sering menggunakan prinsip-prinsip Gestalt pengelompokan visual sebagai launching pad heuristik dari yang untuk melakukan studi tentang preattentive pengelompokan pendengaran (Aksentijevic , A. , Elliot , MA , & Barber , PJ , 2001) . Sebagai contoh, aliran pendengaran dengan frekuensi yang sama dikatakan memiliki kesamaan karena mereka memiliki panjang gelombang yang sama secara fisik . Sebuah kekuatan penting dari sistem pendengaran kita adalah kemampuannya untuk memisahkan aliran dengan kecepatan dan ketepatan relatif. Hal ini memiliki implikasi penting untuk desain setiap lingkungan rumah sakit , di mana salah menafsirkan data yang pendengaran dapat biaya waktu dan menimbulkan risiko kesehatan yang serius .Pengelompokan sequential.

Ini cukup mudah untuk memisahkan aliran pendengaran melalui kualitas spektral , namun itu jauh lebih menantang untuk mengatasi masalah seperti kelelahan pendengaran atau desensitisasi . Solusinya , seperti yang akan kita lihat , terletak pada pengelompokan berurutan .

Pengelompokan Sequential memungkinkan kita untuk mengalami sifat dinamis dari suara dengan menempatkan sumber pendengaran dan mengidentifikasi aliran pendengaran dari waktu ke waktu (Cusack , Deeks , Aikman , & Carlyon , 2004) . Salah satu kunci untuk mengidentifikasi sumber arus pendengaran adalah perbedaan-perbedaan kita fisiologis interaural waktu, atau ITDS . ITDS mengacu pada kemampuan kita neurologis untuk memahami perbedaan waktu antara suara memasuki telinga lebih dekat ke acara tersebut , dibandingkan suara memasuki telinga jauh dari acara tersebut , dan menghitung rasa lokasi spasial (Stevens, 1938 ; Wickens , 2004) . Sayangnya gangguan peristiwa pendengaran dan ilusi pendengaran khusus, seperti mengikat, sangat menghambat kemampuan kita untuk memahami sumber suara . Adalah penting bahwa alarm yang mematikan secara manual , dan pada sumbernya, untuk memastikan bahwa mereka tidak diabaikan atau dilupakan selama gangguan.

Ruang gawat darurat merupakan lingkungan pendengaran yang kaya dibanjiri dengan suara yang kompleks dan gangguan . Konteks lingkungan pendengaran kita sangat mempengaruhi persepsi kita tentang sifat pendengaran spektral dan temporal , dan sebagai hasilnya , kemampuan kita untuk kelompok dan memahami adegan pendengaran kami . Berikut saya daftar beberapa solusi untuk meningkatkan analisis adegan dan pengambilan keputusan .

Monitor dengan frekuensi alarm serupa harus tersebar sehingga aliran pendengaran mereka jelas terpisah , dan alarm sehingga berbeda diidentifikasi secara benar dalam keadaan darurat. Pemantauan juga harus terdesentralisasi dan di mana-mana sehingga peristiwa penting pendengaran , seperti waktu onset, dianggap sebagai berbeda dan lebih baik dipahami .

Sistem pendengaran harus didukung oleh input sensoris lain , seperti visualisasi , sehingga informasi yang diterima secara efektif oleh semua orang . Persepsi visual kami adalah sistem yang jauh lebih kuat untuk memahami informasi yang lebih kompleks.

Tantangan pendengaran kelelahan harus ditangani oleh yang lebih besar , sistem menyeluruh yang memiliki masukan dari berbagai sumber , dan yang lebih penting , memiliki kecerdasan untuk mendistribusikan informasi ini tepat dan efektif .

Menampilkan Dyanamic yang memungkinkan pengguna untuk memahami adegan pendengaran umum , seperti sumber perangkat pemantauan yang berbeda , akan sangat mengurangi beban kerja kognitif mereka dan memungkinkan mereka untuk merespon lebih tepat dan efisien

BAB 3

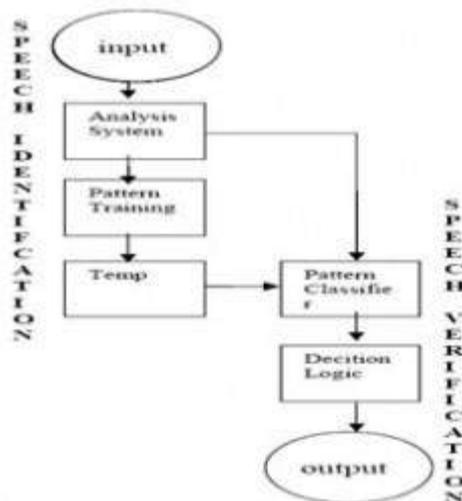
TEKNIK PERBANDINGAN POLA UCAPAN

(COMPARISON OF ENVIRONMENTAL PATTERN TECHNIQUES)

A. Deteksi Sinyal Ucapan (*Detection Of Speech Signals*)

Hampir setiap komputer mampu untuk menyimpan memori ataupun file dalam jumlah yang sangat besar. File lagu yang disimpan dalam suatu PC bisa dalam jumlah yang sangat banyak, sehingga terkadang kita merasa kesulitan untuk mencari lagu yang diinginkan. Oleh karena itu dilakukan berbagai cara untuk mempermudah dan mempercepat dalam mencari ataupun merequest lagu. Salah satu cara yang bisa dilakukan adalah dengan memberikan perintah kepada komputer dengan suara. Sinyal suara merupakan sinyal yang dihasilkan dari pergerakan pita suara dengan otak sebagai pengontrolnya. Kemudian suara ini akan dianalisa berdasarkan analisa harmonik dan diolah melalui beberapa tahapan antara lain: filter anti-aliasing, sinyal mikropon disampling dan dianalisa pada saat windowing dengan durasi yang digambarkan oleh waktu konstan. Kemampuan telinga dalam mengenali kata yang sama dengan variasi panjang berbeda akan ditirukan dalam sistem ini dan dianalisa pada front-end detection. Prinsip dari pemberian perintah melalui suara ini adalah sample suara diambil kemudian diolah menggunakan jaringan saraf tiruan metode propagasi balik dengan menggunakan personal computer. Sinyal suara analog yang digunakan terlebih dahulu

dicuplik menjadi sinyal digital. Untuk mengolah sinyal digital ini kedalam jaringan saraf tiruan, terlebih dahulu sinyal suara ditransformasikan ke domain frekuensi pada Fast Fourier Transform (FFT). Dari FFT ini kemudian sinyal baru bisa diolah ke dalam jaringan saraf tiruan back propagation. Secara keseluruhan sistem kerjanya adalah sebagai berikut:



Identifikasi pengucapan

Masukan ucapan (sampel) direkam dengan menggunakan microphone pada frekuensi 12000Hz dan durasi maksimum 1600 ms. Kemudian sampel ini disimpan dalam bentuk file wav dan nilainya dalam bentuk file txt. Sampel yang diambil untuk setiap judul lagu adalah 10 kali. Dan setiap 10 judul dijadikan dalam satu folder sebagai referensi dalam pola belajar.

Produksi suara yang dihasilkan tadi kemudian dianalisa melalui sistem analisa yang terdiri dari tahapan-tahapan sebagai berikut:

1. Deteksi awal akhir

Dalam tahap ini sebelumnya dilakukan normalisasi amplitudo untuk mengatasi jarak antara mulut dengan microphone dengan cara membagi semua nilai input dengan nilai maksimum dari input itu sendiri. sehingga diperoleh semua sinyal masukan bernilai maksimum yang sama, yaitu 1. Deteksi awal akhir digunakan untuk mendeteksi sinyal ucapan awal dan ucapan yang akhir (tidak diucapkan).

2. Frame blocking, merupakan proses untuk membagi voice menjadi beberapa bagian sehingga proses komputasi bisa lebih cepat. Hasil dari proses ini berupa sinyal terpotong yang discontinue.

3. Windowing

Windowing merupakan suatu fungsi yang berfungsi untuk mengalikan sinyal terpotong yang discontinue dengan fungsi window agar menjadi sinyal continue. Dalam penelitian ini menggunakan fungsi window hamming. Setelah diperoleh sinyal kontinu kemudian dilanjutkan ke proses FFT

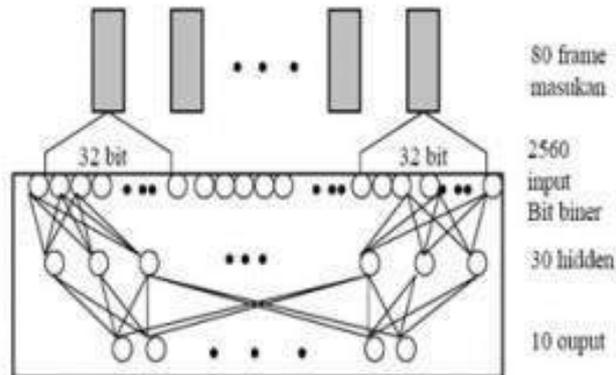
4. Fast Fourier Transform (FFT).

Dalam penelitian ini menggunakan FFT 256 dengan setiap frame mengandung 240 data. Dari data ini disamakan dengan menambahkan zero padding untuk setiap frame blok.

Data yang diambil dari hasil kepstrum adalah 32 data untuk tiap frame, dan total dari FFT ini sebanyak 2560 data. Proses kepstrum ini dilakukan melalui tahapan sebagai berikut:

- a. FFT 256
- b. Log FFT 256
- c. Invers FFT 256
- d. FFT(liftering) 32
5. Bit biner, untuk masukan ke dalam neural network data yang diperoleh diubah kadalam bentuk bit biner dengan menggunakan persamaan: $x_i/x \geq 0,5$ maka $y_i=0$
6. Pola belajar

Dalam proses ini menggunakan jaringan saraf tiruan multilayer persepton dengan metode pembelajaran backpropagation. Gambar jaringan saraf tiruan dapat dilihat sebagai berikut:



□ Verifikasi pengucapan

Dalam tahap ini ucapan yang telah masuk di sistem analisa kemudian dicocokkan denagn template yang sesuai dengan input.

Sehingga output berupa lagu yang diinginkan. Tahapan-tahapan dalam proses ini adalah

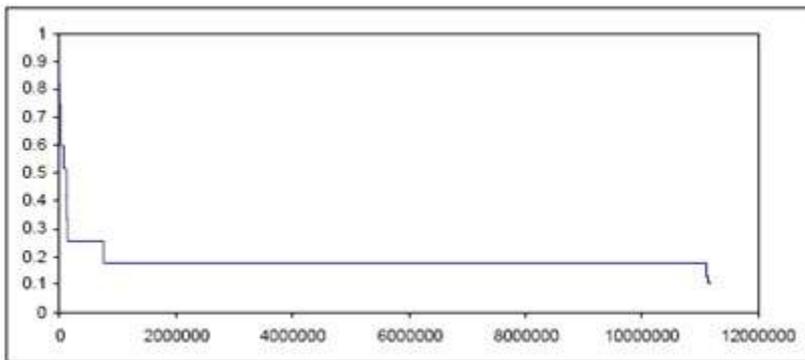
1. Sistem analisa
2. Klasifikasi pola
3. Keputusan logika

Hasil Pengujian

- Pengambilan Suara

PERCOBAAN	PANJANG DATA AWAL	FFT	JUMLAH DATA LEARNING	TARGET ERROR	LEARNING RATE	JUMLAH NODE			ITERASI	WAKTU	GRAFIK ERROR	HASIL WICIBA
						input	hidden	output				
1	tidak seragam	1024	2 X 10	0.1	0.5	32 X 16 X 10	30	10	1000	3jam	til konvergen	0%
2	tidak seragam	512	2 X 10	0.1	0.5	32 X 80 X 10	60	10	56302	33jam	til konvergen	8%
3	seragam	256	3 X 10	0.1	0.8	16 X 10 X 10	30	10	2149	12jam	til konvergen	10%
4	tidak seragam	256	3 X 10	0.1	0.5	16 X 80 X 10	30	10	65516	40jam	konvergen	30%
5	seragam	256	10 X 10	0.1	0.5	32 X 80 X 10	30	10	11165449	70jam	konvergen	98%

- Analisa Sistem



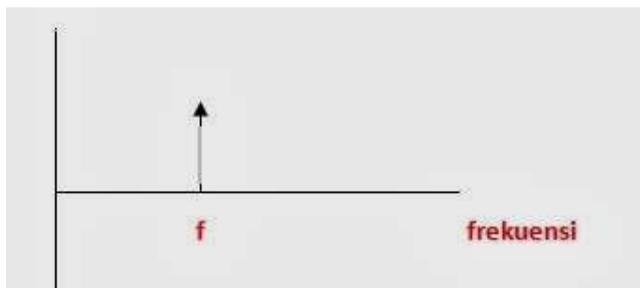
- Pengujian jaringan sistem saraf tiruan
 1. Training data set . dari sini diperoleh besarnya error sebesar 0% sehingga keakuratan untuk mengenali pola training data set mencapai 100%.

2. Blind data set

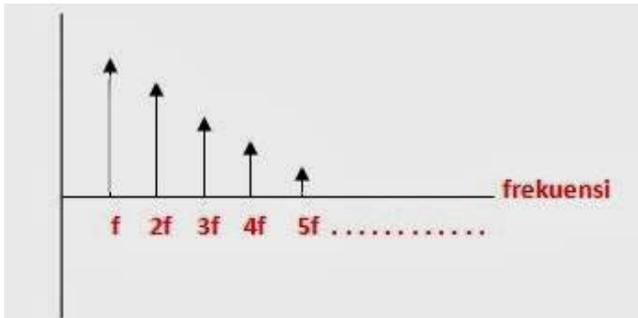
Dari proses ini diperoleh error rate sebesar 2 % sehingga keakuratan sistem dalam pengenalan pola blind data set (pengujian data terhadap sinyal suara secara langsung dari microphone untuk 10 pengucapan judul lagu masing-masing 10 kali) mencapai 98 %

B. Pengukuran Distorsi (*Measurement Of Distortion*)

Sinyal sinus sebenarnya bentuknya tidak sinus, sinyal ini memiliki harmonisa-harmonisa atau biasa disebut dengan cacat atau distorsi pada sinyal. Berikut dibawah ini penggambarannya Pada sinyal sinus kondisi ideal ini idak memiliki frekuensi harmonisa, hanya frekuensi fundamentalnya saja, seperti pada gambar dibawah ini:



Namun kondisi ini tidak sama ketika kita mengamati sinyal sinus yang tidak ideal. Sinyal sinus yang tidak ideal memiliki banyak frekuensi harmonisa seperti pada gambar dibawah ini:



Jika frekuensi yang muncul semakin banyak, maka sinyal tersebut makin tidak sempurna. Ketidaktersempurnaan ini bisa dihitung dengan rumus berikut ini:

$$\text{Cacat (distorsi)} = \left(\frac{\sum \text{Harmonisa}}{\text{fundamental}} \right) (\%)$$

Distorsi suatu sinyal dapat bertambah karena adanya suatu proses.

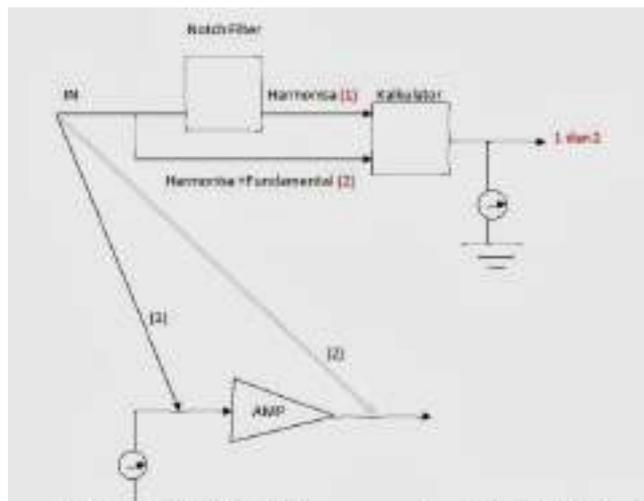
Nilai dari terdistorsi X dan Y disini tidak sama, tetapi Nilai terdistorsi Y > X karena Amplifier menambah cacat atau distorsi, Amplifier disini merupakan proses penguatan. Penambahan cacat Amplifier ini dapat dihitung yaitu dengan cara seperti berikut :

- 1.) Hitung terdistorsi X
- 2.) Hitung terdistorsi Y
- 3.) Carilah selisih antara terdistorsi X dan terdistorsi Y. Pada selisih inilah nilai cacat yang ditambahkan oleh Amplifier.

Apabila kita mengamati perbedaan radio AM dengan FM jika ditinjau dari kualitas suaranya dapat diketahui bahwa radio AM memiliki kualitas suara yang kurang bagus atau tidak jelas. Hal ini dikarenakan AM kaya akan distorsi. Contoh lain pada kehidupan sehari-hari kita yaitu ketika merekam suara musik dari kaset pita melalui Tape Recorder. Output dari rekaman tersebut dimasukkan ke notebook. Hasil rekaman tersebut pastinya tidak sebagus kualitas suara yang ada pada kaset aslinya karena notebook juga menyumbangkan distorsi.

Distorsi dapat diukur dengan menggunakan dua langkah dibawah ini, ada langkah tradisional zaman dahulu dan langkah yang kita temui saat ini atau modern, berikut uraiannya :

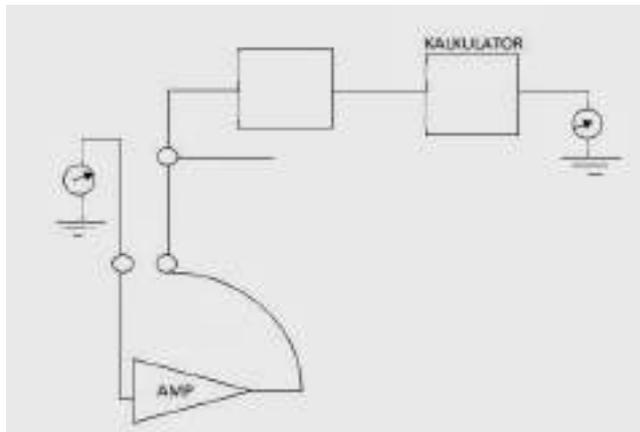
- 1.) Langkah zaman dahulu dengan menggunakan Notch



Karakteristik dari Notch Filter adalah hanya mampu membuang 1 frekuensi saja. Sehingga jika kita menggunakan

langkah ini maka setiap kali pemindahan harus *nge-Tuning* lagi, 1 kali *Tuning* membutuhkan waktu sekitar 5 menit.

Langkah saat ini

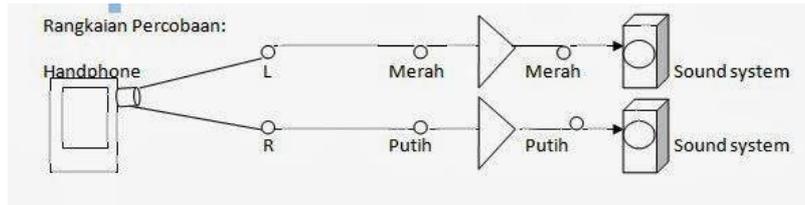


Kelebihan dari langkah ini jika dibandingkan dengan langkah yang sebelumnya adalah :

- 1.) Hanya 1 kali *Tuning* jika ingin melakukan pemindahan
- 2.) Hanya beberapa detik saja untuk 1 kali *Tuning*-nya

Peralatan yang digunakan:

- 1.) Sound System
- 2.) Handphone
- 3.) Kabel penghubung
- 4.) Amplifier



C. Pengukuran Spektral-Distorsi (*spectral-distortion measurements*)

Sebuah **analisa spektrum** mengukur besarnya sinyal masukan terhadap frekuensi dalam rentang frekuensi penuh instrumen. Penggunaan utama adalah untuk mengukur kekuatan dari spektrum sinyal dikenal dan tidak dikenal. Sinyal input langkah-langkah penganalisis spektrum adalah listrik, bagaimanapun, spektral komposisi sinyal lain, seperti akustik gelombang tekanan dan optik gelombang cahaya, dapat dipertimbangkan melalui penggunaan yang sesuai transduser. Analisa spektrum optik juga ada, yang menggunakan teknik optik langsung seperti monokromator untuk membuat pengukuran.

Dengan menganalisis spektrum dari listrik sinyal, dominan frekuensi, daya, distorsi, harmonik, bandwidth, dan lainnya spektral komponen sinyal dapat diamati yang tidak mudah terdeteksi dalam domain waktu gelombang. Parameter ini berguna dalam karakterisasi perangkat elektronik, seperti pemancar nirkabel.

Sejarah

Bagian ini membutuhkan ekspansi .(Desember 2012)



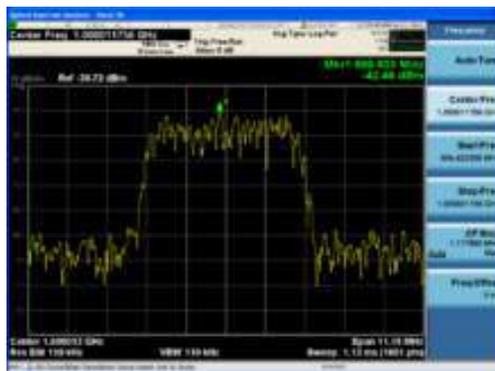
Sebuah analisa spektrum sekitar tahun 1970

The analisa spektrum pertama, pada tahun 1960, tersapu-tuned instrumen.

Menyusul penemuan Fast Fourier Transform pada tahun 1965, analisa FFT berbasis pertama diperkenalkan pada tahun 1967.

Saat ini, ada tiga jenis dasar analyzer: Analyzer menyapu-tuned spektrum, sinyal vektor analyzer, dan spektrum analyzer real-time

Jenis



Sebuah tampilan analisa spektrum yang modern

Jenis analisa spektrum ditentukan oleh metode yang digunakan untuk memperoleh spektrum sinyal. Ada menyapu-tuned dan FFT analisa spektrum berbasis:

- Sebuah spektrum analyzer *menyapu-tuned* menggunakan penerima super heterodyne untuk down-mengkonversi sebagian dari spektrum sinyal input (menggunakan osilator tegangan yang dikendalikan dan mixer) dengan frekuensi pusat dari band-pass filter . Dengan arsitektur superheterodyne, osilator tegangan yang dikendalikan adalah menyapu melalui rentang frekuensi, yang memungkinkan pertimbangan penuh rentang frekuensi dari instrumen.
- Sebuah *FFT* analyzer spektrum menghitung transformasi Fourier diskrit (DFT), proses matematis yang mengubah gelombang ke dalam komponen-komponen yang spektrum frekuensi , dari sinyal input.

Beberapa analisa spektrum, seperti analisa spektrum real-time, menggunakan teknik hibrida di mana sinyal yang masuk pertama turun-dikonversi ke frekuensi yang lebih rendah menggunakan teknik superheterodyne dan kemudian dianalisis dengan menggunakan transformasi Fourier cepat (FFT) teknik.

Faktor bentuk

Spectrum analyzer cenderung jatuh ke dalam tiga faktor bentuk: benchtop, portable dan handheld.

Benchtop

Faktor bentuk ini berguna untuk aplikasi di mana spektrum analyzer dapat dipasang ke listrik AC, yang umumnya berarti dalam lingkungan laboratorium atau area produksi / manufaktur. Bench analisa spektrum atas secara historis menawarkan kinerja dan spesifikasi yang lebih baik dibandingkan dengan faktor bentuk portabel atau genggam. Bench analisa spektrum atas biasanya memiliki beberapa penggemar (dengan ventilasi yang terkait) untuk mengusir panas yang dihasilkan oleh prosesor . Karena arsitektur mereka, bangku analisis spektrum atas biasanya memiliki berat lebih dari 30 pon (14 kg). Beberapa analisa spektrum atas bangku menawarkan opsional kemasan baterai , yang memungkinkan mereka untuk digunakan jauh dari listrik AC . Jenis analyzer sering disebut sebagai “portable” spektrum analyzer ..

Portabel

Faktor bentuk ini berguna untuk setiap aplikasi di mana spektrum analyzer perlu diambil di luar untuk membuat pengukuran atau hanya dilakukan ketika sedang digunakan. Atribut yang berkontribusi terhadap spektrum analyzer portabel berguna meliputi:

- Operasi bertenaga baterai opsional untuk memungkinkan pengguna untuk bergerak bebas di luar.

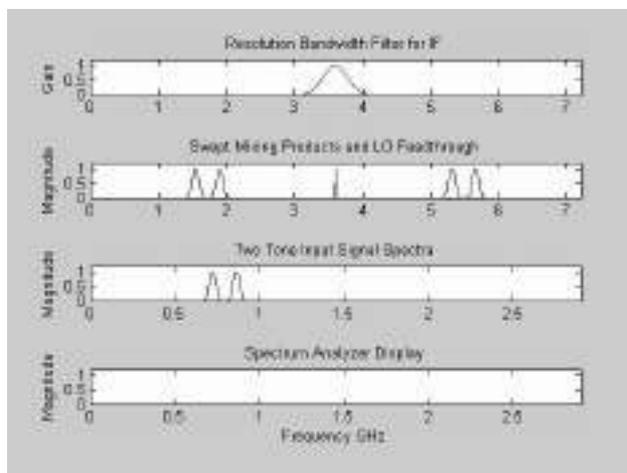
- Display jelas dapat dilihat untuk memungkinkan layar untuk dibaca di bawah sinar matahari yang cerah, kegelapan atau kondisi berdebu ..
- Ringan (biasanya kurang dari 15 pound (6,8 kg)).

Handheld

Faktor bentuk ini berguna untuk aplikasi di mana spektrum analyzer harus sangat ringan dan kecil. Analisis Handheld menawarkan kemampuan terbatas dibandingkan dengan sistem yang lebih besar. Atribut yang berkontribusi terhadap spektrum analyzer genggam berguna meliputi:

- Konsumsi daya sangat rendah.
- Baterai bertenaga operasi sementara di lapangan untuk memungkinkan pengguna untuk bergerak bebas di luar.
- Ukuran yang sangat kecil
- Ringan (biasanya kurang dari 2 pon (0,91 kg)).

Teori operasi



Dibelai-Tuned

Sebagaimana dibahas di atas dalam **jenis**, spektrum analyzer menyapu-tuned down-mengkonversi sebagian dari spektrum sinyal input dengan frekuensi pusat dari band-pass filter dengan menyapu osilator tegangan yang dikendalikan melalui berbagai frekuensi, yang memungkinkan pertimbangan rentang frekuensi penuh instrumen.

Bandwidth dari band-pass filter menentukan resolusi bandwidth, yang berkaitan dengan bandwidth minimum terdeteksi oleh instrumen. Seperti yang ditunjukkan oleh animasi ke kanan, semakin kecil bandwidth, semakin resolusi spektral. Namun, ada trade-off antara seberapa cepat tampilan akan dapat memperbarui rentang frekuensi penuh di bawah pertimbangan dan resolusi frekuensi, yang relevan untuk membedakan komponen-komponen frekuensi yang berdekatan. Untuk arsitektur menyapu-tuned, hubungan ini untuk waktu menyapu berguna:

$$ST = \frac{k(Span)}{RBW^2}$$

Dimana ST menyapu waktu dalam hitungan detik, k adalah konstanta proporsionalitas, Span adalah rentang frekuensi di bawah pertimbangan dalam Hertz, dan RBW adalah resolusi bandwidth di Hertz. Menyapu terlalu cepat, bagaimanapun, menyebabkan penurunan amplitudo ditampilkan dan pergeseran frekuensi ditampilkan.

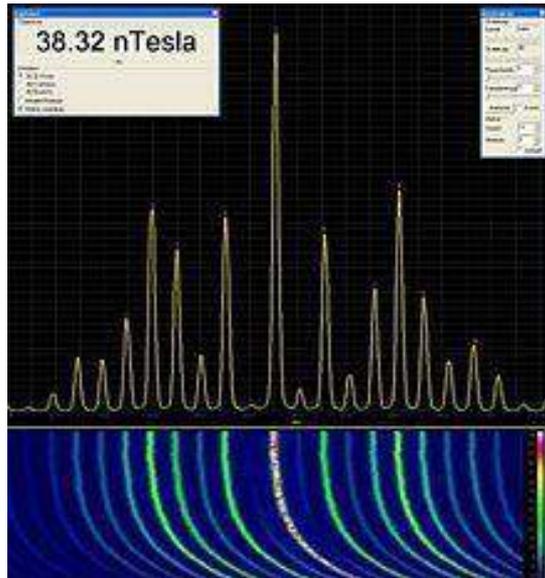
Juga, animasi berisi baik atas dan ke bawah-dikonversi spektrum, yang disebabkan oleh mixer frekuensi menghasilkan baik jumlah

dan frekuensi perbedaan. The osilator lokal feedthrough adalah karena isolasi yang tidak sempurna dari IF jalur sinyal di mixer . Untuk sinyal yang sangat lemah, pre-amplifier yang digunakan, meskipun harmonik dan intermodulasi distorsi dapat mengarah pada penciptaan komponen frekuensi baru yang tidak hadir dalam sinyal asli.

FFT Berbasis

Dengan spektrum analyzer berbasis FFT, resolusi frekuensi $\Delta\nu = 1/T$, Kebalikan dari waktu T di mana gelombang diukur dan Fourier berubah.

Dengan Transformasi Fourier analisis spektrum analyzer digital, maka perlu untuk sampel sinyal input dengan frekuensi sampling ν_s yang setidaknya dua kali bandwidth sinyal, karena batas Nyquist . Sebuah Transformasi Fourier kemudian akan menghasilkan spektrum yang mengandung semua frekuensi dari nol sampai $\nu_s/2$. Hal ini dapat menempatkan tuntutan besar terhadap diperlukan analog-to-digital converter dan daya proses untuk Transformasi Fourier, membuat berbasis FFT spektrum analyzer terbatas dalam rentang frekuensi.



Frekuensi spektrum periode memanas dari power supply beralih (spread spectrum) incl. diagram waterfall selama beberapa menit.

Hybrid Superhetrodyne-FFT

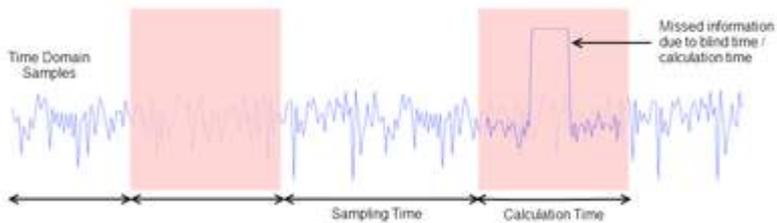
Karena analisis berbasis FFT hanya mampu mempertimbangkan band sempit, salah satu teknik adalah untuk menggabungkan menyapu dan analisis FFT untuk pertimbangan bentang lebar dan sempit. Teknik ini memungkinkan untuk waktu menyapu lebih cepat.

Metode ini dimungkinkan oleh down pertama mengkonversi sinyal, maka digitalisasi frekuensi menengah dan menggunakan superhetrodyne atau FFT teknik untuk memperoleh spektrum.

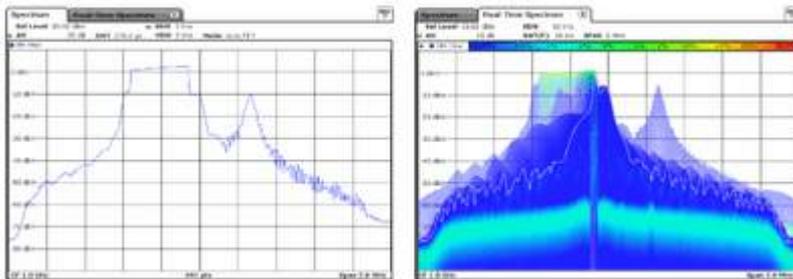
Salah satu manfaat dari digitalisasi frekuensi menengah adalah kemampuan untuk menggunakan filter digital, yang memiliki berbagai keuntungan lebih dari filter analog seperti di dekat faktor

bentuk yang sempurna dan ditingkatkan penyaring settling time. Juga, untuk pertimbangan bentang sempit, FFT dapat digunakan untuk meningkatkan waktu menyapu tanpa distorsi spektrum ditampilkan.

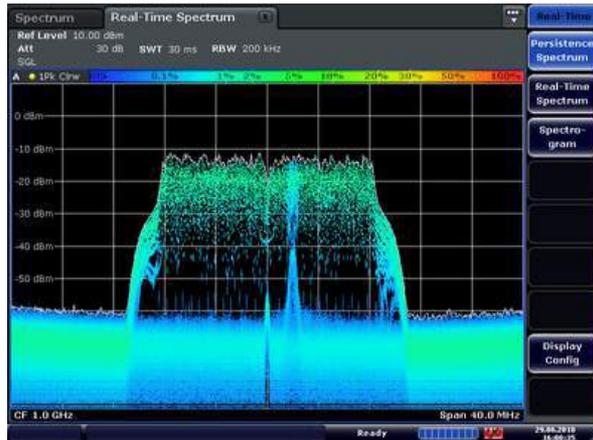
Realtime FFT



Ilustrasi menunjukkan Spectrum Analyzer Blind Waktu



Perbandingan antara Swept Max Tahan dan Realtime Kegigihan menampilkan



Sinyal Bluetooth tersembunyi di balik sinyal LAN nirkabel

Sebagian besar analisa spektrum modern sekarang hampir secara eksklusif Hybrid Superheterodyne-FFT berbasis memberikan peningkatan yang signifikan dalam waktu menyapu. Namun, bahkan dalam kasus-kasus seperti masih ada waktu pemrosesan yang dibutuhkan untuk sampel spektrum dan menghitung FFT. Untuk alasan ini, baik menyapu-tuned dan berdasarkan FFT analyzer menghasilkan “waktu buta” yang berarti bahwa sementara perhitungan spektrum sedang dilakukan, alat memiliki kesenjangan dan merindukan informasi dari spektrum RF yang diterapkan pada front end RF.

Sebuah analisa spektrum realtime tidak memiliki seperti buta waktu sampai beberapa rentang maksimum, yang sering disebut “bandwidth realtime”.Analyzer mampu untuk sampel spektrum RF yang masuk dalam domain waktu dan mengubah informasi ke domain frekuensi menggunakan proses FFT. FFT diproses secara paralel, gapless dan tumpang tindih sehingga tidak ada

kesenjangan dalam RF spektrum dihitung dan tidak ada informasi yang terlewatkan.

Realtime online dan offline realtime

Dalam arti, setiap spektrum analyzer yang memiliki vector signal analyzer adalah kemampuan analisa realtime. Ini Data sampel cukup cepat untuk memenuhi teorema Nyquist Sampling dan menyimpan data dalam memori untuk diproses kemudian. Semacam ini hanya analisa realtime untuk jumlah waktu data / capture dapat menyimpan dalam memori dan masih menghasilkan kesenjangan dalam spektrum dan hasil selama waktu pemrosesan.

FFT tumpang tindih

Meminimalkan distorsi informasi adalah penting dalam semua analisis spektrum. Proses FFT menerapkan teknik windowing untuk meningkatkan spektrum output karena menghasilkan kurang lobus samping. Pengaruh windowing juga dapat mengurangi tingkat sinyal di mana ia ditangkap di perbatasan antara satu FFT dan berikutnya. Untuk alasan ini FFT dalam spektrum analyzer Realtime yang tumpang tindih. Tingkat Tumpang Tindih adalah sekitar 80%. Sebuah analisa yang memanfaatkan proses FFT 1024 poin akan kembali menggunakan sekitar 819 sampel dari proses FFT sebelumnya.

Waktu deteksi sinyal Minimum

Hal ini terkait dengan tingkat sampling dari analisa dan FFT tingkat. Hal ini juga penting bagi penganalisa spektrum realtime untuk memberikan tingkat akurasi yang baik.

Contoh: untuk analyzer dengan 40 MHz dari realtime bandwidth (rentang RF maksimum yang dapat diproses secara realtime) sekitar 50 Msample / detik (kompleks) yang diperlukan. Jika spektrum analyzer menghasilkan 250 000 FFT / sperhitungan FFT diproduksi setiap 4 mikrodetik. Untuk titik FFT 1024 spektrum penuh diproduksi $1024 \times (1/50 \times 10^{-6})$ kira-kira setiap 20 mikrodetik. Ini juga memberikan kita kami tingkat tumpang tindih dari $80\% (20 \text{ mikrodetik} - 4 \text{ mikrodetik}) / 20 \text{ mikrodetik} = 80\%$.

Contoh tampilan realtime

Kegigihan

Analisa spektrum realtime mampu menghasilkan lebih banyak informasi bagi pengguna untuk memeriksa spektrum frekuensi secara lebih rinci. Sebuah menyapu spektrum normal akan menghasilkan max puncak, min puncak menampilkan misalnya tetapi penganalisis spektrum realtime mampu merencanakan semua dihitung FFT selama periode waktu tertentu dengan menambahkan kode warna yang mewakili seberapa sering sinyal muncul. Misalnya, gambar ini menunjukkan perbedaan antara bagaimana spektrum ditampilkan dalam tampilan spektrum

menyapu normal dan menggunakan “Kegigihan” melihat pada spektrum analyzer realtime.

Sinyal tersembunyi

Analisa spektrum realtime dapat melihat sinyal tersembunyi di balik sinyal lain. Hal ini dimungkinkan karena tidak ada informasi yang terlewatkan dan tampilan kepada pengguna adalah output dari perhitungan FFT. Contoh ini bisa dilihat di sebelah kanan.

Fungsi khas

Pusat Frekuensi Dan Rentang

Dalam spektrum analyzer khas ada pilihan untuk mengatur start, stop, dan frekuensi pusat. Frekuensi tengah antara berhenti dan mulai frekuensi pada layar penganalisis spektrum dikenal sebagai **pusat frekuensi**. Ini adalah frekuensi yang ada di tengah-tengah display frekuensi sumbu. **Span** menentukan rentang antara awal dan berhenti frekuensi. Dua parameter ini memungkinkan untuk penyesuaian tampilan dalam rentang frekuensi instrumen untuk meningkatkan visibilitas spektrum diukur.

Resolusi Bandwidth

Sebagaimana dibahas di bagian **operasi**, filter **resolusi bandwidth** atau RBW Filter adalah filter bandpass di IF jalan. Ini adalah bandwidth dari RF rantai sebelum detektor (pengukuran daya perangkat). Ini menentukan RF noise floor dan seberapa

dekat dua sinyal dapat dan masih harus diselesaikan oleh analyzer menjadi dua puncak terpisah. Mengatur bandwidth filter ini memungkinkan untuk diskriminasi sinyal dengan komponen frekuensi berjarak dekat, sementara juga mengubah lantai kebisingan diukur. Penurunan bandwidth dari filter RBW menurunkan lantai kebisingan diukur dan sebaliknya. Hal ini disebabkan filter RBW yang lebih tinggi melewati komponen frekuensi lebih sampai ke detektor amplop dari bandwidth yang lebih rendah filter RBW, karena itu RBW yang lebih tinggi menyebabkan noise diukur tinggi.

Video Bandwidth

Video Filter bandwidth atau VBW Filter adalah filter low-pass langsung setelah detektor amplop . Ini bandwidth dari rantai sinyal setelah detektor. Averaging atau puncak deteksi kemudian mengacu pada bagaimana bagian penyimpanan digital dari catatan perangkat sampel-dibutuhkan beberapa sampel per langkah waktu dan toko hanya satu sampel, baik rata-rata sampel atau yang tertinggi. The video bandwidth menentukan kemampuan untuk membedakan antara dua tingkat daya yang berbeda. Hal ini karena sempit VBW akan menghapus suara dalam output detektor. Filter ini digunakan untuk “halus” layar dengan menghapus suara dari amplop. Serupa dengan RBW, VBW yang mempengaruhi waktu menyapu layar jika VBW kurang dari RBW. Jika VBW kurang dari RBW, hubungan ini untuk waktu menyapu berguna:

$$ST = \frac{k(\text{Span})}{RBW \times VBW}$$

Dimana ST menyapu waktu dalam hitungan detik, k adalah konstanta proporsionalitas, Span adalah rentang frekuensi di bawah pertimbangan dalam Hertz, RBW adalah resolusi bandwidth di Hertz, dan VBW adalah bandwidth video dalam Hertz.

Detector

Dengan munculnya display berbasis digital, beberapa analisis spektrum modern menggunakan analog-ke-digital converter untuk sampel spektrum amplitudo setelah filter VBW. Karena menampilkan memiliki sejumlah diskrit poin, rentang frekuensi yang diukur juga didigitalkan. **Detektor** yang digunakan dalam upaya untuk secara memadai memetakan kekuatan sinyal yang benar ke titik frekuensi yang tepat pada layar. Secara umum ada tiga jenis detektor: sample, puncak, dan rata-rata

- **Deteksi sampel** – sampel hanya menggunakan deteksi titik tengah interval diberikan sebagai nilai titik layar. Sedangkan metode ini tidak mewakili gangguan acak dengan baik, tidak selalu menangkap semua sinyal sinusoidal.
- **Deteksi puncak** – puncak deteksi menggunakan titik maksimum yang diukur dalam interval diberikan sebagai nilai titik layar. Ini menjamin bahwa sinusoid maksimum

diukur dalam interval, namun sinusoid kecil dalam interval tidak dapat diukur. Juga, deteksi puncak tidak memberikan representasi yang baik dari gangguan acak.

- **Deteksi rata** – rata deteksi menggunakan semua titik data dalam interval untuk mempertimbangkan nilai titik layar. Hal ini dilakukan dengan kekuasaan (rms) rata-rata, tegangan rata-rata, atau masuk daya rata-rata.

Ditampilkan Tingkat Kebisingan Rata-Rata

The **Tampilkan Rata-rata Tingkat Kebisingan** (DANL) hanya apa yang dikatakannya itu adalah-tingkat kebisingan rata-rata ditampilkan di analisa. Ini bisa dengan resolusi bandwidth tertentu (biasanya dalam dBm), atau dinormalkan ke 1 Hz (biasanya dalam dBm / Hz)

Menggunakan frekuensi radio

Spectrum analyzer banyak digunakan untuk mengukur respon frekuensi , noise dan distorsi karakteristik semua jenis frekuensi radio (RF) sirkuit, dengan membandingkan spektrum input dan output.

Dalam telekomunikasi , analisa spektrum yang digunakan untuk menentukan bandwidth yang diduduki dan melacak sumber interferensi. Sebagai contoh, perencana sel menggunakan peralatan ini untuk menentukan sumber gangguan pada pita frekuensi GSM dan band frekuensi UMTS .

Dalam pengujian EMC , penganalisis spektrum yang digunakan untuk pengujian precompliance dasar, namun tidak dapat digunakan untuk pengujian dan sertifikasi penuh. Sebaliknya, penerima EMI seperti Rohde & Schwarz ESU EMI Receiver , Agilent Technologies N9038A MXE EMI , atau Gauss Instrumen TDEMI digunakan.

Sebuah analisa spektrum digunakan untuk menentukan apakah pemancar nirkabel bekerja sesuai dengan standar yang ditetapkan pemerintah federal untuk kemurnian emisi. Sinyal output pada frekuensi selain frekuensi komunikasi yang dimaksudkan muncul sebagai garis vertikal (pips) pada tampilan. Sebuah analisa spektrum juga digunakan untuk menentukan, dengan observasi langsung, bandwidth dari sinyal digital atau analog.

Sebuah antarmuka penganalisis spektrum adalah perangkat yang menghubungkan ke penerima nirkabel atau komputer pribadi untuk memungkinkan deteksi visual dan analisis sinyal elektromagnetik lebih band didefinisikan frekuensi. Ini disebut penerimaan panorama dan digunakan untuk menentukan frekuensi sumber gangguan pada peralatan jaringan nirkabel, seperti Wi-Fi dan router nirkabel.

Spectrum analyzer juga dapat digunakan untuk menilai RF perisai. RF shielding adalah penting bagi penentuan tapak mesin pencitraan resonansi magnetik karena ladang RF liar akan menghasilkan artefak dalam gambar MR.

Menggunakan audio-frekuensi

Analisis spektrum dapat digunakan di frekuensi audio untuk menganalisis harmonik dari sinyal audio. Sebuah aplikasi khas adalah untuk mengukur distorsi dari nominal sinewave sinyal, sebuah sinewave sangat-rendah distorsi digunakan sebagai input ke peralatan yang diuji, dan penganalisis spektrum dapat memeriksa output, yang akan telah menambahkan produk distorsi, dan menentukan distorsi persentase pada setiap harmonik fundamental. Analisa seperti itu pada satu waktu digambarkan sebagai “gelombang analisis”. Analisis dapat dilakukan dengan tujuan umum komputer digital dengan kartu suara yang dipilih untuk kinerja yang sesuai dan perangkat lunak yang sesuai. Alih-alih menggunakan sinewave rendah distorsi, input dapat dikurangi dari output, dilemahkan dan fase-dikoreksi, untuk memberikan hanya distorsi ditambahkan dan kebisingan, yang dapat dianalisis.

Sebuah teknik alternatif, pengukuran total distorsi harmonik, membatalkan fundamental dengan filter takik dan mengukur sinyal yang tersisa total, yang total harmonic distortion ditambah noise, ia tidak memberikan detail harmonik-by-harmonik analyzer.

Spektrum optik analyzer

Sebuah analisa spektrum optik menggunakan teknik reflektif dan reflektif untuk memisahkan panjang gelombang cahaya. Sebuah

detektor elektro-optik digunakan untuk mengukur intensitas cahaya, yang kemudian biasanya ditampilkan pada layar dengan cara yang mirip dengan spektrum analyzer radio-atau audio-frekuensi.

Input ke spektrum analyzer optik mungkin hanya melalui aperture dalam kasus instrumen, serat optik atau konektor optik yang kabel serat optik dapat dilampirkan.

Teknik yang berbeda ada untuk memisahkan panjang gelombang. Salah satu metode adalah dengan menggunakan monokromator, misalnya desain Czerny-Turner, dengan detektor optik ditempatkan di celah output. Sebagai kisi-kisi di monokromator bergerak, band frekuensi yang berbeda (warna) yang 'dilihat' oleh detektor, dan sinyal yang dihasilkan kemudian dapat diplot pada layar.

Respon frekuensi dari analisa spektrum optik cenderung relatif terbatas, misalnya 1600 – 800 nm (inframerah untuk red), tergantung pada tujuan yang telah ditetapkan, meskipun (agak) instrumen tujuan umum yang lebih luas bandwidth yang tersedia.

D. Time Alignment Dan Normalisasi (Time Alignment And Normalization)

Masalah adanya time alightmnet adalah karena posisi duduk kita dimobil tidak ideal. tidak seperti yang dirumah dimana kita duduk diantara 2 speaker, sehingga staging, tonal dan imaging dll di dapat dengan maksimal, dengan asumsi kondisi

yang mendukung home audio itu maksimal, seperti ruangan, akustik position dll.

Sedangkan di mobil posisi duduk kita tidak di posisi ideal dimana salah satu speaker selalu lebih dekat dengan kita, sehingga dengan demikian waktu tiba dari suara yang di hasilkan oleh speaker yang dekat dengan pendengar akan selalu lebih cepat, lebih keras dll.

Time delay dimaksudkan sedemikian rupa sehingga pada saat kita duduk speaker terdekat dengan pendengar di delay (di mundurkan) dengan hitungan sedemikian rupa sehingga waktu tiba ke kuping kita bisa sama dengan speaker terjauh dari pendengar sehingga idealnya bisa sama dengan posisi home audio.

Normalisasi database bagi saya merupakan suatu bentuk pengaplikasian / cara dalam memetakan suatu informasi data menjadi beberapa bagian / entitas agar menjadi lebih mudah dalam pengelolaan data.

Hubungan tiap entitas tersebutlah yang dimaksud sebagai sebuah normalisasi, biasanya normalisasi database terbagi menjadi beberapa tingkatan bergantung dengan seberapa kompleksnya data tersebut, seperti bentuk normalisasi pertama (1NF) , normalisasi kedua (2NF), 3NF dan BCNF. Sebelum melangkah lebih jauh, saya ingin menjelaskan bahwa dalam konteks RDBMS istilah-istilah normalisasi tersebut merujuk kepada:

- **Entitas** , yakni sebuah table, hubungan antar entitas berarti hubungan antara tiap table

- **Atribut**, yakni kelengkapan sebuah table, dalam database hal ini berarti adalah sebuah field yang tentu saja beragam
- **Identifier**, yakni atribut pengidentifikasian setiap entitas, dalam Bahasa mudahnya kita mengenalnya sebagai ID
- **Relasi/Hubungan** tiap table yang menjadi sebuah nilai *foreign key*

Oke langsung saja pada prakteknya (dalam hal ini saya menggunakan database MySQL) kita memiliki sebuah data sebagai berikut :

Data Buku Terjual CV. Buku Kita Raya

ID Buku	Judul Buku	Jenis Buku	Tahun Buku	Penerbit	Alamat Penerbit
1	Mengubah Wajah Dunia	Filsafat	2015	Mulia Jaya	Jl. Asem Bener 2
2	Belajar Pemrograman PHP 5	IT	2016	Tekno Buku	Jl. Budidaya No 44
3	Ayah	Otobiografi	2016	Mulia Jaya	Jl. Asem Bener 2
4	Belajar MySQL Pemula	IT	2017	Komputek	Jl. Jatiwaringin 52

Data diatas masih berupa bentuk Normal Data, yang tentunya dapat membuat fungsi *DML (Data Manipulation Language)* tidak berjalan sempurna, karena adanya redundansi/duplikasi data. Sehingga misalkan saya ingin mengubah salah satu buku berdasarkan jenis buku IT, maka ada 2 row yang akan terupdate.

Bentuk 1NF (First Normal Form)

Sehingga kita perlu mengubahnya dalam bentuk 1NF (First Normal Form), yakni memetakan data redundansi tersebut menjadi beberapa entitas. Berikut hasil 1NF nya :

Entitas : Buku

ID Buku	Judul Buku	Tahun Buku
1	Mengubah Wajah Dunia	2015
2	Belajar Pemrograman PHP 5	2016
3	Ayah	2016
4	Belajar MySQL Pemula	2017

Entitas : Jenis Buku

ID Jenis Buku	Jenis Buku
JB01	Filsafat
JB02	IT
JB03	Otobiografi

Entitas : Penerbit

ID Penerbit	Penerbit	Alamat Penerbit
PB01	Mulia Jaya	Jl. Asem Bener 2
PB02	Tekno Buku	Jl. Budidaya No 44
PB03	Komputek	Jl. Jatiwaringin 52

Dapat kita lihat data normal form sebelumnya dapat kita pecah menjadi 3 entitas baru beserta atributnya. Dan perlu diperhatikan bahwa setiap entitas memiliki identifier / ID masing, masing sehingga menghindarkan lagi terjadinya redundancy data.

Bentuk 2NF (Second Normal Form)

Jika data sudah bisa dibentuk dalam bentuk 1NF, jika kita perlu melakukan hubungan/relasi antara entitas tersebut, (pada contoh nyata : membuat table transaksi untuk menyimpan order) nah bentuk 2NF ini sudah mengakomodir hal tersebut berdasarkan identifier dari masing-masing entitas. Lihat table 2NF nya sebagai berikut :

Entitas : Transaksi

ID Transaksi	ID Buku	ID Jenis Buku	ID Penerbit
1		1 JB01	PB01
2		2 JB02	PB02
3		3 JB03	PB01
4		4 JB02	PB03

Bentuk 3NF (Third Normal Form)

Bentuk ini adalah pengembangan dari 2NF, Hal ini yakni dengan membuang field/atribut yang tidak bergantung pada identifier / Primary key ID. Itu artinya kita masih bisa menambahkan field jika itu sesuai/berhubungan dengan primary key Misalkan sebagai berikut :

Entitas : Transaksi

ID Transaksi	ID Buku	ID Jenis Buku	ID Penerbit	Jumlah Order	Harga Buku	Total
1	1	JB01	PB01	1	90000	
2	2	JB02	PB02	1	40000	
3	3	JB03	PB01	1	55000	
4	4	JB02	PB03	2	35000	

BAB 4

TEORI HIDDEN MARKOV MODELS (HMM) (*THEORY OF HIDDEN MARKOV MODELS (HMM)*)

A. Discrete-time Markov Processes (*Discrete-Time Markov Processes*)

Rantai Markov waktu diskrit adalah rantai Markov yang memiliki parameter waktu diskrit. Dalam rantai Markov waktu diskrit terdapat beberapa macam kedudukan yang saling berhubungan satu dengan yang lain. Pada skripsi ini dibahas masing-masing kedudukan tersebut dan penerapan rantai Markov waktu diskrit pada bidang Ekonomi. Kedudukan rantai Markov waktu diskrit dibagi beberapa macam yakni kedudukan accessible, kedudukan communicate, kedudukan irreducible, first passage time (kedudukan pertama kalinya x ke y), kedudukan absorbing, kedudukan rekuren dan transient, kedudukan periodik dan aperiodik, kedudukan ergodik, kedudukan steady state. Apabila rantai Markov waktu diskrit mencapai kedudukan steady state maka dapat ditentukan kedudukan jangka panjang dari rantai Markov tersebut. Rantai Markov waktu diskrit dapat diterapkan dalam bidang ekonomi yaitu untuk mengetahui apakah perpindahan konsumen mengalami penurunan atau peningkatan suatu produk bakery pada jangka panjang. Hasil yang diperoleh terdapat kedudukan setimbang dari pangsa pasar produk bakery A, B dan C merupakan pangsa pasar dalam jangka panjang atau perkiraan dari pangsa pasar atau rata-rata pangsa pasar produk

bakery A, B dan C. Pangsa pasar untuk produk bakery A mengalami kenaikan 5,6 %, produk bakery B mengalami kenaikan 11 %, produk bakery C mengalami penurunan 26,7 %.

Markov chain (rantai markov) merupakan salah satu model yang sering digunakan untuk menggambarkan proses-proses stokastik. Istilah Markov chain pertama kali dicetuskan oleh seorang profesor Rusia yakni Prof. Andrei A. Markov (1856-1922). Riset-nyalah yang mendasari studi tentang proses-proses stokastik yang pada perkembangannya mempunyai banyak sekali aplikasi di dunia nyata.

Sebuah proses stokastik dikatakan termasuk Markov chain apabila memenuhi *Markovian property* (sifat Markovian). *Markovian property* menyatakan bahwa probabilitas bersyarat (*conditional probability*) dari sebuah kejadian masa depan, dengan diketahui kejadian masa lampau dan keadaan masa kini, adalah tidak tergantung oleh kejadian masa lampau dan hanya tergantung oleh keadaan masa kini.

Dalam Markov Chain ada tabel probabilitas transisi (transition probability) yang kira-kira memiliki hubungan sebagai berikut:

Keadaan pada saat $t + 1$ (masa depan) = Probabilitas transisi *
Keadaan pada saat t (masa kini)

Markov Chain secara umum dapat digolongkan menjadi dua, yakni Discrete Markov Chain dan Continuous Markov Chain. Markov chain dikatakan diskrit apabila perpindahan keadaan terjadi dengan interval waktu diskrit yang tetap. Sebaliknya, Markov chain dikatakan kontinyu apabila perpindahan keadaan terjadi dengan rentang waktu dengan variabel random yang kontinyu. Pada perkembangannya, ada yang dinamakan Hidden Markov Model (HMM), Higher order Markov Chain dan

Multivariate Markov Chain. Perlu dicatat bahwa model-model pengembangan dari Markov Chain dapat menyalahi Markovian property. keluarga Markov Chain ini cukup memiliki aplikasi yang luas di bidang Teknik Industri. Beberapa contohnya adalah analisa pangsa pasar (termasuk juga pangsa pasar partai politik), sistem re-manufaktur (ada material/part/produk dengan masalah kualitas kemudian diperbaiki dengan masuk kembali ke jalur produksi), klasifikasi pelanggan, peramalan permintaan pelanggan, dan lain-lain.

File template excel tersebut menyediakan dua worksheet. Worksheet pertama dapat digunakan untuk mencari n-steps transition probability. Worksheet kedua dapat digunakan untuk mencari steady state probability. Worksheet pertama hanya menggunakan perkalian matriks biasa, sedangkan worksheet kedua menggunakan algoritma eliminasi Gauss-Jordan. Selain itu, worksheet pertama dapat juga digunakan untuk mencari steady state probability dengan mengeset nilai n yang sangat besar (misalkan 100).

Sebagai contoh, kita memiliki probabilitas transisi sebagai berikut:

No Of States	6	Multiply The Matrix					
No Of Multiplication	3						
		Transition probability from row i to column j					
		1	2	3	4	5	6
	1	0.1	0.26	0.31	0.13	0.11	0.09
	2	0.26	0.31	0.02	0.22	0.11	0.08
	3	0.03	0.19	0.15	0.21	0.11	0.31
	4	0.22	0.23	0.13	0.2	0.09	0.13
	5	0.31	0.1	0.15	0.12	0.11	0.21
	6	0.19	0.15	0.31	0.02	0.22	0.11

Kemudian dengan mengeset No of Multiplication menjadi 3, kita mendapatkan probabilitas transisi 3 langkah sebagai berikut:

Number Of States	6								
Number Of Multiplications	3								
		Transition pro	1	2	3	4	5	6	
	1	0.1804838	0.21821001	0.17159398	0.15699976	0.12279093	0.14992152		
	2	0.18195217	0.21964251	0.16981313	0.15774563	0.12274355	0.14810301		
	3	0.18015003	0.2172925	0.17198072	0.15627741	0.12362902	0.15067032		
	4	0.18156245	0.21852546	0.17106294	0.1567942	0.12322673	0.14882822		
	5	0.18176394	0.21685419	0.17252284	0.15511085	0.12440416	0.14934402		
	6	0.18023801	0.21666682	0.17241139	0.15590944	0.12366697	0.15110737		

Selain itu, dengan worksheet yang kedua, kita dapat mendapatkan steady state probability sebagai berikut:

No Of States	6								
						Solve Steady State			
No	Steady state	Transition probability from row i to column j	1	2	3	4	5	6	
1	0.181036742	1	0.1	0.26	0.31	0.13	0.11	0.09	
2	0.21801639	2	0.26	0.31	0.02	0.22	0.11	0.08	
3	0.171425681	3	0.03	0.19	0.15	0.21	0.11	0.31	
4	0.156610305	4	0.22	0.23	0.13	0.2	0.09	0.13	
5	0.123322514	5	0.31	0.1	0.15	0.12	0.11	0.21	
6	0.149588367	6	0.19	0.15	0.31	0.02	0.22	0.11	

Bisa anda bandingkan steady state probability dengan probabilitas transisi 3 langkah pada gambar nomor dua.

B. Ekstensi dan Jenis pada HMM (*Extensions and Types on HMM*)

Ekstensi (Extensi) file adalah sebuah penanda yang telah ditetapkan sebagai akhiran untuk sebuah nama file yang bisa dioperasikan pada perangkat teknologi seperti komputer atau smartphone. Ekstensi file memperlihatkan karakteristik dari isi file serta tujuan penggunaannya.

Secara umum, ekstensi file mengandung 3 (tiga) karakter. Terdiri dari gabungan antara huruf dan angka yang secara textual mewakili jenis dari file tersebut. Ekstensi file terletak di akhir nama file dengan karakter titik sebagai pemisah keduanya.

Contoh dari ekstensi file tersebut adalah mp3, mp4, xml, php, psd dan lain sebagainya. Perlu kalian ketahui, bahwa file ekstensi ini dapat berjalan dengan baik pada perangkat jika terdapat aplikasi lain yang mendukung format file ekstensi tersebut.

Misalnya file ekstensi .docx hanya bisa kalian buka dengan baik pada software microsoft office word dan jika kalian buka pada aplikasi lain, maka file ini secara otomatis akan menolak untuk membuka data atau sisi dalam file ekstensi tersebut.

Ekstensi file sering dianggap sebagai metadata yang digunakan oleh sistem operasi untuk dapat mengetahui informasi dari file tersebut.



Pentingnya Memahami Extensi File

Apabila ada pengguna komputer yang mengatakan bahwa ekstensi file bukanlah hal penting, mungkin saat itu dia sedang bercanda. Mengapa demikian? Karena sistem komputer dapat diibaratkan sebagai dunia yang luas, flexibel dan penuh perubahan.

Ada begitu banyak file yang berada dalam dunia yang bernama komputer. Baik itu file yang berfungsi sebagai penyusun sistem maupun file-file yang dihasilkan oleh sistem komputer itu sendiri.

Pada kenyataannya, tidak semua dari file-file tersebut memiliki jenis dan karakteristik yang sama. Sama seperti manusia, sistem komputer juga membutuhkan tanda khusus untuk dapat mengenali adanya perbedaan antara jenis file yang satu dengan yang lainnya. Tanda khusus itulah yang selanjutnya kita kenal dengan istilah ekstensi file.

Dengan adanya tanda tersebut sistem komputer dapat dengan cepat mengetahui apa jenis dan karakteristiknya serta bagaimana menangani file tersebut. Itulah mengapa ekstensi file menjadi begitu penting.

Daftar Nama / Jenis Ekstensi (Extensi) File

Extension Type	Description
.h	Header file
.h!	On-line help file (Flambeaux Help! Display Engine)
.h++	Header file (C++)
.h-	Header file (Sphinx C-)
.ha	Compressed file archive created by HA (ha098.zip)
.hap	Compressed file archive created by HAP (hap303re.zip)
.hbk	Handbook (Mathcad)
.hdf	Hierarchical Data File graphics (SDSC Image Tools)
.hdf	Help file (Help Development Kit)
.hdl	Alternate download file listing (Procomm Plus)
.hdr	Pc-File+ Database header
.hdr	Datafile (Egret)
.hdr	Message header text (Procomm Plus – 1st Reader)
.hdw	Vector graphics (Harvard Draw)
.hdx	Help index (AutoCAD – Zortech C++)
.hex	Hex dump

.hfi	Hp Font Info file (GEM)
.hgl	Hp Graphics Language graphics
.hh	C++ header file
.hhh	Precompiled header file (Power C)
.hhp	Help information for remote users (Procomm Plus)
.hin	Molecule (HyperChem)
.hlb	Help library (VAX)
.hlp	Help information
.hmm	Alternate Mail Read option menu (Procomm Plus)
.hnc	CNC program files Heidenhain (?) dialog
.hof	Hall Of Fame (game scores)
.hp8	Ascii text HP Roman8 character set (NewWave Write)
.hpf	Hp LaserJet fonts (PageMaker)
.hpg	HPGL plotter file vector graphics (AutoCad – Harvard Graphics)
.hpi	Font information file (GEM)
.hpj	Help project (MS Help Compiler)
.hpk	Compressed file archive created by HPACK (hpack75.zip)
.hpm	Emm text (HP NewWave)
.hpm	Alternate Main menu for privileged users (Procomm Plus)

.hpp	C++ header file (Zortech C++)
.hqx	Compressed Macintosh ASCII archive created by BINHEX (xbin23.zip)
.hrf	Graphics (Hitachi Raster Format)
.hrm	Alternate Main menu for limited/normal users (Procomm Plus)
.hs2	Monochrome image (Postering)
.hsi	Handmade Software Inc. graphics – almost JPEG (Image Alchemy)
.hst	History file (Procomm Plus)
.htm	HyperText Markup Language document
.htx	Hypertext file
.hwd	Presentation (Hollywood)
.hxm	Alternate Protocol Selection menu for all users (Procomm Plus)
.hxx	C++ header file
.hy1	Hyphenation algorithms (Ventura Publisher)
.hy2	Hyphenation algorithms (Ventura Publisher)
.hyc	Data (WordPerfect)
.hyd	Hyphenation dictionary (WordPerfect for Win)
.hyp	Compressed file archive created by HYPER (hyper25.zip)
.i	Intermediate file (Borland C++ 4.5)

.iax	Bitmap graphics (IBM Image Access eXecutive)
.ibm	Compressed file archive created by ARCHDOS (Internal IBM only)
.ica	Bitmap graphics (Image Object Content Architecture)
.icb	Bitmap graphics
.icl	icon library
.icn	Icon source code file
.ico	Icon (Windows 3.x)
.id	Disk identification file
.ide	Project (Borland C++ 4.5)
.idw	Vector graphics (IntelliDraw)
.idx	Index (many – FoxPro)
.ifd	Form (JetForm Design)
.iff	Interchange File Format bitmap graphics/sound (Amiga)
.iff	Sun TAAC Image File Format (SDSC Image Tool)
.ifp	Script (KnowledgeMan)
.ifs	Fractal image compressed file (Yuvpak)
.ifs	System file (OS/2) hpfs.ifs
.ilb	Data (Scream Tracker)
.ilk	Outline of program's format (MS ILink incremental linker)

.im8	Sun raster graphics
.ima	Mirage vector graphics (EGO, Chart, Autumn)
.img	Bitmap graphics (Ventura Publisher – GEM Paint)
.imp	Spreadsheet (Lotus Improv)
.imq	Image presentation (ImageQ)
.in\$	Installation file (HP NewWave)
.in3	Input device driver (Harvard Graphics 3.0)
.inb	Test script (Vermont HighTest)
.inc	Include file (several programming languages)
.ind	Index (dBASE IV)
.inf	Type 1 LaserJet font information file (soft font installers)
.inf	Information text file (ASCII)
.inf	Install script
.ini	Initialization file
.ink	Pantone reference fills file (CorelDRAW)
.ins	Data (WordPerfect)
.ins	Installation script (1st Reader)
.ins	Instrument music file (Adlib)
.int	Borland Interface Units
.int	Program saved in Internal (semi-compiled) format (Signature)
.inx	Index (Foxbase)

.io	Compressed file archive created by CPIO
.iob	3d graphics database in TDDD format
.ioc	Organizational chart (Instant ORGcharting!)
.ion	4dos descript.ion file (file descriptions)
.ipl	Pantone Spot reference palette file (CorelDRAW)
.irs	Resource (WordPerfect)
.isd	Spelling Checker dictionary (RapidFile)
.ish	Compressed file archive created by ISH
.iso	ISO-9660 table
.it	Settings (intalk)
.itf	Interface file (JPI TopSpeed Pascal)
.iw	Presentation flowchart (IconAuthor – HSC InterActive)
.iwa	Text file (IBM Writing Assistant)
.iwp	Text file (Wang)
.izt	Izl binary token file (IZL)
.jas	Graphics
.jav	Java source code file
.jbd	Datafile (SigmaScan)
.jbx	Project file (Project Scheduler 4)
.jet	Fax (Hybrid JetFax)
.jff	Bitmap graphics (JPEG File Interchange Format)
.jor	Journal file SQL

.jou	Journal backup (VAX Edt editor)
.jpc	Graphics (Japan PIC)
.jpg	Bitmap graphics (Joint Photography Experts Group)
.jtf	Fax (Hayes JT Fax)
.jtf	Bitmap graphics (JPEG Tagged Interchange Format)
.jw	Text document (JustWrite)
.jwl	Library (JustWrite)
.jzz	Spreadsheets (Jazz)
.kar	Midi file with karaoke word track
.kb	Keyboard script (Borland C++ 4.5)
.kb	Program source (Knowledge Pro)
.kbd	Keyboard mapping (LocoScript – Signature – Procomm Plus)
.kbm	Keyboard mapping (Reflection 4.0)
.kcl	Lisp source code (Kyoto Common Lisp)
.kex	Macro (KEDIT)
.key	Datafile (Forecast Pro)
.key	Keyboard macros
.key	Security file eg. Shareware Registration info
.kpp	Toolpad (SmartPad)
.kps	Ibm KIPS bitmap graphics

.kyb	Keyboard mapping (FTP Software PC/TCP)
.l	Lex source code file
.l	Lisp source code file
.l	Linker directive file (WATCOM wlink)
.lab	Datafile (NCSS – SOLO)
.lab	Mailing labels (Q+E for MS Excel)
.lay	Word chart layout (APPLAUSE)
.lbg	Label generator data (dBASE IV)
.lbl	Label (dBASE IV – Clipper 5 – dBFast)
.lbm	Bitmap graphics (DeluxePaint)
.lbm	Linear bitmap graphics (XLib)
.lbo	Compiled label (dBASE IV)
.lbr	Compressed file archive created by LU (lue220.arc)
.lbr	Display driver (Lotus 1-2-3)
.lbt	Label memo (FoxPro)
.lbr	Label (FoxPro)
.lcf	Linker Control File (Norton Guides compiler)
.lck	Lockfile (Paradox)
.lcl	Data (FTP Software PC/TCP)
.lcn	Lecture (WordPerfect)
.lcs	Datafile (ACT! History Files)
.lcw	Spreadsheet (Lucid 3-D)

.ld	Long Distance codes file (Telix)
.ld1	Overlay file (dBASE)
.ldb	Data (MS Access)
.ldf	Library definition file (Geoworks Glue)
.les	Lesson (check *.cvt)
.lev	Level file (NetHack 3.x)
.lex	Lexicon (dictionary) (many)
.lft	Laser printer font (ChiWriter)
.lg	Logo procedure definitions (LSRHS Logo)
.lgo	Logo for header and footer (SuperFax)
.lgo	Startup logo code (Windows 3.x)
.lha	Compressed file archive created by LHA/LHARC (lha255b.exe)
.lhw	Compressed Amiga file archive created by LHWARP
.lib	Library file (several programming languages)
.lif	Logical Interchange Format data file (Hewlett- Packard)
.lif	Compressed file archive
.lim	Compressed file archive created by LIMIT (limit12.zip)
.lin	Line types (AutoCAD)
.lis	Listing (VAX)

.lj	Text file for HP LJ II printer
.ll3	Laplink III related file (document) (LapLink III)
.lnk	Linker response file (.RTLlink)
.lod	Load file
.log	Log file
.lpc	Printer driver (TEKO)
.lrf	Linker response file (MS C/C++)
.lrs	Language Resource File (WordPerfect for Win)
.lsl	Lotus Script Library
.lsp	Lisp source code file (Xlisp)
.lss	Spreadsheet (Legato)
.lst	Keyboard macro (1st Reader)
.lst	List file (archive index – compiler listfile)
.lst	Spool file (Oracle)
.ltm	Form (Lotus Forms)
.lwd	Text document (LotusWorks)
.lwp	Lotus Wordpro 96/97 File
.lwz	MS Linguistically Enhanced Sound File
.lyr	DataCAD Layer File
.lzd	Difference file for binaries (Ldiff 1.20)
.lzh	Compressed file archive created by LHA/LHARC (lha255b.exe)

.lzs	Compressed file archive created by LARC (larc333.zip)
.lzw	Compressed Amiga file archive created by LHWARP
.lzx	Compressed File
.m	Function (program) (Matlab)
.m	Macro module (Brief)
.m	Standard package (Mathematica)
.m11	Text file (MASS11)
.m3	Modula 3 source code file
.m3d	3d animation macro
.m4	M4 preprocessor file (unix)
.m_u	Backup of boot sector, FAT and boot dir (MazeGold)
.ma3	Macro (Harvard Graphics 3.0)
.mac	Bitmap graphics (Macintosh MacPaint)
.mac	Macro
.mag	Woody Lynn's MAG graphics format (MPS Magro Paint System)
.mai	Mail (VAX)
.mak	Makefile
.mak	Project file (Visual Basic)
.man	Command manual

.map	Color palette
.map	Format data (Micrografx Picture Publisher)
.map	Linker map file
.map	Map (Atlas MapMaker)
.map	Network map (AccView)
.mar	Assembly program (VAX Macro)
.mas	Smartmaster set (Freelance Graphics)
.mat	Data file (Matlab)
.max e	Max source code fil
.mb	Memo field values for database (Paradox)
.mbk	Multiple index file backup (dBASE IV)
.mbx	Mailbox (Eudora/Zerberus)
.mcc	Configuration file (Mathcad)
.mcd	Document (Mathcad)
.mcf	Font file (Mathcad)
.mci	Mci command script (Media Control Interface)
.mcp	Application script (Capsule)
.mcp	Printer driver (Mathcad)
.mcw	Text file (MacWrite II)
.md	Compressed file archive created by MDCCD (mdcd10.arc)
.mda	Data (MS Access)
.mdb	Database (MS Access)

.mdl	Model (3D Design Plus)
.mdl	Spreadsheet (CA-Compete!)
.mdm	Modem definition (TELEX)
.mdt	Data table (MS ILink incremental linker)
.mdx	Multiple index file (dBASE IV)
.me	Usually ASCII text file READ.ME
.meb	Macro Editor bottom overflow file (WordPerfect Library)
.med	Macro Editor delete save (WordPerfect Library)
.med	Music (OctaMED)
.mem	Macro Editor macro (WordPerfect Library)
.mem	Memory variable save file (Clipper – dBASE IV – FoxPro)
.meq	Macro Editor print queue file (WordPerfect Library)
.mer	Macro Editor resident area (WordPerfect Library) (value)
.mes	Macro Editor work space file (WordPerfect Library)
.mes	Message
.met	Macro Editor top overflow file (WordPerfect Library)
.met	Document (Omnipage Pro)
.meu	Menu group (DOS Shell)

.mex	Mex file (executable command) (Matlab)
.mex	Macro Editor expound file (WordPerfect Library)
.mf	Metafont text file
.mgf	Font (Micrografx)
.mht	MS MHTML Document
.mib	Snmp MIB file
.mid	Standard MIDI file (music synthetizers)
.mif	Maker Interchange Format (FrameMaker)
.mii	Datafile (MicroStat-II)
.mix	Object file (Power C)
.mk	Makefile
.mke	Makefile (MS Windows SDK)
.mki	Japanese graphics MAKIchan format (MagView 0.5)
.mks	Data (TACT)
.ml3	Project (Milestones 3.x)
.mlb	Macro library file (Symphony)
.mm	Text file (MultiMate Advantage II)
.mmf	Mail message file (MS Mail)
.mmm	Movie (RIFF RMMP format) (MacroMind Director 3.x)
.mmo	Memo writer file (RapidFile)
.mmp	Output video format from Bravado board

.mnd	Menu source (AutoCAD Menu Compiler)
.mng	Map (DeLorme Map'n'Go)
.mnt	Menu memo (FoxPro)
.mnu	Advanced macro (HP NewWave)
.mnu	Menu (AutoCAD Menu Compiler – Norton Commander – Signature)
.mnx	Compiled menu (AutoCAD)
.mnx	Menu (FoxPro)
.mny	Account book (MS Money)
.mob	Device definition (PEN Windows)
.mod	Modula-2 source code file (Clarion Modula-2)
.mod	Windows kernel module
.mod	Music (FastTracker – many)
.mon	Monitor description (ReadMail)
.mov	Movie (AutoCAD AutoFlix)
.mp2	Mpeg audio file (xing)
.mpc	Calender file (MS Project)
.mpg	Mpeg animation
.mpm	Mathplan macro (WordPerfect Library)
.mpp	Project file (MS Project)
.mpr	Generated program (FoxPro)
.mpt	Bitmap graphics (Multipage TIFF)
.mpv	View file (MS Project)

.mpx	Compiled menu program (FoxPro)
.mrb	Multiple Resolution Bitmap graphics (MS C/C++)
.mrs	Macro Resource file (WordPerfect for Win)
.msc	Microsoft C makefile
.msg	Message
.msp	Bitmap graphics (Microsoft Paint)
.mss	Manuscript text file (Perfect Writer – Scribble – MINCE – Jove)
.mst	Minispecification file (Prosa)
.mst	Setup script (MS Windows SDK)
.msw	Text file (MS Word)
.msx	Compressed CP/M file archive created by MSX
.mth	Math file (Derive)
.mtm	Multitracker Module music
.mtw	Datafile (Minitab)
.mu	Menu (Quattro Pro)
.mus	Sound file (MusicTime)
.mvf	Stop frame file (AutoCAD AutoFlix)
.mvi	Movie command file (AutoCAD AutoFlix)
.mvw	Log file (Saber LAN)
.mwf	Animation (ProMotion)
.mxt	Data (MS C)
.myp	Presentation (MM Make Your Point)

.nap	Naplps file (VideoShow) (EnerGraphics)
.nb	Text file (Nota Bene)
.nc	Graphics (netcdf)
.nc	Instructions for NC (Numerical Control) machine (CAMS)
.ncc	Cnc (Computer Numeric Control) control file (CamView 3D)
.ncd	Norton Change Directory support file (Norton Commander)
.ndb	Network database (Intellicom – Compex)
.ndx	Index file (dBASE II – III – IV – dBFast)
.neo	Raster graphics (Atari Neochrome)
.net	Network configuration/info file
.new	New info
.nfo	Info file
.ng	Online documentation database (Norton Guide)
.nlm	Netware Loadable Module
.nlx	Form (FormWorx 3.0)
.np	Project schedule (Nokia Planner) (Visual Planner 3.x)
.npi	Source for DGEN.EXE interpreter (dBASE Application Generator)
.nsf	Lotus Notes / Domino database

.nst	Music (NoiseTracker)
.nt	Startup files (Windows NT)
.ntf	Lotus Notes / Domino template file
.ntr	Executable ASCII text file (strip header and rename) (netrun31.zip)
.nts	Tutorial (Norton)
.nts	Executable ASCII text file (strip header and rename) (netsend1.zip)
.ntx	Index (Clipper 5)
.nuf	Message for new users on their 1st call (Procomm Plus)
.nws	Info text file (latest news) (ASCII)
.nxt	Sound (NeXT format)
.o	Object file (unix – Atari – GCC)
.o\$\$	Outfile (Sprint)
.oaz	Fax (NetFax Manager)
.ob	Object cut/paste file (IBM LinkWay)
.obj	Object code (Intel Recolatable Object Module)
.obr	Object browser data file (Borland C++)
.obs	Script (ObjectScript)
.obv	Visual interface (ObjectScript)
.ocf	Object Craft File (Object Craft)
.ocr	Incoming fax transcribed to text (FAXGrapper)

.ocx	OLE custom control
.odl	Type library source (Visual C++)
.ofd	Form definition (ObjectView)
.off	Object File Format vector graphics
.okt	Music (Oktalizer)
.olb	Object library (VAX)
.old	Backup file
.oli	Text file (Olivetti)
.oom	Swap file (Shroom)
.opn	Active options (Exact)
.opt	Optimize support file (QEMM)
.opw	Organization chart (Org Plus for Windows)
.opx	Inactive options (Exact)
.ora	Parameter file (Oracle)
.org	Calendar file (Lotus Organizer)
.ost	Microsoft Outlook Offline file
.otl	Outline font description (Z-Soft Type Foundry)
.otx	Text file (Olivetti Olitext Plus)
.out	Output file
.ov1	Overlay file (part of program to be loaded when needed)
.ov2	Overlay file (part of program to be loaded when needed)

.ovd	Datafile (ObjectVision)
.ovl	Overlay file (part of program to be loaded when needed)
.ovr	Overlay file (part of program to be loaded when needed)
.p	Pascal source code file
.p	Rea-C-Time application parameter file (ReaGeniX code generator)
.p	Picture file (APPLAUSE)
.p16	Music (16 channels) (ProTracker Studio 16)
.p22	Patch file (Patch22)
.p65	Adobe Pagemaker v6.5
.pa1	Worktable (PageAhead)
.pab	Microsoft Outlook personal address book
.pac	Stad Image (graphics ?)
.pac	Package (SBStudio II)
.pad	Keypad definition (Telemate)
.pak	Compressed file archive created by PAK (pak251.exe)
.pal	Color palette
.pan	Printer-specific file (copy to coreldrw.ink) (CorelDRAW)
.par	Parts application (Digitalk PARTS)

.par	Parameter file (Fractint)
.par	Permanent output file (Windows 3.x)
.pas	Pascal source code file
.pat	Hatch patterns (AutoCAD – Photostyler)
.pat	Vector fill files (CorelDRAW)
.pb	Fax (FAXability Plus)
.pb	Phonebook (WinFax Pro)
.pb	Setup file (PixBase)
.pb1	Document (First Publisher for Windows)
.pba	Powerbasic BASIC source code (Genus)
.pbd	Phone book (FaxNOW! – Faxit)
.pbi	Powerbasic include file (Genus)
.pbi	Profiler Binary Input (MS Source Profiler)
.pbl	Powerbasic library (Genus)
.pbm	Pbm Portable Bit Map graphics
.pbm	Planar bitmap graphics (XLib)
.pbo	Profiler Binary Output (MS Source Profiler)
.pbt	Profiler Binary Table (MS Source Profiler)
.pc	Text file containing IBM PC specific info
.pc3	Custom palette (Harvard Graphics 3.0)
.pc8	Ascii text IBM8 character set (NewWave Write)
.pcc	Cutout picture vector graphics (PC Paintbrush)
.pcd	Graphics (Kodak PhotoCD)

.pcf	Profiler Command File (MS Source Profiler)
.pch	Patch file
.pch	Precompiled header (MS C/C++)
.pcj	Multimedia authoring tool graphics (IBM's Linkaway-Live)
.pck	Pickfile (Turbo Pascal)
.pcl	HP-PCL graphics data (HP Printer Control Language)
.pct	Bitmap graphics (Macintosh b&w PICT1 – color PICT2)
.pcw	Text file (PC Write)
.pcx	Bitmap graphics (PC Paintbrush)
.pda	Bitmap graphics
.pdb	Data (TACT)
.pdd	Adobe PhotoDeluxe Image
.pdf	Package Definition File
.pdf	Graphics file (ED-SCAN 24bit format)
.pdl	Project Description Language file (Borland C++ 4.5)
.pds	Pds graphics
.pds	Pldasm source code file (hardware assembly)
.pdv	Printer driver (Paintbrush)
.pdw	Document (Professional Draw)

.peb	Program Editor bottom overflow file (WordPerfect Library)
.ped	Program Editor delete save (WordPerfect Library)
.pem	Program Editor macro (WordPerfect Library)
.peq	Program Editor print queue file (WordPerfect Library)
.per	Program Editor resident area (WordPerfect Library) (vakioalue)
.pes	Program Editor work space file (WordPerfect Library)
.pet	Program Editor top overflow file (WordPerfect Library)
.pfa	Type 3 font file (unhinted PostScript font)
.pfb	Type 1 font file
.pfc	Text file (First Choice)
.pfk	Programmable function keys (XTreePro)
.pfm	Windows Type 1 font metric file
.pfs	Database (PFS:FILE) – text file (PFS:Write)
.pft	Printer font (ChiWriter)
.pg	Page cut/paste file (IBM LinkWay)
.pgi	Printer Graphics File device driver (PGRAPH library)
.pgm	Portable Grayscale bitMap graphics

.pgm	Program (Signature)
.pgp	Support file (Pretty Good Privacy RSA System)
.pgs	Manual page (man4dos)
.ph	Optimized .goh file (Geoworks)
.ph	Perl header file
.ph	Phrase-table (MS C/C++)
.phn	Phone list (UltraFax – QmodemPro)
.pho	Phone database (Metz Phone for Windows)
.phr	Phrases (LocoScript)
.pic	Pixar picture file (SDSC Image Tool)
.pic	Bitmap graphics (Macintosh b&w PICT1 – color PICT2)
.pic	Bitmap graphics (many eg. Lotus 1-2-3 – PC Paint)
.pif	Program Information File (Windows 3.x)
.pif	Vector graphics GDF format (IBM mainframe computers)
.pit	Compressed Mac file archive created by PACKIT (unpackit.zoo)
.pix	Alias image file (SDSC Image Tool)
.pj	Project (CA-SuperProject)
.pjt	Project memo (FoxPro)
.pjx	Project (FoxPro)

.pk	Packed bitmap font bitmap file (TeX DVI drivers)
.pka	Compressed file archive created by PKARC
.pkg	Installer script (Next)
.pkt	Fidonet packet
.pl	Perl source code file
.pl	Prolog source code file
.pl	Property List font metric file (TeX)
.pl	Palette (Harvard Graphics)
.pl1	Room plan (3D Home Architect)
.pl3	Chart palette (Harvard Graphics 3.0)
.plb	Library (FoxPro)
.plc	Add-in file (functions – macros – applications) (Lotus 1-2-3)
.pll	Pre-linked library (Clipper 5)
.pln	Spreadsheet (WordPerfect for Win)
.plt	Hpgl plotter file vector graphics (AutoCAD)
.plt	Palette
.plt	Pre-linked transfer file (Clipper 5)
.ply	Data (PopMail)
.ply	Presentation screen (Harvard Spotlight)
.pm	Bitmap graphics (Presentation Manager)
.pm3	Document (PageMaker 3)
.pm4	Document (PageMaker 4)

.pm5	Document (PageMaker 5)
.pmc	Graphics (A4TECH Scanner)
.pmm	Program file (Amaris BTX/2)
.pn3	Printer device driver (Harvard Graphics 3.0)
.png	Bitmap graphics (Portable Network Graphics)
.pnm	Pbm Portable aNy Map (PNM) graphics
.pnt	Macintosh painting
.pnt	Qwk reader pointer file (MarkMail 2.x)
.poh	Optimized .goh file (Geoworks)
.pop	Messages index (PopMail)
.pop	Pop-up menu object (dBASE Application Generator)
.pov	Raytraced graphics image (Persistence Of Vision)
.pow	Chord chart (PowerChords)
.pp	Compressed Amiga file archive created by POWERPACKER
.ppb	Button bar for Print Preview (WordPerfect for Win)
.ppl	Polaroidpaletteplus ColorKey device driver (Harvard Graphics 3.0)
.ppm	Portable Pixel Map graphics
.ppo	Pre-processor output (Clipper 5)
.ppp	Publication (PagePlus)
.pps	Storyboard (Personal Producer)

.ppt	General file extension (PowerPoint)
.pr2	Presentation (Aldus Persuasion 2.x)
.pr2	Printer driver (dBASE IV)
.pr3	Postscript printer driver (dBASE IV)
.pr3	Presentation (Aldus Persuasion 3.x)
.prd	Printer driver (many)
.pre	Presentation (Freelance Graphics)
.pre	Settings (Programmer's WorkBench – MS C/C++)
.prf	Pixel Run Format graphics (Improces – Fastgraph)
.prf	Printer driver (dBASE IV)
.prf	Profiler output
.prg	Program (Atari)
.prg	Program source (dBASE IV – FoxPro – Clipper 5 – dBFast)
.pri	Printer definitions (LocoScript)
.prj	Project
.prm	Parameters
.prn	Printer driver (Signature)
.prn	Text file (Lotus 1-2-3 – Symphony)
.pro	Prolog source code file
.pro	Graphics profile file (DOS)
.prs	Printer Resource eg. fonts (WordPerfect for Win)

.prs	Presentation (Harvard Graphics Win)
.prs	Procedure (dBASE IV)
.prt	Printer driver (Dr.Halo)
.prx	Compiled program (FoxPro)
.ps	Postscript file (text/graphics) (ASCII)
.psd	Graphics (Photoshop 3.0)
.pse	Bitmap graphics (IBM printer Page SEgment)
.psf	Outline PostScript printer font (ChiWriter)
.psm	Music (MASI – ProTracker)
.psm	Symbol table of IDE (Turbo Pascal)
.psp	Procedure (Prodea Synergy)
.pst	Microsoft Outlook personal folder
.pt3	Device driver (Harvard Graphics 3.0)
.pt3	Template (PageMaker 3)
.pt4	Template (PageMaker 4)
.ptb	Script (PubTech BatchWorks)
.ptm	Macro (PubTech BatchWorks)
.ptr	Qwk reader pointer file (QMail)
.pub	Page template (MS Publisher)
.pub	Public key ring file (Pretty Good Privacy RSA System)
.pub	Publication (Ventura Publisher – 1st Publisher)

.put	Compressed file archive created by PUT (put334.zip)
.pvd	Script (Instalit)
.pvl	Library (Instalit)
.pvt	Local Fidonet pointlist
.pw	Text file (Professional Write)
.pwl	Password List
.pwp	Text document (Professional WritePlus)
.px	Primary database index (Paradox)
.py	Python script file
.pyc	Compiled PYTHON script file
.pzd	Default settings (Pizazz Plus)
.pzo	Overlay file (Pizazz Plus)
.pzp	Palette (Pizazz Plus)
.pzs	Settings (Pizazz Plus)
.pzt	Transfer file (Pizazz Plus)
.pzx	Swap file (Pizazz Plus)
.qap	Application (Omnis Quartz)
.qbe	Saved query (Query By Example) (dBASE IV – Quattro Pro)
.qbo	Compiled query (dBASE IV)
.qbw	Spreadsheet (QuickBooks for Windows)
.qd0	Data file – segment 10 (Omnis Quartz)

.qd1	Data file – segment 1 (Omnis Quartz)
.qd2	Data file – segment 2 (Omnis Quartz)
.qd3	Data file – segment 3 (Omnis Quartz)
.qd4	Data file – segment 4 (Omnis Quartz)
.qd5	Data file – segment 5 (Omnis Quartz)
.qd6	Data file – segment 6 (Omnis Quartz)
.qd7	Data file – segment 7 (Omnis Quartz)
.qd8	Data file – segment 8 (Omnis Quartz)
.qd9	Data file – segment 9 (Omnis Quartz)
.qdk	Backup of startup files created by Optimize (QEMM)
.qdv	Graphics (Steve Blackstock Giffer)
.qef	Query file (Q+E for MS Excel)
.qfx	Fax (QuickLink)
.qlb	Quick library (MS C/C++)
.qlc	Data (PostScript help file) atmfonts.qlc
.qlp	Printer driver (QuickLink)
.qm4	Options or services file (QMail 4.x Mail Door)
.qpr	Generated query program (FoxPro)
.qpr	Print queue device driver (OS/2)
.qpx	Compiled query program (FoxPro)
.qrs	Equation Editor support file (WordPerfect for Win)
.qrt	Qrt ray tracing graphics

.qry	Query (dBASE IV)
.qt	Quicktime movie (animation)
.qwk	Qwk reader message file
.qxd	Document (QuarkXPress)
.qxl	Element library (QuarkXPress)

C. Masalah Dasar Pada HMM (*Basic Problems With HMM*)

Komputer Tidak Mau Hidup

Cara mengatasinya

- Cek koneksi kabel (dari power outletnya ke tombol power pada PC)
- Cek apakah stabilizer berfungsi atau tidak (jika memakai stabilizer)
- Cek kabel power pada CPU
- Jika masih juga tidak mau hidup permasalahannya mungkin terletak pada power supply atau MB2.

Komputer Mau Hidup Tetapi Tidak Mau Booting

Cara Mengatasinya dengan melihat suara beep yang muncul :

- Beep 1 kali saja Tanda bahwa kondisi komputer baik
- Beep 1 kali, panjang Terdapat problem di memory
- Beep 1 kali panjang dan 3 kali pendek Kerusakan di VGA card
- Beep 1 kali panjang dan 2 kali pendek Kerusakan di DRAM parity

- Beep terus menerus Kerusakan dimodul memory atau memory video
- Cek dengan menggunakan software diagnosa seperti sisoft sandra, PC mark04, PC mark05 dll.

Komputer Mau Booting Tetapi Selalu “Safe Mode “ (untuk masuk ke safe mode tekan F8)

Cara Mengatasinya :

- Restart kembali komputer anda
- Jika masih trouble intall ulang windows anda
- Jika masih safe mode juga, berarti HD anda bermasalah
- Cek dengan : Scan Disk.

Komputer Sering Hang

Cara Mengatasinya :

- Disebabkan software mengalami crash
- Tekan ctrl + alt + del >> klik End task pada program yang “Not Responding”
- Tekan tombol restart pada CPU
- Disebabkan hardware mengalami konflik (adanya penambahan hardware baru)
- Konflik antar hardware sering terjadi pada sistem operasi windows
- Install ulang windows anda, tetapi yang perlu diingat sebelum reinstall windows anda, lepaskan dulu hardware baru anda

- Jalankan fasilitas “add new“ hardware yang terdapat pada control panel.

Keyboard Tidak Dikenali Oleh Komputer

Cara Mengatasinya

- Cek Apakah Keyboard Anda Sudah Terpasang Dengan Benar
- Jika Sudah Tapi Masih Juga Keyboard Tidak Terdeteksi Maka Kemungkinan Keyboard Anda Bermasalah.
- Coba Ganti Keyboard Anda, Jika Sudah Diganti Tapi Juga Masih Bermasalah Maka Kemungkinan Besar Yang Rusak Adalah Di Bagian Port Keyboard Di MB Anda.
- Jika Memang Sudah Di Ganti Keyboard Baru Tapi Tetap Tidak Terdeteksi Juga Coba Ganti Dengan Keyboard USB Dan Apabila Tidak Terdeteksi Juga Berarti Ada Yang Salah Pada Sitem Windows Sobat6 Mouse Tidak Dikenali Oleh Komputer (Sama Denagn Kasus Keyboard).

Pointer Mouse Selalu Meloncat-Loncat

Cara Mengatasinya

- Mouse kotor segera di Bersihkan (khususnya pada bola mouse).

Komputer Sering Crash

Cara Mengatasinya :

- Cek semua posisi kabel, hardware, dan juga tegangan pada casing, cek suhu pada CPU dan juga cek ram, processor dan juga vga.

Bila Produsen MotherBoard(MB) Tidak Diketahui

Cara Mengatasinya :

- Membuka casing, dan cek CPU anda biasanya sebuah MB memiliki label produsen yang sekaligus berisi spesifikasi tipe Mbnya.
- Lihat pada manual book
- Cari data Mb lewat internet, cocokan ID yang tercetak pada sticker board dengan daftar yang terdapat pada situs www.fcc.gov/oet/fccid, dan cari daftar nomor ID yang dikeluarkan oleh lembaga perijinan untuk perangkat elektronik di Amerika
- Gunakan software analisa, seperti sandra99 dll.

Lupa Password BIOS

Cara Mengatasinya :

- Cabut battery cmos pada cpu, kemudian nyalakan kembali
- Atau dengan cara mencoba menebak beberapa password default untuk beberapa produsen bios misalkan AMI dan AWARD (contoh : A.M.I, AMI, AMI_SW, ALLY, 589589 dll)

Jam dan setting tanggal BIOS Selalu Berubah-Rubah

Cara Mengatasinya :

- Baterai cmos sudah tidak berfungsi (mati), ganti dengan batteray yang baru.

Menambah Perangkat Hardware Baru, Tp Tidak Terdeteksi Oleh BIOS

Cara Mengatasinya :

- Kemungkinan besar bios anda sudah kuno sehingga tidak dapat mendeteksi hardware yang baru, maka segera update bios anda (bisa download melalui internet, mis : www.windrivers.com).

Melacak Kerusakan Card Pada MB

Cara Mengatasinya :

- Cobalah Dengan mencabut dan Menancapkan Beberapa Card Pada MB Anda
- Jika Booting Berhasil Maka Card Anda Tidak Bermasalah Begitu Juga Sebaliknya.

Pasang Processor Baru Tp Tidak Terdeteksi

Cara Mengatasinya :

- Cek Apakah Anda Sudah Memasang Processor Denan Benar

- Cek Apakah Posisi Jumper Pada Processor Sudah Benar (Tentang Jumper Pada Processor Bisa Anda Periksa Pada Manual Booknya).

Crash Setelah Memasang RAM Baru

Cara Mengatasinya :

- Kemungkinan ram yang anda pasang tidak kompatibel dengan komputer anda (cabut ram tersebut).

Menambah RAM Tapi Tidak Terdeteksi

Cara Mengatasinya :

- Lakukan pengecekan seperti ketika kasus sebelumnya
- Pastikan slot yan dipakai sesuai, misalnya : SD RAM memiliki slot yang hampir sama dengan RD RAM tetapi RD RAM, tidak bisa terdeteksi meskipun bisa dipasang pada slot jenis SD RAM.

Setelah Menambah RAM Proses Komputer Manjadi Semakin Lambat

Cara Mengatasinya :

- Perhatikan batas kapasitas ram anda, misalnya ram jenis EDO batas maksimalnya adalah 64 MB, maka ketika dipaksakan untuk ditambah maka komputer anda menjadi semakin lambat.

Virtual Ram

Cara Mengatasinya :

- Klik kanan icon My computer, pilih properties, kemudian pilih tab performance dan klik VIRTUAL MEMORY
- Pilih item let me specify my own virtual memory setting (pilih HD yang akan digunakan sebagai virtual memory)
- Klik OK.

Monitor Tidak Mau Nyala

Cara Mengatasinya :

- Pastikan semua kabel power maupun konektor yang berhubungan dengan monitor ok
- Pastikan juga pin yang ada pada port VGA masuk dengan sempurna tidak ada yang bengkok apalagi tidak masuk semua/salah satu pin ke port VGA
- Pastikan juga VGA card anda ok.
- 19. Monitor Menjadi Gelap Saat Loading Windows
- Cara Mengatasinya :
- Kemungkinan disebabkan karena setup driver untuk monitor tidak tepat(setting frekuensinya terlalu tinggi)
- Masuk dulu ke dalam kondisi safe mode (tekan F8)
- Install ulang driver VGAny.

Tampilan Tiba-Tiba Rusak Dan Komputer Manjadi Hang

Cara Mengatasinya :

- Mungkin suhu (pada VGA card) sangat panas.

Ukuran Tampilan monitor Tidak Sesuai Keinginan

Cara Mengatasinya :

- Masuk ke display properties (klik kanan sembarang tempat pilih properties)
- Tekan tab settings dan atur ukuran tampilan sesuai dengan keinginan (pada screen area).

Monitor Seperti Berkedip Saat Digunakan

Cara Mengatasinya :

- Masuk ke display properties (klik kanan sembarang tempat pilih properties)
- Tekan tab settings dan klik advanced, kemudian klik adapter, pada bagian ini ditampilkan refresh rate yang diinginkan

Sound Card Baru Tidak Terdeteksi

Cara Mengatasinya :

- Crash dengan sound card yang lama
- Cek pada manual booknya, apakah soundcard onboardnya perlu dimatikan atau tidak jika hendak menginstall ulang soundcard yang baru (biasanya bisa dimatikan lewat jumper atau bios).

D. Variasi Pada Struktur HMM (*Variation in HMM structure*)

Sistem dibangun dapat mengenali 3 jenis emosi yaitu marah, bahagia dan netral. Fitur yang digunakan dalam sistem ini

adalah pitch, energi dan formant. Database yang digunakan adalah suara dari rekaman film. Dari hasil obeservasi probabilitas emosi marah sebesar 0.196, bahagia 0.254 dan netral 0.045. Sistem memiliki tingkat akurasi rata-rata sebesar 86.66%. Rata waktu eksekusi sistem dalam mendeteksi dan mengklasifikasikan emosi sebesar 21.6ms. Kata kunci: suara, emosi, HMM, klasifikasi

Abstract This research aims to recognize human emotions through voice using HMM algorithm. The system can confirm three types of emotions: anger, happiness and neutrality. The features used in this system are pitch, energy and formant. From the results, the emotional probability of angry is 0.196, happy is 0.254 and neutral is 0.045. Base on testing result, the system has an average accuracy of 86.66% and average execution time of the system in detecting and classifying emotions of 21.6ms. Keywords: voice, emotion, HMM, classification.

BAB 5

SPEECH RECOGNITION

A. Model Kata Terhubung (*Connected word*)

Komunikasi adalah "suatu proses di mana seseorang atau beberapa orang, kelompok, organisasi, dan masyarakat menciptakan, dan menggunakan informasi agar terhubung dengan lingkungan dan orang lain".. Pada umumnya, komunikasi dilakukan secara lisan atau verbal yang dapat dimengerti oleh kedua belah pihak.¹ Apabila tidak ada bahasa verbal yang dapat dimengerti oleh keduanya, komunikasi masih dapat dilakukan dengan menggunakan gerak-gerik tubuh atau menunjukkan sikap tertentu, misalnya tersenyum, menggelengkan kepala, dan mengangkat bahu. Cara seperti ini disebut komunikasi dengan bahasa nonverbal.

Komunikasi adalah suatu proses penyampaian informasi (pesan, ide, gagasan) dari satu pihak kepada pihak lain. Pada umumnya, komunikasi dilakukan secara lisan atau verbal yang dapat dimengerti oleh kedua belah pihak. apabila tidak ada bahasa verbal yang dapat dimengerti oleh keduanya, komunikasi masih dapat dilakukan dengan menggunakan gestur tubuh, menunjukkan sikap tertentu, misalnya tersenyum, menggelengkan kepala, mengangkat bahu. Cara seperti ini disebut komunikasi nonverbal.

Komunikasi atau *communicaton* berasal dari bahasa Latin *communis* yang berarti 'sama'. *Communico*, *communicatio* atau *communicare* yang berarti membuat sama (*make to common*). Secara

sederhana komunikasi dapat terjadi apabila ada kesamaan antara penyampaian pesan dan orang yang menerima pesan. Oleh sebab itu, komunikasi bergantung pada kemampuan kita untuk dapat memahami satu dengan yang lainnya (*communication depends on our ability to understand one another*).

Pada awalnya, komunikasi digunakan untuk mengungkapkan kebutuhan organis. Sinyal-sinyal kimiawi pada organisme awal digunakan untuk reproduksi. Seiring dengan evolusi kehidupan, maka sinyal-sinyal kimiawi primitif yang digunakan dalam berkomunikasi juga ikut berevolusi dan membuka peluang terjadinya perilaku yang lebih rumit seperti tarian kawin pada ikan.

Manusia berkomunikasi untuk membagi pengetahuan dan pengalaman. Bentuk umum komunikasi manusia termasuk bahasa sinyal, bicara, tulisan, gerakan, dan penyiaran. Komunikasi dapat berupa interaktif, komunikasi transaktif|transaktif, komunikasi bertujuan|bertujuan, atau komunikasi tak bertujuan|tak bertujuan.

Melalui komunikasi, sikap dan perasaan seseorang atau sekelompok orang dapat dipahami oleh pihak lain.¹ Akan tetapi, komunikasi hanya akan efektif apabila pesan yang disampaikan dapat ditafsirkan sama oleh penerima pesan tersebut.

Walaupun komunikasi sudah dipelajari sejak lama dan termasuk “barang antik”, topik ini menjadi penting khususnya pada abad 20 karena pertumbuhan komunikasi digambarkan sebagai “penemuan yang revolusioner”, hal ini dikarenakan peningkatan teknologi komunikasi yang pesat seperti radio. Televisi, telepon, satelit dan jaringan komuter seiring dengan industrialisasi bidang usaha yang besar dan politik yang mendunia. Komunikasi dalam tingkat akademi mungkin telah memiliki departemen sendiri di mana komunikasi dibagi-bagi menjadi komunikasi masa, komunikasi bagi pembawa acara,

humas dan lainnya, namun subyeknya akan tetap. Pekerjaan dalam komunikasi mencerminkan keberagaman komunikasi itu sendiri.

Komponen komunikasi adalah hal-hal yang harus ada agar komunikasi bisa berlangsung dengan baik. Menurut Laswell komponen-komponen komunikasi adalah:

- Pengirim atau komunikator (*sender*) adalah pihak yang mengirimkan pesan kepada pihak lain.
- Pesan (*message*) adalah isi atau maksud yang akan disampaikan oleh satu pihak kepada pihak lain.
- Saluran (*channel*) adalah media di mana pesan disampaikan kepada komunikan. dalam komunikasi antar-pribadi (tatap muka) saluran dapat berupa udara yang mengalirkan getaran nada/suara.
- Penerima atau komunikan (*receiver*) adalah pihak yang menerima pesan dari pihak lain
- Umpan balik (*feedback*) adalah tanggapan dari penerimaan pesan atas isi pesan yang disampaikannya.
- Aturan yang disepakati para pelaku komunikasi tentang bagaimana komunikasi itu akan dijalankan ("Protokol")

Komponen komunikasi adalah hal-hal yang harus ada agar komunikasi bisa berlangsung dengan baik. Menurut Laswell komponen-komponen komunikasi adalah:

- Pengirim atau komunikator (*sender*) adalah pihak yang mengirimkan pesan kepada pihak lain.
- Pesan (*message*) adalah isi atau maksud yang akan disampaikan oleh satu pihak kepada pihak lain.

- Saluran (*channel*) adalah media di mana pesan disampaikan kepada komunikan. dalam komunikasi antar-pribadi (tatap muka) saluran dapat berupa udara yang mengalirkan getaran nada/suara.
- Penerima atau komunikan (*receiver*) adalah pihak yang menerima pesan dari pihak lain
- Umpan balik (*feedback*) adalah tanggapan dari penerimaan pesan atas isi pesan yang disampaikannya.
- Aturan yang disepakati para pelaku komunikasi tentang bagaimana komunikasi itu akan dijalankan ("Protokol")

Secara ringkas, proses berlangsungnya komunikasi bisa digambarkan seperti berikut.

1. Komunikator (*sender*) yang mempunyai maksud berkomunikasi dengan orang lain mengirimkan suatu pesan kepada orang yang dimaksud. Pesan yang disampaikan itu bisa berupa informasi dalam bentuk bahasa ataupun lewat simbol-simbol yang bisa dimengerti kedua pihak.
2. Pesan (*message*) itu disampaikan atau dibawa melalui suatu media atau saluran baik secara langsung maupun tidak langsung. Contohnya berbicara langsung melalui telepon, surat, e-mail, atau media lainnya.

media (*channel*) alat yang menjadi penyampai pesan dari komunikator ke komunikan.

1. Komunikan (*receiver*) menerima pesan yang disampaikan dan menerjemahkan isi pesan yang diterimanya ke dalam bahasa yang dimengerti oleh komunikan itu sendiri.
2. Komunikan (*receiver*) memberikan umpan balik (*feedback*) atau tanggapan atas pesan yang dikirimkan kepadanya, apakah dia mengerti atau memahami pesan yang dimaksud oleh si pengirim

Dari berbagai model komunikasi yang sudah ada, di sini akan dibahas tiga model paling utama, serta akan dibicarakan pendekatan yang mendasarinya dan bagaimana komunikasi dikonseptualisasikan dalam perkembangannya

Model komunikasi linear

Model komunikasi ini dikemukakan oleh Claude Shannon dan Warren Weaver pada tahun 1949 dalam buku *The Mathematical of Communication*. Mereka mendeskripsikan komunikasi sebagai proses linear karena tertarik pada teknologi radio dan telepon dan ingin mengembangkan suatu model yang dapat menjelaskan bagaimana informasi melewati berbagai saluran (*channel*). Hasilnya adalah konseptualisasi dari komunikasi linear (*linear communication model*). Pendekatan ini terdiri atas beberapa elemen kunci: sumber (*source*), pesan (*message*) dan penerima (*receiver*). Model linear berasumsi bahwa seseorang hanyalah pengirim atau penerima. Tentu saja hal ini merupakan pandangan yang sangat sempit terhadap partisipan-partisipan dalam proses komunikasi. Suatu konsep penting dalam model ini adalah

gangguan (*noise*), yakni setiap rangsangan tambahan dan tidak dikehendaki yang dapat mengganggu kecermatan pesan yang disampaikan. Gangguan ini selalu ada dalam saluran bersama sebuah pesan yang diterima oleh penerima.

Model interaksional

Model interaksional dikembangkan oleh Wilbur Schramm pada tahun 1954 yang menekankan pada proses komunikasi dua arah di antara para komunikator. Dengan kata lain, komunikasi berlangsung dua arah: dari pengirim dan kepada penerima dan dari penerima kepada pengirim. Proses melingkar ini menunjukkan bahwa komunikasi selalu berlangsung. Para peserta komunikasi menurut model interaksional adalah orang-orang yang mengembangkan potensi manusiawinya melalui interaksi sosial, tepatnya melalui pengambilan peran orang lain. Patut dicatat bahwa model ini menempatkan sumber dan penerima mempunyai kedudukan yang sederajat. Satu elemen yang penting bagi model interkasional adalah **umpan balik** (*feedback*), atau tanggapan terhadap suatu pesan.

Model transaksional

Model komunikasi transaksional dikembangkan oleh Barnlund pada tahun 1970. Model ini menggarisbawahi pengiriman dan penerimaan pesan yang berlangsung secara terus-menerus dalam sebuah episode komunikasi.¹ Komunikasi bersifat transaksional adalah proses kooperatif: pengirim dan

penerima sama-sama bertanggungjawab terhadap dampak dan efektivitas komunikasi yang terjadi. Model transaksional berasumsi bahwa saat kita terus-menerus mengirimkan dan menerima pesan, kita berurusan baik dengan elemen verbal dan nonverbal. Dengan kata lain, peserta komunikasi (komunikator) melakukan proses negosiasi makna

Manfaat Jaringan Komputer :

1. Jaringan komputer membantu seseorang berhubungan dengan orang lain dari berbagai negara dengan mudah.
2. Melalui jaringan komputer, Anda bisa melakukan proses pengiriman data secara cepat dan efisien.
3. Anda dapat mengakses berita atau informasi dengan sangat mudah melalui internet dikarenakan internet merupakan salah satu contoh jaringan komputer.
4. Dengan jaringan komputer, Anda bisa mengakses file yang dimiliki sekaligus file orang lain yang telah disebarluaskan melalui suatu jaringan, semisal internet.



Dua Jenis Jaringan Komputer

Pada dasarnya setiap jaringan komputer ada yang berfungsi sebagai client dan juga server. Tetapi ada jaringan yang memiliki komputer yang khusus didedikasikan sebagai server sedangkan yang lain sebagai client. Ada juga yang tidak memiliki komputer yang khusus berfungsi sebagai server saja. Untuk itu, berdasarkan fungsinya jaringan komputer dibedakan menjadi dua jenis:

1. Client-server

Yaitu jaringan komputer dengan komputer yang didedikasikan khusus sebagai server. Sebuah service/layanan bisa diberikan oleh sebuah komputer atau lebih. Contohnya adalah sebuah domain seperti www.detik.com yang dilayani oleh banyak komputer web server. Atau bisa juga banyak service/layanan yang diberikan oleh satu komputer. Contohnya adalah server jtk.polban.ac.id yang merupakan satu komputer dengan multi service yaitu mail server, web server, file server, database server dan lainnya.

2. Peer-to-peer

Yaitu jaringan komputer dimana setiap host dapat menjadi server dan juga menjadi client secara bersamaan. Contohnya dalam file sharing antar komputer di Jaringan Windows Network Neighbourhood ada 5 komputer (kita beri nama A,B,C,D dan E) yang memberi hak akses terhadap file yang dimilikinya. Pada satu saat A mengakses file share dari B bernama `data_nilai.xls` dan juga memberi akses file `soal_uas.doc` kepada C. Saat A

mengakses file dari B maka A berfungsi sebagai client dan saat A memberi akses file kepada C maka A berfungsi sebagai server. Kedua fungsi itu dilakukan oleh A secara bersamaan maka jaringan seperti ini dinamakan peer to peer.

B. Model Kata Terpisah (*Isolated word*)

Pengenalan ucapan atau pengenalan wicara—dalam istilah bahasa Inggrisnya, *automatic speech recognition* (ASR)—adalah suatu pengembangan teknik dan sistem yang memungkinkan komputer untuk menerima masukan berupa kata yang diucapkan. Teknologi ini memungkinkan suatu perangkat untuk mengenali dan memahami kata-kata yang diucapkan dengan cara digitalisasi kata dan mencocokkan sinyal digital tersebut dengan suatu pola tertentu yang tersimpan dalam suatu perangkat. Kata-kata yang diucapkan diubah bentuknya menjadi sinyal digital dengan cara mengubah gelombang suara menjadi sekumpulan angka yang kemudian disesuaikan dengan kode-kode tertentu untuk mengidentifikasi kata-kata tersebut. Hasil dari identifikasi kata yang diucapkan dapat ditampilkan dalam bentuk tulisan atau dapat dibaca oleh perangkat teknologi sebagai sebuah komando untuk melakukan suatu pekerjaan, misalnya penekanan tombol pada telepon genggam yang dilakukan secara otomatis dengan komando suara.

Alat pengenalan ucapan, yang sering disebut dengan *speech recognizer*, membutuhkan sampel kata sebenarnya yang diucapkan dari pengguna. Sampel kata akan didigitalisasi,

disimpan dalam komputer, dan kemudian digunakan sebagai basis data dalam mencocokkan kata yang diucapkan selanjutnya. Sebagian besar alat pengenalan ucapan sifatnya masih tergantung kepada pembicara. Alat ini hanya dapat mengenal kata yang diucapkan dari satu atau dua orang saja dan hanya bisa mengenal kata-kata terpisah, yaitu kata-kata yang dalam penyampaiannya terdapat jeda antar kata. Hanya sebagian kecil dari peralatan yang menggunakan teknologi ini yang sifatnya tidak tergantung pada pembicara. Alat ini sudah dapat mengenal kata yang diucapkan oleh banyak orang dan juga dapat mengenal kata-kata kontinu, atau kata-kata yang dalam penyampaiannya tidak terdapat jeda antar kata.

Pengenalan ucapan dalam perkembangan teknologinya merupakan bagian dari pengenalan suara (proses identifikasi seseorang berdasarkan suaranya). Pengenalan suara sendiri terbagi menjadi dua, yaitu pengenalan pembicara (identifikasi suara berdasarkan orang yang berbicara) dan pengenalan ucapan (identifikasi suara berdasarkan kata yang diucapkan).

Perkembangan alat pengenalan ucapan

Sejak tahun 1940, perusahaan *American Telephone and Telegraph Company* (AT&T) sudah mulai mengembangkan suatu perangkat teknologi yang dapat mengidentifikasi kata yang diucapkan manusia. Sekitar tahun 1960-an, para peneliti dari perusahaan tersebut sudah berhasil membuat suatu perangkat yang dapat mengidentifikasi kata-kata terpisah dan pada tahun

1970-an mereka berhasil membuat perangkat yang dapat mengidentifikasi kata-kata kontinu. Alat pengenalan ucapan kemudian menjadi sangat fungsional sejak tahun 1980-an dan masih dikembangkan dan terus ditingkatkan keefektifannya hingga sekarang.

Jenis-jenis pengenalan ucapan

Berdasarkan kemampuan dalam mengenal kata yang diucapkan, terdapat 5 jenis pengenalan kata, yaitu:

1. Kata-kata yang terisolasi

Proses pengidentifikasian kata yang hanya dapat mengenal kata yang diucapkan jika kata tersebut memiliki jeda waktu pengucapan antar kata

2. Kata-kata yang berhubungan

Proses pengidentifikasian kata yang mirip dengan kata-kata terisolasi, namun membutuhkan jeda waktu pengucapan antar kata yang lebih singkat

3. Kata-kata yang berkelanjutan

Proses pengidentifikasian kata yang sudah lebih maju karena dapat mengenal kata-kata yang diucapkan secara berkesinambungan dengan jeda waktu yang sangat sedikit atau tanpa jeda waktu. Proses pengenalan suara ini sangat rumit karena membutuhkan metode khusus untuk membedakan kata-kata yang diucapkan tanpa jeda waktu. Pengguna perangkat ini dapat mengucapkan kata-kata secara natural

4. Kata-kata spontan

Proses pengidentifikasian kata yang dapat mengenal kata-kata yang diucapkan secara spontan tanpa jeda waktu antar kata

5. Verifikasi atau identifikasi suara

Proses pengidentifikasian kata yang tidak hanya mampu mengenal kata, namun juga mengidentifikasi siapa yang berbicara.

Proses kerja alat pengenalan ucapan

Alat pengenalan ucapan memiliki empat tahapan dalam prosesnya, yaitu:

1. Tahap penerimaan masukan

Masukan berupa kata-kata yang diucapkan lewat pengeras suara.

2. Tahap ekstraksi

Tahap ini adalah tahap penyimpanan masukan yang berupa suara sekaligus pembuatan basis data sebagai pola. Proses ekstraksi dilakukan berdasarkan metode Model Markov Tersembunyi atau *Hidden Markov Model* (HMM), yang merupakan model statistik dari sebuah sistem yang diasumsikan oleh Markov sebagai suatu proses dengan parameter yang tidak diketahui. Tantangan dalam model statistik ini adalah menentukan parameter-parameter tersembunyi dari parameter yang dapat diamati. Parameter-parameter yang telah kita tentukan kemudian

digunakan untuk analisis yang lebih jauh pada proses pengenalan kata yang diucapkan. Berdasarkan HMM, proses pengenalan ucapan secara umum menghasilkan keluaran yang dapat dikarakterisasikan sebagai sinyal. Sinyal dapat bersifat diskrit (karakter dalam abjad) maupun kontinu (pengukuran temperatur, alunan musik). Sinyal dapat pula bersifat stabil (nilai statistiknya tidak berubah terhadap waktu) maupun nonstabil (nilai sinyal berubah-ubah terhadap waktu). Dengan melakukan pemodelan terhadap sinyal secara benar, dapat dilakukan simulasi terhadap masukan dan pelatihan sebanyak mungkin melalui proses simulasi tersebut sehingga model dapat diterapkan dalam sistem prediksi, sistem pengenalan, maupun sistem identifikasi. Secara garis besar model sinyal dapat dikategorikan menjadi dua golongan, yaitu: model deterministik dan model statistikal. Model deterministik menggunakan nilai-nilai properti dari sebuah sinyal seperti: amplitudo, frekuensi, dan fase dari gelombang sinus. Model statistikal menggunakan nilai-nilai statistik dari sebuah sinyal seperti: proses Gaussian, proses Poisson, proses Markov, dan proses Markov Tersembunyi. Suatu model HMM secara umum memiliki unsur-unsur sebagai berikut:

- a. N, yaitu jumlah bagian dalam model. Secara umum bagian tersebut saling terhubung satu dengan yang lain, dan suatu bagian bisa mencapai semua bagian

yang lain, serta sebaliknya (disebut dengan model ergodik). Namun hal tersebut tidak mutlak karena terdapat kondisi lain dimana suatu bagian hanya bisa berputar ke diri sendiri dan berpindah ke satu bagian berikutnya. Hal ini bergantung pada implementasi dari model.

- b. M , yaitu jumlah simbol observasi secara unik pada tiap bagiannya, misalnya: karakter dalam abjad, dimana bagian diartikan sebagai huruf dalam kata.
- c. Probabilita Perpindahan Bagian $\{ \} = ij A a$
- d. Probabilita Simbol Observasi pada bagian $j, \{ \} () = j Bb$
k
- e. Inisial Distribusi Bagian $i p p$

Setelah memberikan nilai $N, M, A, B,$ dan $p,$ maka proses ekstraksi dapat diurutkan. Berikut adalah tahapan ekstraksi pengenalan ucapan berdasarkan HMM:

- 6. Tahap ekstraksi tampilan
Penyaringan sinyal suara dan pengubahan sinyal suara analog ke digital
- 7. Tahap tugas pemodelan
Pembuatan suatu model HMM dari data-data yang berupa sampel ucapan sebuah kata yang sudah berupa data digital
- 8. Tahap sistem pengenalan HMM
- 9. Penemuan parameter-parameter yang dapat merepresentasikan sinyal suara untuk analisis lebih lanjut.

3. Tahap perbandingan

Tahap ini merupakan tahap pencocokan data baru dengan data suara (pencocokan tata bahasa) pada pola. Tahap ini dimulai dengan proses konversi sinyal suara digital hasil dari proses ekstraksi ke dalam bentuk spektrum suara yang akan dianalisis dengan membandingkannya dengan pola suara pada basis data. Sebelumnya, data suara masukan dipilah-pilah dan diproses satu per satu berdasarkan urutannya. Pemilihan ini dilakukan agar proses analisis dapat dilakukan secara paralel. Proses yang pertama kali dilakukan ialah memproses gelombang kontinu spektrum suara ke dalam bentuk diskrit. Langkah berikutnya ialah proses kalkulasi yang dibagi menjadi dua bagian:

1. Transformasi gelombang diskrit menjadi data yang terurut
Gelombang diskrit berbentuk masukan berukuran n yang menjadi objek yang akan dibagi pada proses konversi dengan cara pembagian rincian waktu
2. Menghitung frekuensi pada tiap elemen data yang terurut

Selanjutnya tiap elemen dari data yang terurut tersebut dikonversi ke dalam bentuk bilangan biner. Data biner tersebut nantinya akan dibandingkan dengan pola data suara dan kemudian diterjemahkan sebagai keluaran yang dapat berbentuk tulisan ataupun perintah pada perangkat.

4. Tahap validasi identitas pengguna

Alat pengenalan ucapan yang sudah memiliki sistem verifikasi/identifikasi suara akan melakukan identifikasi orang yang berbicara berdasarkan kata yang diucapkan setelah menerjemahkan suara tersebut menjadi tulisan atau komando.

Perintah Suara

Perintah Suara (komando suara) adalah suatu program pada komputer yang melakukan perintah berdasarkan perintah suara dari pengguna. Contohnya pada aplikasi *Microsoft Voice* yang berbasis bahasa Inggris. Ketika pengguna mengatakan “Mulai kalkulator” dengan intonasi dan tata bahasa yang sesuai, komputer akan segera membuka aplikasi kalkulator. Jika komando suara yang diberikan sesuai dengan daftar perintah yang tersedia, aplikasi akan memastikan komando suara dengan menampilkan tulisan “Apakah Anda meminta saya untuk ‘mulai kalkulator?’”. Untuk melakukan verifikasi, pengguna cukup mengatakan “Lakukan” dan komputer akan langsung beroperasi.

Pendiktean

Pendiktean adalah sebuah proses mendikte yang sekarang ini banyak dimanfaatkan dalam pembuatan laporan atau penelitian. Contohnya pada aplikasi *Microsoft Dictation* yang merupakan aplikasi yang dapat menuliskan apa yang diucapkan oleh pengguna secara otomatis.

Telepon

Pada telepon, teknologi pengenalan ucapan digunakan pada proses penekanan tombol otomatis yang dapat menelepon nomor tujuan dengan komando suara.

Kelebihan alat pengenalan ucapan

Kelebihan dari peralatan yang menggunakan teknologi ini adalah:

1. Cepat

Teknologi ini mempercepat transmisi informasi dan umpan balik dari transmisi tersebut. Contohnya pada komando suara. Hanya dalam selang waktu sekitar satu atau dua detik setelah kita mengkomandokan perintah melalui suara, komputer sudah memberi umpan balik atas komando kita.

2. Mudah digunakan

Kemudahan teknologi ini juga dapat dilihat dalam aplikasi komando suara. Komando yang biasanya kita masukkan ke dalam komputer dengan menggunakan tetikus atau papan ketik kini dapat dengan mudahnya kita lakukan tanpa perangkat keras, yakni dengan komando suara.

Kekurangan alat pengenalan ucapan

Kekurangan dari peralatan yang menggunakan teknologi ini adalah:

1. Rawan terhadap gangguan

Hal ini disebabkan oleh proses sinyal suara yang masih berbasis frekuensi. Ketika sebuah informasi dalam sinyal suara mempunyai komponen frekuensi yang sama banyaknya dengan komponen frekuensi gangguannya, akan sulit untuk memisahkan gangguan dari sinyal suara

2. Jumlah kata yang dapat dikenal terbatas
Hal ini disebabkan pengenalan ucapan bekerja dengan cara mencari kemiripan dengan basis data yang dimiliki.

C. Aplikasi Perekaman Suara (*Voice Recording Application*)

Di era modern saat ini banyak sekali bermunculan aplikasi yang sangat membantu dalam kegiatan sehari-hari kita misalnya aplikasi menulis, game, video, suara. Ada banyak sekali pilihan untuk anda ketika memilih aplikasi yang sedang anda butuhkan. Nah! Berikut ini adalah artikel yang akan membahas **software rekaman suara** yang mumpuni.

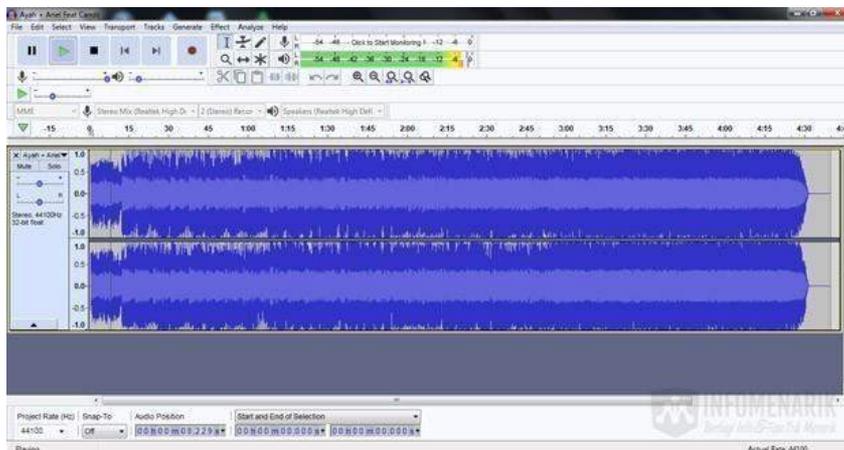
Software Perekam Suara yang Mumpuni

1. Free Studio Recorder



Free Studio Recorder merupakan *software* yang dikembangkan oleh Coolmedia, *software* ini secara khusus memang untuk kebutuhan merekam suara dengan format audio apa saja, menggunakan aplikasi ini anda bisa merekam suara anda dengan beberapa pilihan format audio yang bisa disesuaikan kebutuhan antara lain : MP3, WAV, WMA. Meng-unduh *software* ini dengan mengunjungi website www.sound-recorder.biz.

2. Audacity



Audacity merupakan *software* edit dan rekam suara yang dikembangkan oleh *Audacity Development Team*. *Software* satu ini sangat cocok dengan kebutuhan rekaman dan editing suara anda, anda bisa merekam dan mengedit suara anda dalam *software* ini dengan tampilan yang simple GUI (*Graphic User Interface*). Dan siapa tentunya yang tidak suka dengan yang namanya gratis, tentu senang bukan? *Software* ini bisa dipakai

cukup men-download di website www.audacity.org/download/windows/.

Untuk *voice editor* pemula alangkah baiknya anda menggunakan *software* ini untuk edit dan rekaman suara selain tampilannya bagus, gratis, mudah dioperasikan dan banyak sekali pilihan tools Plug-in yang tentunya akan memberikan kreasi editing sesuai yang anda harapkan. Dan satu lagi, aplikasi ini dapat anda operasikan pada sistem operasi windows, linux ataupun MAC OS.

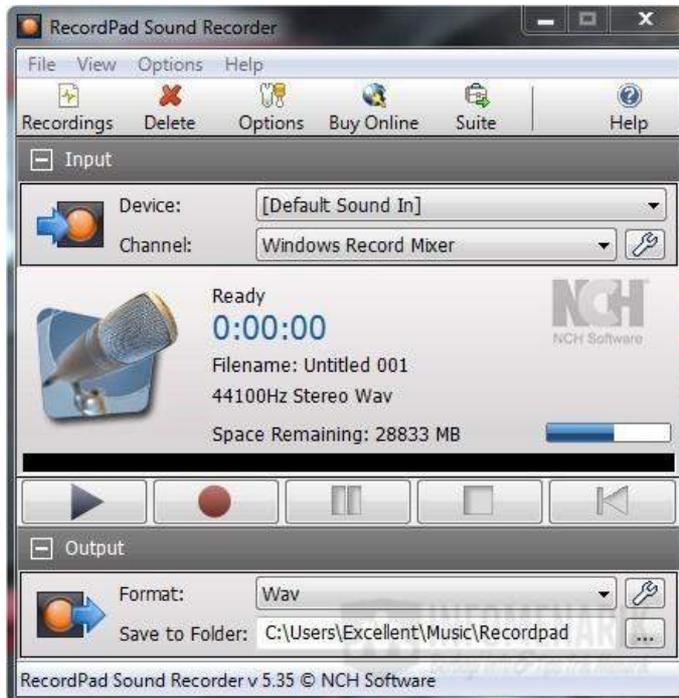
3. Jet Audio Basic



Jet Audio Basic merupakan software dengan berbagai fitur, selain untuk merekam suara, Jet Audio Basic juga mempunyai fitur memainkan video dan *music*, konversi format video, variasi sound effect. Audio CD RIPPING, dan masih banyak lagi. Banyak kalangan yang menyebutkan bahwa software ini bekerja pada dua dunia yakni dunia video dan audio. Seperti yang dijelaskan di atas

fitur dari Jet Audia Basic banyak sekali dan itu mencakup kebutuhan dasar video dan audio, bisa mendapatkan *software* ini di www.jetaudio.com/download/.

4. Recordpad Studio Recording



Free Studio Recorder merupakan *software* yang dikembangkan oleh NCH Software, *software* ini dioperasikan pada *Personal Computer* (PC). *Software* ini akan menghasilkan suara yang jernih tanpa proses editing terlebih dahulu, karena memang *software* ini hanya untuk rekaman saja, *software* ini sering digunakan untuk tes suara musik ataupun untuk merekam

suara presentasi. bisa men-download *software* ini di website www.nchsoftware.com.

5. Audio Recorder For Free



Audio Recorder for Free merupakan *software* yang gratis yang bisa digunakan untuk kebutuhan rekaman suara anda, anda bisa merekam dan menyimpannya dalam format MP3, WAV, dan WMA. Software ini pun sangat mudah digunakan di samping itu juga *software* ini sangat ringan jadi tidak memakan size memori!. Untuk kebutuhan rekaman suara saja *software* ini *recommended*.

6. Free Sound Recorder



Sesuai nama dari *software* ini, aplikasi Free Sound Recorder merupakan aplikasi rekaman suara yang gratis tanpa harus registrasi atau harus *download* tambahan, aplikasi ini berfungsi penuh tanpa batasan. Free Sound Recorder merupakan software yang sempurna untuk digunakan karena *software* ini bisa langsung merekam ataupun melalui input audio external. Anda juga bisa menyimpan hasil rekaman anda dengan berbagai format file seperti MP3, WMA dan WAV yang tentunya bisa nada sesuaikan sendiri.

Software ini juga memungkinkan anda untuk merekam suara dari berbagai sumber lho!, mulai dari Streaming online, Mikrofon, Line-in, stereo, pemutar LP/Kaset/CD, Softphones seperti skype. Pemutar media seperti Real Audia dan Windows Media Player dll. Semua bisa anda rekam dengan menggunakan

software ini. Bisa mendownload software ini di website www.coolrecordedit.com/freestudio recorder/.

7. Hi-Q MP3 Voice Recorder



Aplikasi satu ini sangat cocok buat yang suka merekam suara tanpa harus ribet-ribet dan tentunya tanpa banyak aturan. Aplikasi ini berbasis android. Jadi bisa mendownload di Google Play. Aplikasi ini selain simpel juga bagus dan memiliki kualitas suara yang jernih. bisa men-setting rekaman disesuaikan kebutuhan anda saat rekaman suara.

BAB 6

SPEECH DATABASE (BASIS DATA UCAPAN)

A. Perancangan *Speech Database (Designing A Speech Database)*

Pengukuran waktu kerja merupakan bagian penting dalam proses standarisasi kerja. Secara umum pengukuran kerja dapat dibagi menjadi dua yaitu pengukuran secara langsung dan tidak langsung. Pada pengukuran kerja secara tidak langsung, terdapat beberapa metode pengukuran salah satunya adalah *Methods Time Measurements (MTM)-1*. Untuk pengukuran secara manual, perhitungan dengan *MTM-1* dilakukan dengan memilih daftar operasi yang sesuai dengan operasi kerja amatan. Penggunaan *table MTM* dapat dilakukan oleh pengguna yang telah memahami gerakan kerja dan *MTM* serta memerlukan durasi waktu tertentu dalam proses penetapan waktu standar yang harus tepat dan akurat menggunakan *table MTM*. Pengembangan perangkat lunak untuk mempermudah dan meningkatkan keakuratan penggunaan metode *MTM* merupakan kontribusi utama dalam penelitian ini. Aplikasi perangkat lunak (*SperofiM-1*) ini dirancang untuk seorang observer dalam pengukuran kerja secara tidak langsung. Teknologi *speech recognition* dikembangkan dalam aplikasi ini untuk meningkatkan kecepatan dan keakuratan pemilihan operasi kerja yang bersesuaian pada *table MTM*. Suara observer dibaca sebagai sebuah informasi dan dikonversi dalam spektrum suara yang dianalisa dan dibandingkan dengan *template suara* dan data

waktu sesuai tabel MTM-1 pada database sistem. Dari hasil pengujian antara metode manual, metode click, dan SperofiM-1 diperoleh bahwa SperofiM-1 mampu memberikan performansi waktu terbaik dalam proses pengukuran kerja dengan tabel MTM-1 sebesar 74,14%

B. Sumber-sumber *Speech Database (Sources Of Speech Databases)*

Teknologi speech recognition menemukan akarnya dari pelbagai penemuan sejak abad ke-18. tirto.id - "Hallo, HAL. Apakah kamu mendengarkanku, HAL?" "Afirmatif, Dave. Saya mendengarkanmu." "Tolong bukakan pintu pod, HAL." "Maaf, Dave. Saya takut tidak bisa melakukan itu untukmu." Percakapan di atas terasa alamiah. Masing-masing pihak memahami dan merespon percakapan dengan lugas. Namun, HAL, yang diminta Dave membukakan pintu pod, bukanlah manusia. HAL atau HAL 9000 ialah mesin, mesin pintar yang menemani Dave Bowman dalam perjalanan luar angkasa untuk mencari tahu asal-usul manusia. Keengganan HAL membuka pintu terjadi karena menurutnya akan membahayakan misi mereka. Percakapan ini tak terjadi di dunia nyata, melainkan dari adegan 2001: A Space Odyssey, film garapan Stanley Kubrick yang rilis pada 1968. Dalam dunia fiksi, mesin atau robot yang bisa memahami serta merespons percakapan manusia tidak hanya disajikan Kubrick. Dalam saga Star Wars, ada R2D2 dan C3PO yang memiliki kemampuan itu. Di dunia nyata, percakapan antara mesin dan

manusia kini makin mudah dijumpai, misalnya pada produk asisten digital seperti Google Assistant, Alexa, hingga Siri. Kemampuan dasar utama yang memungkinkan HAL, R2D2, hingga C3PO, serta Google Assistant, Alexa, dan Siri mampu memahami dan merespons percakapan disebut speech recognition.

Teknologi ini dikembangkan sebagai bentuk baru komunikasi antara manusia dengan mesin bernama voice user interface (VUI). Katanya, VUI ialah masa depan bagaimana manusia berinteraksi dengan mesin. Ia lebih sederhana, efisien, dan lebih mudah dibandingkan bentuk interaksi manusia-mesin yang kini populer digunakan, graphical user interface (GUI). Sayangnya, speech recognition sebagai teknologi dasar VUI belum sempurna. Sebagaimana dilansir The Washington Post, Google Assistant, VUI yang tersimpan dalam Google Home, masih sulit memahami dialek dan bahasa selain Inggris. Dalam ujicoba menggunakan 70 perintah suara yang dilakukan manusia pada Google Home, secara keseluruhan Google Assistant memiliki tingkat akurasi hingga 83 persen. Namun, pada akses pada bahasa tertentu nilainya lebih rendah. Akses bahasa Inggris yang dituturkan di Pantai Timur Amerika Serikat, misalnya, memiliki tingkat akurasi 2,5 persen lebih rendah dibandingkan akses bahasa Inggris yang digunakan di wilayah Barat Amerika Serikat.

Sementara itu, tingkat akurasi perintah yang dilakukan menggunakan bahasa Mandarin memiliki tingkat akurasi lebih

rendah 2,6 persen. Dengan penutur utama lebih dari 800 juta jiwa, bahasa Mandarin rupanya memiliki tingkat akurasi yang lebih rendah dibandingkan bahasa Inggris. Bahasa-bahasa yang penuturnya lebih sedikit bernasib lebih buruk. Meski belum sempurna, tulis Richard Lippmann dalam "Speech Recognition by Machine and Humans" (2006), teknologi speech recognition terus mengalami peningkatan kemampuan dengan tingkat kesalahan yang masih bisa ditoleransi. Lippman juga menyebutkan bahwa beberapa sistem speech recognition yang masih diuji, yakni TI, Alphabet, Resource Management, North America Business News, dan Switchboard, menggunakan speech corpus, suatu database percakapan audio. Dari lebih dari 2.000 kosakata yang diumpankan, sistem Switchboard, misalnya, hanya mengalami 80 hingga 150 kebingungan.

Sistem-sistem speech recognition yang ada memiliki tingkat kesalahan yang tak terlalu jauh dibandingkan kemampuan manusia. Manusia, jika diuji dengan speech corpus, memiliki tingkat kesalahan sebesar 0,105 persen. Sementara sistem yang ada rata-rata memiliki tingkat kesalahan sebesar 0,72 persen. Nilai yang cukup bagus itu baru muncul manakala uji dilakukan dalam kondisi "baik", yakni di ruangan tertutup tanpa kebisingan. Kesalahan sistem speech recognition dalam mengenali suara meningkat antara 20 hingga 40 persen jika kondisi ujicoba di luar kondisi "baik" itu. Bermula dari Abad ke-18 Merujuk sejarahnya, sebagaimana ditulis B.H. Juang dalam "Automatic Speech Recognition: A Brief History of the Technology Development"

(2018), penciptaan teknologi speech recognition dimulai setidaknya sejak paruh kedua abad ke-18. Kala itu, ilmuwan Rusia bernama Christian Kratzenstein mencoba membuat mesin yang mampu meniru suara memanfaatkan tabung resonansi. Pada 1881, penemu telepon Alexander Graham Bell dan rekannya membuat Dictaphone, alat perekam suara yang menggunakan silinder berputar. Lantas, pada 1930, Homer Dudley, seorang ilmuwan Amerika Serikat, menemukan teknologi peniru suara (speech synthesizer) bernama VODER (Voice Operating Demonstrator). Infografik Speech Recognition Sistem yang lebih mirip dengan speech recognition hari ini baru muncul pada 1960-an. Kala itu, ilmuwan dari MIT Lincoln Lab menciptakan sistem yang mampu mengenali 10 suku kata yang diucapkan manusia. Teknologi yang lebih maju muncul pada 1992. Kala itu, AT&T merilis layanan bernama Voice Recognition Call Processing (VRCP), teknologi yang mampu mentranskrip percakapan via jaringan AT&T yang digunakan para pelanggannya.

Di masa modern kini, setidaknya sejak 1997, mesin-mesin speech recognition sanggup mengenali 65 ribu suku kata secara simultan. Ketika Lembaga riset militer Amerika Serikat Defense Advanced Research Projects Agency (DARPA) meluncurkan program bernama EARS (Effective, Affordable, and Reusable Speech-to-text). Secara umum, penciptaan sistem speech recognition dilakukan dalam beberapa tahap. Beberapa tahap tersebut, sebagaimana dipaparkan Li Deng dalam penelitiannya yang bertajuk “Challenges in Adopting Speech Recognition”

(2004), adalah mencari ciri khas dari suatu kosakata ketika diucapkan, memahami artikulasi kosakata, hingga mengurutkan kata dalam skema tertentu. Markus Forsberg menyatakan tantangan penyempurnaannya terletak pada fakta bahwa speech recognition adalah bahasa tutur (spoken language) alih-alih bahasa tulis (written language).

Dalam bahasa yang ditulis hanya tercipta komunikasi satu arah. Bahasa tutur mengandaikan komunikasi dua arah. Walhasil, selain harus mampu memahami kosakata yang diucapkan, mesin juga harus mampu merespons lawan bicara secara langsung. Keragaman bahasa dunia berikut dialektanya juga menyulitkan. Mesin harus bisa membedakan "the tail of a dog" (ekor anjing) dengan "the tale of the dog" (kisah anjing), yang bunyinya mirip ketika diucapkan.

C. Implementasi *Speech Database (Implementation Of Speech Databases)*

Di Era digital seperti ini sudah banyak teknologi yang membantu kita dalam melakukan kegiatan. Tapi apakah anda tau kalau terdapat sebuah sistem yang membungkus teknologi tersebut? Yap! Sistem itu disebut **Kecerdasan Buatan**, kecerdasan buatan terbagi menjadi 4, yaitu :

- Sistem pakar
- Robotika dan sistem sensor
- Google playing
- Speech recognize

Nah gais, saya akan membahas salah satu dari hal tersebut yaitu tentang Speech Recognize beserta Implementasi atau penerapan Speech Recognize dalam kehidupan sehari – hari.

Speech Recognize

Dikenal juga dengan pengenalan suara otomatis (automatic speech recognition) atau pengenalan suara komputer (computer speech recognition). Merupakan salah satu fitur antarmuka telematika yang merubah suara menjadi tulisan. Istilah ‘voice recognition’ terkadang digunakan untuk menunjuk ke speech recognition dimana sistem pengenalan dilatih untuk menjadi pembicara istimewa, seperti pada kasus perangkat lunak untuk komputer pribadi, oleh karena itu disana terdapat aspek dari pengenalan pembicara, dimana digunakan untuk mengenali siapa orang yang berbicara, untuk mengenali lebih baik apa yang orang itu bicarakan. Speech recognition merupakan istilah masukan yang berarti dapat mengartikan pembicaraan siapa saja.

Speech recognition atau pengenalan pembicaraan mengkonversi diucapkan kata-kata untuk teks. The “pengenalan suara” istilah kadang-kadang digunakan untuk merujuk kepada sistem pengakuan yang harus dilatih untuk kasus-speaker tertentu seperti untuk perangkat lunak pengenalan yang paling desktop. Menyadari pembicara dapat menyederhanakan tugas menerjemahkan pidato. Pengenalan pembicaraan adalah solusi yang lebih luas yang mengacu pada teknologi yang dapat mengenali pidato tanpa ditargetkan pada pembicara tunggal

seperti sistem call center yang dapat mengenali suara sewenang-wenang.

Aplikasi pengenalan pembicaraan termasuk user interface seperti suara panggilan suara, call routing, kontrol alat domotic, pencarian, sederhana entri data, persiapan dokumen terstruktur, pengolahan pidato-ke-teks, dan pesawat udara.

Secara umum, speech recognizer memproses sinyal suara yang masuk dan menyimpannya dalam bentuk digital. Hasil proses digitalisasi tersebut kemudian dikonversi dalam bentuk spektrum suara yang akan dianalisa dengan membandingkan dengan template suara pada database sistem. Sebelumnya, data suara masukan dipilah-pilah dan diproses satu per satu berdasarkan urutannya. Pemilahan ini dilakukan agar proses analisis dapat dilakukan secara paralel.

Fungsi Automatic Speech Recognition (ASR)

Voice Recognition adalah suatu sistem yang dapat mengidentifikasi seseorang melalui suaranya, pada saat saya mencari pengertian dari voice recognition terdapat juga pengertian Speech Recognition yang hampir sama secara fungsinya. Tapi terdapat perbedaan antara keduanya, Voice Recognition mengidentifikasi siapa yang berbicara, tetapi Speech Recognition mengidentifikasi apa yang diucapkan.

Voice recognition dibagi menjadi dua jenis, yaitu :

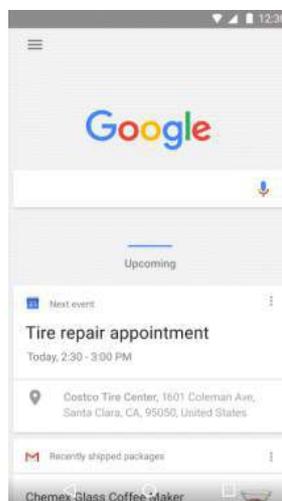
- **Speech recognition** : merupakan proses yang dilakukan computer untuk identifikasi suara yang diucapkan oleh seseorang

tanpa mempedulikan identitas orang terkait. Implementasi speech recognition misalnya perintah suara untuk menjalankan aplikasi komputer. Parameter yang dibandingkan ialah tingkat penekanan suara yang kemudian akan dicocokkan dengan template database yang tersedia.

- **Speaker recognition** : Merupakan sistem pengenalan identitas yang diklaim oleh seseorang dari suaranya atau berdasarkan orang yang berbicara. Misalnya berupa intonasi suara, tingkat kedalaman suara, dan sebagainya. Speech recognition juga dikenal sebagai automatic speech recognition atau computer speech recognition yaitu penerjemah perkataan yang diucapkan menjadi text. Teknologi speech recognition ini sudah ada sejak lama dan sekarang banyak sekali jenis aplikasi yang dikembangkan menggunakan teknologi ini.

Implementasi Speech Recognize

Google Now



Google Now adalah asisten pribadi cerdas yang dikembangkan oleh Google. Google Now tersedia dalam aplikasi bergerak Google Search untuk Android dan iOS, serta penjelajah web Google Chrome di komputer pribadi. Google Now menggunakan bahasa antarmuka pengguna alami untuk menjawab pertanyaan, membuat rekomendasi, dan melakukan tindakan dengan mendelegasikan permintaan untuk satu set layanan web. Seiring dengan menjawab pertanyaan-pengguna dimulai, Google Now secara proaktif memberikan kepada pengguna informasi yang memprediksi (berdasarkan kebiasaan pencarian mereka) yang mereka mungkin inginkan. Ini pertama kali disertakan dalam Android 4.1 (“Jelly Bean”), yang diluncurkan pada tanggal 9 Juli 2012, dan pertama kali didukung pada telepon cerdas Galaxy Nexus. Layanan ini tersedia untuk iOS pada tanggal 29 April 2013 tanpa sebagian besar fitur-fiturnya. Google Now dirilis untuk Google Chrome pada tanggal 24 Maret 2014. Popular Science menyebut Google Now sebagai *“Innovation of the Year”* untuk tahun 2012. Google Now terutama bersaing dengan asisten seperti *Apple Siri* dan *Microsoft Cortana*.

Fungsi Google Now.

Google Now diimplementasikan sebagai aspek dari aplikasi Google Search. Ini mengakui tindakan berulang yang pengguna melakukan pada perangkat (lokasi umum, janji kalender berulang, permintaan pencarian, dll) untuk menampilkan informasi yang lebih relevan kepada pengguna dalam bentuk “kartu”. Sistem ini

memanfaatkan proyek Grafik Pengetahuan Google, sistem yang digunakan untuk merakit hasil pencarian yang lebih rinci dengan menganalisis makna dan hubungan mereka.

Kartu khusus saat ini terdiri:

- Ringkasan aktivitas (berjalan/bersepeda)
- Pas naik
- Konser
- Mata uang
- Mengembangkan cerita dan berita terkini
- Acara
- Peningat acara
- Penerbangan
- Ulang tahun teman
- Hotel
- Peningat lokasi
- Film
- Tempat wisata
- Acara sekitar
- Bintik foto sekitar
- Album/buku/permainan video/episode TV baru
- Topik berita
- Pertemuan berikutnya
- Paket
- Lokasi parkir
- Tempat
- Daftar produk
- Peringatan publik
- Angkutan umum
- Topik penelitian
- Pemesanan restoran
- Olahraga
- Saham
- Waktu ke rumah
- Peningat waktu
- Lalu lintas dan transit
- Terjemahan
- Cuaca
- Pemutakhiran situs web
- Apa yang ingin ditonton
- Ulang tahun Anda

Google mengumumkan bahwa mereka akan memulai dukungan untuk kartu dari aplikasi pihak ketiga. Ini termasuk Airbnb, Duolingo, Fandango, Lyft, Pandora, The Guardian, dan Zillow, dan lainnya.

Aplikasi Google Now terdapat fitur antarmuka telematika yang merubah suara menjadi tulisan. Google Now dalam penggunaannya cukup mudah, kita hanya mengatakan “Ok, Google” maka akan muncul kotak *voice search*, kita tinggal ucapkan apa yang mau kita cari, maka si Google Now yang cerdas

ini akan menampilkan hasil yang dia dapatkan dari internet ke layar *smartphone*. Tidak hanya itu Google Now dapat menjadi pengingat saat kita lupa layaknya seorang asisten sungguhan, seperti mengingatkan untuk pergi ke kantor dan secara otomatis memberi tahu rute menuju kantor, kelebihan lainnya adalah apabila ada kecelakaan disekitar kita Google Now akan memberitahunya melalui notifikasi.

BAB 7

TEXT TO SPEECH (TTS), SERTA MASALAH TEXT TO SPEECH (TTS)

(TEXT TO SPEECH (TTS), AND THE PROBLEM TEXT TO SPEECH (TTS))

A. Definisi TTS (*Definition of TTS*)

TTSpedia merupakan situs terlengkap dan terbaik untuk memecahkan teka teki silang.

- TTSpedia memuat lebih dari 61.688 data pertanyaan dan jawaban TTS.
- Proses pencarian yang sangat cepat.
- Dapat dipergunakan secara GRATIS.

Data dari berbagai sumber seperti: majalah, koran-koran populer (Jawa Pos, Kompas, Tempo, Pikiran Rakyat dll), kamus arti kata, dan kamus thesaurus. Data-data dari berbagai sumber tersebut kemudian diolah sedemikian rupa sehingga dapat menghasilkan jawaban yang tepat untuk setiap pertanyaan TTS.

B. Fonologi, Morphologi, Sintaks, Semantik, dan Pragmatik (*Phonology, Morphology, Syntax, Semantics, and Pragmatics*)

1. Fonologi

Fonologi menurut bahasa berasal dari kata fon (bunyi) dan logi (ilmu). Menurut istilah fonologi adalah bagian dari ilmu bahasa yang mempelajari bunyi bahasa secara umum. Fonologi terbagi

menjadi 2 yaitu Fonetik dan Fonemik. Adapun contoh fonologi yaitu :

- a. Kata pola yang membedakan /o/ dan /u/ pula
- b. Kata panci yang membedakan /c/ dan /t/ panti
- c. Kata pegal yang membedakan /g/ dan /d/ pedal
- d. Diftong ai yaitu pantai, santai, petai, lantai
- e. Diftong au yaitu imbau, pulau, surau, gurau

2. Morfologi

Morfologi menurut bahasa berasal dari kata morph (bentuk) + o (pembentuk stem) dan logos (ilmu). Morfologi adalah bagian dari ilmu bahasa yang mempelajari seluk beluk bentuk kata, perubahan bentuk kata, serta perubahan bentuk kata terhadap arti dan golongan kata. Contoh morfologi :

- a. Dari kata Jalan, dijalankan, menjalankan, berjalan, dijalani, perjalanan
- b. Kata Ragu, diragukan, meragukan
- c. Kata makan, dimakan, memakan
- d. Dari kata lempar, dilempar, dilempari, dilemparkan, melemparkan, melempari, pelemparan
- e. Kata baca, dibaca, membaca, dibacakan, membacakan, pembacaan

3. Sintaksis

Sintaksis berasal dari bahasa belanda yaitu syntaxis. Sintaksis adalah bagian dari ilmu bahasa yang mempelajari

prinsip dan peraturan dalam membuat kalimat. Sintaksis tersusun dari Subjek, Predikat, Objek dan Keterangan. Contoh sintaksis

a. Seorang penjual sedang menjual di pasar

S P O

b. Para siswa sedang belajar di kelas

S P O

c. Ibu pergi ke pasar membeli beras

S P O K

d. Adik berangkat ke sekolah pada pukul 7

S P O K

e. Saya berjalan menuju Kampus

S P O

4. Semantik

Semantik berasal dari bahasa Yunani yaitu semantikos berarti memberikan tanda, penting. Semantik adalah cabang dari ilmu bahasa yang mempelajari makna yang terkandung dalam suatu bahasa. Contoh semantik:

a. Cakar ayam yang berarti orang yang tulisannya jelek

b. Otak Udang bermakna orang yang bodoh. Contohnya, Rani akan belajar giat karena tak mau dipanggil otak udang oleh teman-temannya

c. Banting tulang yang bermakna kerja keras. Contohnya, Bapak banting tulang setiap hari demi keluarganya

d. Darah biru artinya orang keturunan bangsawan

- e. Buah Bibir artinya bahan pembicaraan. Contohnya, Dia selalu menjadi buah bibir para mahasiswa di kampusnya

5. Pragmatik

Pragmatik adalah ilmu bahasa yang mempelajari hubungan antara konteks dan makna. Pragmatik mengkaji kondisi-kondisi penggunaan bahasa manusia yang ditentukan oleh konteks kemasyarakatan. Contoh pragmatik yaitu

- a. "Saya izin ke belakang" . Secara pragmatik izin ke belakang bermakna saya izin ke toilet, tapi kata "toilet" sepertinya kurang sopan didengar jadi digunakanlah bahasa itu.
- b. "wajahnya biasa saja". Secara pragmatik berarti wajah orang itu tidak cantik.
- c. "Ia memiliki kemampuan berpikir di bawah rata-rata". Artinya orang tersebut tidak pintar/ bodoh
- d. "Ada oknum yang tidak bertanggung jawab dibalik peristiwa tersebut". Artinya, oknum yg dimaksud ini adalah orang yang dalam konteksnya 'kurang baik'.
- e. si A berhasil mempersunting wanita yang dicintainya

C. Ucapan Dan Tulisan, Organisasi Sistem TTS, Sistem Tts Yang Sederhana, Konsep Ucapan (*Speeches And Writings, The Organization Of The Tts System, The Tts System Is A Simple, Conceptual Greeting*)

TTS (Text to Speech) adalah suatu sistem yang dapat melakukan konversi dari teks menjadi ucapan. Saat ini terdapat

berbagai sistem yang dapat melakukan fungsi seperti itu, misalnya sistem IVR (interactive voice response) yang banyak digunakan untuk layanan informasi otomatis melalui telpon (call center). Sistem IVR biasanya menggunakan rekaman kata atau kalimat yang direkam secara utuh.

Suatu sistem TTS dapat didefinisikan sebagai “Sistem yang dapat mengubah suatu teks menjadi ucapan secara otomatis dengan cara fonetisasi (penyusunan fonem-fonem untuk membentuk ucapan)”.

Teknik fonem to speech:

1. Merecord potongan-potongan suara manusia
Suara seseorang direkam, kemudian untuk menghasilkan suatu ucapan potongan suara tersebut digabungkan.
2. Recording Diphone
Perekaman menggunakan 2 fonem
3. Learnings
Melalui pembelajaran
4. Artikulatory
Persamaan matematik yang merepresentasikan gelombang fisika

Speech Recognition

Speech Recognition adalah proses identifikasi suara berdasarkan kata yang diucapkan dengan melakukan konversi sebuah sinyal akustik, yang ditangkap oleh audio device (perangkat input suara).

Speech Recognition juga merupakan sistem yang digunakan untuk mengenali perintah kata dari suara manusia dan kemudian diterjemahkan menjadi suatu data yang dimengerti oleh komputer. Pada saat ini, sistem ini digunakan untuk menggantikan peranan input dari keyboard dan mouse.

Terdapat 4 langkah utama dalam sistem pengenalan suara:

- Penerimaan data input
- Ekstraksi, yaitu penyimpanan data masukan sekaligus pembuatan database untuk template.
- Perbandingan / pencocokan, yaitu tahap pencocokan data baru dengan data suara (pencocokan tata bahasa) pada template.
- Validasi identitas pengguna.

D. Kendala dalam TTS (*Constraints in TTS*)

Speech recognition, ya baru kali ini saya mendengar software semacam ini. software ini merupakan perintah suara pengganti keyboard atau mouse. Perintah suara yang kita perintahkan ke komputer dapat diterjemahkan oleh komputer kemudian akan disajikan dalam bentuk output.

Penggunaan speech recognition ini dapat mempersingkat waktu, namun input suara yang kita berikan harus jelas, setidaknya berbahasa Inggris.

Hambatan yang ditemui dalam menggunakan Speech Recognition Software antara lain :

1. Pengucapan kalimat atau kata dalam bahasa inggris harus benar (pronunciation).
2. Suara harus keras dan jelas (loud).
3. Pelafalan juga harus jelas (speelling).
4. Bila suara tidak jelas maka perintah yang dijalankan komputer tidak benar/salah(eror).
5. Begitu juga dengan pengucapannya jika salah dan tidak jelas, komputer tidak melaksanakan instruksi yang diberikan.

Keuntungan yang paling menguntungkan menurut saya adalah masalah masalah yang kita hadapi dalam berkomputer dapat kita tanyakan menggunakan bahasa kita dengan software ini. Misalnya "how can i change my desktop?". Kemudian speech recognition akan mengirimkan perintah ini untuk diterjemahkan sebagai output.

Ikuti langkah berikut ini :

1. Tempatkan mikrofon sehingga sekitar satu inci atau lebar jempol ke sisi mulut. Pastikan bahwa itu tidak secara langsung di depan mulut, dan tidak bernapas langsung ke dalamnya. CATATAN: Jika secara tidak sengaja memindahkan mikrofon saat berbicara, ingatlah untuk membawa kembali ke posisi yang benar. Jalankan program di mana ingin menggunakan pengenalan suara, dan kemudian klik dokumen untuk menempatkan titik penyisipan ke dalam dokumen. Jika membuka topik Bantuan sementara bekerja atau jika pesan

ditampilkan pada layar, klik lagi dokumen untuk terus menggunakan pengenalan suara. Pada bar Bahasa, klik Mikrofon (jika mikrofon belum dihidupkan).

Catatan: Bar menampilkan Bahasa label masing-masing di samping tombol pada panel secara default. Untuk menyembunyikan atau menampilkan label teks, klik kanan bar Bahasa, dan kemudian klik Text Labels. Beralih antara Dikte dan Voice Command mode saat Anda bekerja.

Catatan: Dapat menghemat waktu jika Anda menyelesaikan dikte pertama, review file, dan kemudian memformat teks atau membuat koreksi. Bila melakukannya, harus beralih antara modus Dikte dan Voice modus Command kurang sering. Untuk mengubah mode:

Menggunakan modus Dikte: Untuk mengaktifkan kata yang berbicara ke dalam teks, klik Dikte pada bar Bahasa.

2. Ketika berbicara, bar biru ditampilkan, ini berarti bahwa komputer adalah pengolahan suara. Seperti kata-kata diakui, teks akan ditampilkan pada layar. dapat terus berbicara sementara komputer proses suara, tidak harus menunggu sampai bar biru menghilang untuk berbicara lagi. Catatan: Sementara bar biru ditampilkan, hindari menggunakan mouse atau keyboard untuk mengetik atau mengambil tindakan lain. Jika melakukannya, pengenalan suara terganggu, dan kata-kata tidak diproses.

Menggunakan Voice Command Mode: Untuk memilih menu, toolbar, kotak dialog (US hanya dalam bahasa Inggris), dan task pane (hanya dalam bahasa Inggris AS) item, di bar Bahasa, klik Voice Command. Misalnya, untuk mengubah format font, bisa mengatakan “font” atau “font” untuk membuka kotak Font pada toolbar Formatting, dan kemudian mengatakan nama font. Atau jika ingin memformat teks yang dipilih, mengatakan “berani” atau “garis bawah.” Mikrofon Klik pada bar Bahasa untuk menghidupkan mikrofon mati saat selesai berbicara dengan komputer.

BAB 8

TEKNIK MENGUBAH DARI TEKS KE FONEM

(TECHNIQUES TO CHANGE FROM TEXT TO FONEM)

A. Grammars, Inference, Parsing, and Transduction *(Grammars, Inference, Parsing, and Transduction)*

Dalam ilmu komputer dan linguistik , parsing, atau lebih formal disebut analisis sintaksis, adalah proses menganalisis teks, terbuat dari urutan token (misalnya, kata-kata), untuk menentukan struktur gramatikal terhadap hal yang diberikan (lebih atau kurang) tata bahasa formal . Parsing juga merupakan istilah awal untuk diagram kalimat dari bahasa alam , dan masih digunakan untuk diagram dari infleksi bahasa, seperti bahasa Romance atau Latin . Dalam komputasi, parser adalah salah satu komponen dalam sebuah interpreter atau kompiler , yang memeriksa sintaks yang benar dan membangun struktur data (sering beberapa jenis pohon parse , pohon sintaks abstrak atau struktur hirarkis lainnya) tersirat dalam token masukan. Program pendeteksi kombinasi tombol sering menggunakan yang terpisah penganalisa leksikal untuk membuat token dari urutan karakter masukan. Parser dapat diprogram dengan tangan atau mungkin (semi-) otomatis dihasilkan (dalam beberapa bahasa pemrograman) dengan alat (seperti Yacc) dari tata bahasa yang ditulis dalam bentuk Backus-Naur . Dalam beberapa terjemahan mesin dan bahasa alami pengolahan sistem, bahasa manusia parsing oleh program komputer.kalimat Manusia tidak mudah diurai oleh program,

karena ada substansial ambiguitas di struktur bahasa manusia, yang penggunaannya untuk menyampaikan makna (atau semantik) di antara rentang terbatas berpotensi kemungkinan tetapi hanya beberapa yang erat dengan kasus tertentu. Untuk mengurai data bahasa alami, peneliti pertama harus setuju pada tata bahasa untuk digunakan.

Pemilihan sintaks dipengaruhi oleh linguistik dan komputasi perhatian, misalnya beberapa sistem parsing menggunakan tata bahasa fungsional leksikal , tetapi secara umum, parsing untuk tata bahasa jenis ini dikenal sebagai NP-lengkap . Kepala berbasis frase struktur tata bahasa linguistik lain formalisme yang telah populer di masyarakat parsing, namun upaya penelitian lain telah terfokus pada formalisms kurang kompleks seperti yang digunakan dalam Penn Treebank . parsing Dangkal hanya bertujuan untuk menemukan batas dari unsur utama seperti frasa nominal. Strategi lain yang populer untuk menghindari kontroversi linguistik adalah ketergantungan tata bahasa parsing.

Sebagian parser modern setidaknya sebagian statistik , yaitu, mereka bergantung pada korpus data pelatihan yang telah dijelaskan (parsing dengan tangan). Pendekatan ini memungkinkan sistem untuk mengumpulkan informasi tentang frekuensi yang berbagai konstruksi terjadi dalam konteks tertentu Lihat. (mesin belajar) Pendekatan yang telah digunakan termasuk mudah PCFGs (probabilistik tata bahasa bebas konteks), entropi maksimum , dan jaringan syaraf . Sebagian besar sistem yang

lebih sukses menggunakan statistik leksikal (yaitu, mereka mempertimbangkan identitas kata-kata yang terlibat, serta mereka bagian dari pidato). Namun sistem ini rentan terhadap overfitting dan membutuhkan beberapa jenis smoothing efektif. Parsing algoritma untuk bahasa alam tidak bisa mengandalkan pada tata bahasa memiliki 'baik' properti sebagai dengan tata bahasa secara manual dirancang untuk bahasa pemrograman. Seperti yang disebutkan sebelumnya beberapa formalisms tata bahasa sangat sulit untuk mengurai komputasi, secara umum, bahkan jika struktur yang diinginkan tidak bebas konteks, semacam aproksimasi bebas konteks untuk tata bahasa digunakan untuk melakukan lulus pertama. Algoritma yang menggunakan tata bahasa bebas konteks seringkali bergantung pada beberapa varian dari algoritma CKY, biasanya dengan beberapa heuristik untuk memangkas pergi analisis tidak mungkin untuk menghemat waktu Lihat. (parsing grafik .) Namun beberapa sistem perdagangan kecepatan untuk akurasi menggunakan, misalnya, linear-time versi mengurangi pergeseran- algoritma. Sebuah perkembangan yang agak baru-baru ini telah reranking mengurai di mana parser mengusulkan beberapa sejumlah besar analisis, dan kompleks sistem yang lebih memilih pilihan terbaik.

Bahasa Pemrograman Yang umum menggunakan sebagian besar parser adalah sebagai komponen dari suatu compiler atau interpreter . Hal ini mem-parsing kode sumber dari bahasa pemrograman komputer untuk membuat beberapa bentuk representasi internal. Bahasa Pemrograman cenderung akan

ditentukan dalam istilah dari tata bahasa bebas konteks karena cepat dan efisien parser dapat ditulis untuk mereka . Parser yang ditulis oleh tangan atau yang dihasilkan oleh generator parser . Konteks tata bahasa bebas terbatas sejauh mana mereka bisa mengekspresikan semua persyaratan dari sebuah bahasa.. Informal, alasannya adalah bahwa memori seperti bahasa terbatas. tata bahasa tidak dapat mengingat keberadaan membangun atas masukan sewenang-wenang panjang; ini diperlukan suatu bahasa yang, misalnya, nama harus dinyatakan sebelum dapat dirujuk. tata bahasa kuat lainnya yang dapat mengungkapkan kendala ini, bagaimanapun, tidak bisa diurai efisien..

Oleh karena itu, strategi umum untuk membuat parser santai untuk tata bahasa bebas konteks yang menerima superset dari bahasa yang dikehendaki konstruksi (yaitu, ia menerima beberapa konstruksi tidak valid), kemudian, konstruksi yang tidak diinginkan dapat disaring. Sekilas proses Contoh berikut menunjukkan kasus umum dari penguraian sebuah bahasa komputer dengan dua tingkat tata bahasa: leksikal dan sintaksis. Tahap pertama adalah generasi token, atau analisis leksikal , di mana aliran input karakter dibagi menjadi simbol bermakna yang didefinisikan oleh tata bahasa dari kalimat biasa . Sebagai contoh, sebuah program kalkulator akan melihat masukan seperti " $12*(3+4)^2$ " dan membagi ke dalam token 12 , $*$, $($, 3 , $+$, 4 , $)$, $^$ dan 2 , masing-masing yang merupakan sebuah simbol yang bermakna dalam konteks ekspresi aritmatika. lexer itu akan berisi

aturan untuk mengatakan bahwa karakter $*$, $+$, $^$, $($ dan $)$ menandai awal dari sebuah token baru, berarti tanda jadi seperti " 12* "atau" (3 "tidak akan dihasilkan. Tahap berikutnya adalah parsing atau analisa sintaksis , yang memeriksa bahwa bentuk ekspresi token yang diijinkan. Hal ini biasanya dilakukan dengan mengacu pada tata bahasa bebas konteks yang secara rekursif mendefinisikan komponen yang dapat membuat ekspresi dan urutan di mana mereka harus muncul. Namun, tidak semua aturan menentukan bahasa pemrograman dapat dinyatakan dengan tata bahasa bebas konteks saja, untuk jenis validitas dan deklarasi yang tepat contoh pengenalan.

Aturan-aturan ini dapat dinyatakan secara resmi dengan tata bahasa atribut . Tahap akhir adalah semantik parsing atau analisis, yang bekerja di luar implikasi dari ungkapan hanya disahkan dan mengambil tindakan yang tepat. Dalam kasus kalkulator atau interpreter, tindakan ini adalah untuk mengevaluasi ekspresi atau program, compiler, di sisi lain, akan menghasilkan semacam kode. Atribut tata bahasa juga dapat digunakan untuk mendefinisikan tindakan tersebut. Jenis dari parser Tugas parser pada dasarnya adalah untuk menentukan apakah dan bagaimana input dapat diturunkan dari simbol awal tata bahasa. Hal ini dapat dilakukan pada dasarnya dua cara: - Top-down parsing - Top-down parsing dapat dilihat sebagai upaya untuk menemukan kiri-derivasi sebagian besar aliran-masukan dengan mencari pohon parse menggunakan-down ekspansi yang diberikan atas tata bahasa formal aturan. Inklusif pilihan

digunakan untuk mengakomodasi ambiguitas dengan memperluas semua alternatif tangan kanan-sisi aturan tata bahasa.

Bottom-up parsing - parser A dapat memulai dengan masukan dan berusaha untuk menulis ulang ke simbol awal. Intuitif, parser upaya untuk menemukan elemen paling dasar, maka unsur-unsur yang mengandung, dan seterusnya. parser LR adalah contoh-up parser bawah. Istilah lain yang digunakan untuk jenis parser ini Shift-Mengurangi parsing. Parser LL dan keturunan-parser recursive adalah contoh dari top-down parser yang tidak dapat mengakomodasi rekursif kiri produksi. Meskipun telah percaya bahwa implementasi sederhana-down parsing atas tidak dapat mengakomodasi langsung dan tidak langsung kiri rekursi dan mungkin memerlukan waktu eksponensial dan kompleksitas ruang sementara parsing ambigu -bebas tata bahasa konteks, canggih algoritma lebih untuk-down parsing atas telah diciptakan oleh Frost, Hafiz, dan Callaghan yang mengakomodasi ambiguitas dan rekursi kiri dalam waktu polinomial dan yang menghasilkan ukuran representasi polinom-eksponensial dari jumlah potensial pohon parse. algoritma mereka mampu menghasilkan paling baik kiri-dan-yang paling derivasi kanan input sehubungan dengan diberikan CFG. Sebuah perbedaan penting sehubungan dengan parser adalah apakah sebuah parser menghasilkan derivasi paling kiri atau paling kanan derivasi (lihat tata bahasa bebas konteks). LL parser akan menghasilkan paling kiri derivasi dan parser LR akan

menghasilkan paling kanan derivasi (walaupun biasanya secara terbalik).

B. Arsitektur Pemrosesan Bahasa Alamiah Untuk *TTS Synthesis* (*Architecture processes natural languages for TTS Synthesis*)

Natural language processing (NLP) atau dalam bahasa Indonesia berarti Pemrosesan Bahasa Alami (PBA) merupakan cabang dari ilmu komputer dan linguistik yang membahas tentang interaksi antara bahasa manusia (bahasa alami) dan komputer. NLP juga sering disebut sebagai cabang dari kecerdasan buatan (Artificial Intelligence) dan memiliki kajian yang berhubungan dengan linguistik komputasional.

Bahasa dapat dibedakan menjadi 2 yaitu Bahasa Alami dan Bahasa Buatan. Bahasa Alami adalah bahasa yang digunakan manusia untuk berkomunikasi, misalnya Bahasa daerah, bahasa Inggris, Jepang dan sebagainya. Sedangkan Bahasa buatan merupakan bahasa yang sengaja dibuat secara khusus untuk kebutuhan tertentu, seperti bahasa-bahasa pemrograman komputer atau bahasa pemodelan.

Sedikit melihat sejarahnya, NLP dimulai pada tahun 1950-an. Bapak ilmu komputer "Alan Turing" menerbitkan sebuah artikel. Dalam artikel terkenalnya ini Alan Turing mengusulkan sebuah tes yang sekarang dikenal Turing Test. Sebuah tes yang akan mengukur kemampuan mesin (program komputer) untuk menunjukkan kecerdasannya. Ilustrasinya seorang juri (manusia)

akan melakukan percakapan dengan manusia dan mesin yang dites tersebut. Peserta akan dipisahkan satu sama lain. Mesin tersebut akan dinyatakan lulus tes jika juri tidak bisa membedakan antara manusia dan mesin.

Tujuan dari bidang Bahasa Alami ini adalah untuk melakukan proses pembuatan model komputasi dari bahasa manusia, sehingga antara manusia dan komputer dapat melakukan interaksi dengan perantaranya bahasa alami. Sebuah Sistem Bahasa Alami harus memperhatikan pengetahuan dari bahasa itu sendiri baik dari segi kata yang digunakan, apa arti dari sebuah kata, fungsi kata dari sebuah kalimat dan bagaimana dari kata-kata tersebut dapat membentuk sebuah kalimat.

Bahasa alami pada prinsipnya merupakan bentuk dari representasi suatu pesan yang ingin dikomunikasikan oleh antar manusia, bisa berupa suara atau ucapan tetapi juga bisa dalam bentuk tulisan. Intinya adalah dalam pemrosesan bahasa alami, penguraian dari parser atau yang lebih dikenal dengan kalimat. Parser itu sendiri berfungsi untuk membaca kalimat dengan kata demi kata serta menentukan jenis kata apa yang dapat mengikuti kata tersebut.

NLP memiliki beberapa kajian antara lain mencakup, segmentasi teks, segmentasi tuturan, penandaan kelas kata, dan pengawataksaan makna.

Ada 2 bagian utama dari pengolahan bahasa alami, yaitu : parser, sistem representasi pengetahuan dan pengolahan output. Penjelasaannya sebagai berikut :

1. Parser : merupakan suatu sistem yang mengambil kalimat input bahasa alami dan menguraikannya kedalam bagian gramatikal (kata kerja, kata benda, kata sifat dan sebagainya).
 2. Sistem representasi pengetahuan dan pengolahan output : Sistem representasi pengetahuan merupakan sistem yang menganalisis output parser sehingga dapat menentukan maknanya.
- Pengolahan Natural Language Processing
Natural Language Processing memiliki beberapa tingkatan pengolahan yaitu :
 1. Fonetik dan Fonologi : berhubungan dengan hasil kata yang dikenali melalui suara yang diinputkan. Sangat penting untuk aplikasi yang memakai metoda speech bases system.
 2. Morfologi : manfaatnya untuk membedakan antara kata yang satu dan kata yang lainnya. Antara kata, tanda baca atau lainnya akan dipisahkan.
 3. Sintaksis : pengetahuan tentang pemahaman urutan kata untuk membentuk kalimat dan proses perubahan hubungan antar kata yang membentuk kalimat menjadi bentuk yang terstruktur (sistematis).
 4. Semantik : tidak bergantung pada struktur kalimat, bentuk struktur sintaksis akan dipetakan berdasarkan tiap kata ke dalam bentuk yang lebih mendasar. Pada tingkatan ini akan mempelajari arti suatu kata dan bagaimana arti dari

setiap kata tersebut membentuk suatu arti dalam suatu kalimat yang utuh.

5. Pragmatik : pada tingkat ini, konteksnya akan berbeda-beda tergantung dari sistem yang dibuat.
6. Discourse Knowledge : pengetahuan tentang pengenalan apakah kalimat yang sudah diinputkan akan terbaca dan dikenali serta apakah arti dari sebelumnya akan mempengaruhi arti kalimat yang selanjutnya.
7. World Knowledge : pada tahap ini apakah dalam suatu kalimat tersebut ada arti khusus bagi suatu kata. World knowledge mencakup arti kata secara umum.

Tingkatan pengolahan bahasa alami ini definisinya tidak bersifat kaku. Bisa lebih spesifik dengan menambahkan tambahan proses sesuai dengan karakter dari bahasa yang digunakan dan sistem yang dibentuk atau pada beberapa masalah mungkin akan mengambil beberapa pendekatannya saja.

Makna ganda dari suatu kata atau suatu kalimat, atau biasa disebut Ambiguitas, juga merupakan kendala dalam NLP, misalnya bisa saja akan memiliki arti yang berbeda dengan nilai yang sama benar dari masukan yang sama, hal ini tergantung dari keperluan yang memaknai.

- Aplikasi Natural Language

Dalam bidang Natural Language pada umumnya ada 2 jenis aplikasi yang bisa dibuat yaitu :

1. Text – based application : mencakup segala macam aplikasi yang melakukan proses dengan text tertulis misalnya berupa e-mail, berita di surat kabar, buku dan sebagainya. Sebagai contoh : mencari judul buku pada perpustakaan. Dengan menggunakan Text – based application pencarian akan lebih efisien dari sistem database.

Dialogue – based applications : berhubungan dengan pengenalan suara atau bisa juga dengan cara memasukkan text melalui keyboard. Contoh penerapannya pada sistem tanya jawab, dimana bahasa alami digunakan dalam mendapatkan informasi dalam database. Selain itu penerapannya juga bisa pada kontrol suara peralatan elektronik.

Natural Language Processing (NLP) merupakan cabang ilmu AI yang berfokus pada pengolahan bahasa alami. Bahasa alami adalah bahasa yang secara umum digunakan oleh manusia dalam berkomunikasi satu sama lain. Bahasa yang diterima oleh komputer harus diproses dan dipahami terlebih dahulu agar maksud dari pengguna bisa dipahami dengan baik oleh komputer. Beberapa terapan aplikasi dari NLP, Diantaranya Chatbot (membuat pengguna seolah-olah melakukan komunikasi interaktif dengan computer), Stemming/Lemmatization (pemotongan kata menjadi bentuk dasar pengenalan fungsi setiap kata dalam kalimat), Summarization (ringkasan dari bacaan), Translation Tools (menterjemahkan bahasa).

Klasifikasi Natural Language Processing Pustejovsky dan Stubbs (2012) menjelaskan beberapa area utama penelitian pada NLP, diantaranya:

- Question Answering Systems (QAS). Kemampuan komputer untuk menjawab pertanyaan yang diberikan oleh pengguna.
- Summarization. Pembuatan ringkasan dari sekumpulan konten dokumen atau email. Dengan menggunakan aplikasi ini, pengguna dibantu untuk mengkonversikan dokumen teks yang besar ke dalam bentuk slide presentasi.
- Translation Machine. Aplikasi yang dapat memahami bahasa manusia dan menterjemahkannya ke dalam bahasa lain.
- Speech Recognition. Kemampuan telpon/komputer dalam mengenali bahasa yang diucapkan. Bahasa yang sering digunakan biasanya berupa pertanyaan dan perintah.
- Document classification. Menentukan dimana lokasi terbaik dokumen yang baru diinput ke dalam penyimpanan. Hal ini sangat berguna pada aplikasi news article classification, spam filtering, dan movie review.

Terminologi Natural Language Processing Poole dan Mackworth (2010) menjelaskan bahwa ada 3 aspek utama pada teori pemahaman mengenai Bahasa alami:

- Syntax: menjelaskan bentuk dari bahasa. Syntax sering dispesifikasikan sebagai sebuah pola kalimat (grammar).

Bahasa alami jauh lebih daripada Bahasa formal yang digunakan untuk logika kecerdasan buatan dan program computer

- Semantics: menjelaskan arti dari sebuah kalimat dalam satu Bahasa. Teori semantics secara umum sudah ada, pada saat membangun sistem natural language understanding untuk sebuah aplikasi, akan digunakan representasi yang paling sederhana.
- Pragmatics: menjelaskan bagaimana pernyataan yang ada berhubungan dengan dunia. Untuk memahami bahasa, pengembang harus mempertimbangan lebih dari hanya sekedar kalimat. Pengembang aplikasi harus melihat lebih ke dalam konteks kalimat, keadaan dunia, tujuan dari penutur dan pendengar, konvensi khusus, dan sejenisnya.

Contoh Aplikasi NLP

Pada tahun 2013 Suhartono, Christiandy, dan Rolando melakukan penelitian dengan merancang sebuah algoritma lemmatization untuk Bahasa Indonesia. Algoritma itu dibuat untuk menambahkan fungsionalitas pada algoritma Stemming yang sudah pernah dikerjakan sebelumnya yaitu Enhanced Confix-Stripping Stemmer (ECS) (2009). ECS merupakan pengembangan dari algoritma Confix-Stripping Stemmer (2007). Pengembangan yang dikerjakan terdiri dari penambahan beberapa rule dan modifikasi dari rule yang ada. Langkah untuk

melakukan suffix backtracking juga ditambahkan. Hal ini untuk menambah akurasi algoritma.

Secara mendasar, algoritma lemmatization ini tidak bertujuan untuk mengembangkan dari metode ECS, karena tujuannya berbeda. Algoritma lemmatization bertujuan untuk mengoptimalkan ECS, supaya lebih tepat dengan konsep lemmatization. Namun, masih ada beberapa kemiripan pada proses yang ada pada ECS. Pada beberapa kasus dimana ECS belum berhasil untuk digunakan, namun bisa diselesaikan dengan algoritma lemmatization ini.

Bahasa Pemrosesan Alami atau Natural Language processing adalah salah satu bidang ilmu komputer maupun kecerdasan buatan yang berhubungan dengan interaksi komputer dengan bahasa alami manusia, misalnya seperti bahasa inggris, indonesia, china dan sebagainya. Adapun interaksi tersebut dapat berupa suara maupun tulisan.

Adapun tujuan utama dibuatnya BPA adalah untuk membuat mesin-mesin yang mengerti bahasa atau perintah manusia sehingga nantinya diharapkan dapat memberikan respon yang sesuai dengan kebutuhan manusia. BPA sendiri mulai muncul pada tahun 1950-an, meskipun pada tahun-tahun sebelumnya sudah ada penelitian yang mirip.

Alan Turing atau sering dikenal sebagai bapak Ilmu Komputer, melalui artikelnya yang berjudul “ Computing Machinery and Intelligence”, mengusulkan adanya “Turing Test”.Yaitu sebuah tes yang bertujuan untuk mengukur

kemampuan komputer dalam berperilaku cerdas atau sesuai dengan pemikiran manusia. Konsepnya, Seorang penilai (Juri) yang merupakan manusia akan melakukan penilaian terhadap percakapan antara manusia dan komputer yang dipisahkan satu sama lain. Jika Juri kesulitan untuk membedakan antara manusia dan komputer maka dapat dikatakan bahwa komputer tersebut telah lulus tes.

Bahasa Pemrosesan Alami sendiri muncul setelah munculnya pemikiran dari seorang tokoh bernama Chomsky. Ia mempresentasikan bahasa sebagai rangkaian simbol, kemudian ia berhasil membuktikan bahwa bahasa apapun dapat direpresentasikan dengan sebuah cara yang universal. Menurutnya, dengan menggunakan simbol-simbol dan aturan untuk mengatur susunan simbol-simbol ini akan membuka peluang untuk melakukan proses bahasa secara simbolik dengan teknologi komputer.

Sistem pemrosesan bahasa alami secara lisan terbagi menjadi tiga, yaitu:

1. Sub-system Natural Language Processing (NLP), dimana fungsinya adalah untuk melakukan pemrosesan secara simbolik terhadap bahasa tulisan. Contoh bentuk aplikasinya adalah sistem translator dan sistem pemeriksaan sintaks bahasa.
2. Sub-System Text to Speech (TTS), berguna untuk mengubah text menjadi suara/ucapan.

3. Sub-System Speech Recognition (SR), fungsinya adalah untuk mengubah atau mengenali suara/ ucapan menjadi text.

Pada Pemrosesan Bahasa Alami (PBA), teknik-teknik yang digunakan bersifat sangat 'Language Independent', maksudnya adalah teknik yang berlaku pada sistem tersebut tidak mudah untuk diterapkan di sistem lain, dimana umumnya bersifat generik.

System Speech Recognition

Pada System Speech Recognition (Sistem Pengenalan Ucapan), inputnya berupa ucapan manusia, setelah diproses dan diidentifikasi oleh sistem maka akan mendapatkan output berupa text yang sesuai dengan kata yang diucapkan.

Penganalisis inputan akan mentransformasikan sinyal ucapan dari domain waktu ke domain frekuensi. Setiap sinyal akan mempunyai ciri yang unik. Setiap fonem (pengucapan suatu unit bunyi ucapan) kebanyakan akan berbeda antara orang yang satu dengan yang lainnya, tergantung pada fonem-fonem disekitarnya, situasi emosi, noise (kebisingan), dan faktor-faktor lainnya. Sistem akan mengidentifikasi ucapan/suara kemudian fonem itu akan dicari kombinasinya, sehingga hasil ucapannya dapat diterima/ sesuai dengan apa yang diucapkan oleh manusia tersebut.

Pada umumnya Speech Recognition dioperasikan dengan dua mode, yakni mode belajar dan mode produksi. Mode belajar akan melatih sistem menggunakan beberapa kata atau kalimat

dengan kriteria tertentu. Setiap kata atau kalimat akan menghasilkan pola-pola tertentu untuk dipelajari oleh sistem dan disimpan sebagai sebuah referensi. Sedangkan mode produksi (pengenalan ucapan), masing-masing kalimat yang akan diidentifikasi akan di analisis polanya, setelah itu, hasil perbandingan referensi, modul klasifikasi pola, dan pengambilan keputusannya akan mengidentifikasi kata atau kalimat yang diucapkan.

Kemampuan sistem untuk mengenali ucapan sangat bergantung pada referensi yang dihasilkannya melalui proses belajar dalam sistem tersebut. Namun, kemampuan ini dapat dilatih menggunakan korpus (rekaman suara yang heterogen dalam volume yang sangat besar serta memenuhi kriteria teknis).

System Text to Speech

Sistem ini mempunyai proses yang berkebalikan dengan Sistem Pengenalan Ucapan, namun pendekatannya sangatlah berbeda. Pada Umumnya sistem ini terdiri dari dua bagian sebagai berikut:

1. Bagian Konverter Text ke Fonem, yang berfungsi untuk mengubah inputan text menjadi kode-kode suara yang direpresentasikan dengan kode-kode fonem, durasi, serta pitch-nya
2. Bagian Konverter Fonem ke Ucapan, berfungsi untuk menerima masukan kode-kode dari proses sebelumnya

kemudian mengubahnya menjadi bunyi atau ucapan sesuai dengan inputan textnya.

System Natural Language Prosessing

Bahasa Pemrograman Alami tumbuh secara alami guna memenuhi kebutuhan manusia, sehingga akan sulit diproses dibandingkan dengan bahasa buatan karena bahasa alami tidak dirancang dengan memperhatikan kendala-kendala untuk kemudahan proses. Hingga kini, masalah-masalah mendasar mengenai bahasa pemrosesan alami masih belum terselesaikan. Berikut adalah beberapa alasan yang menyulitkan pemrosesan bahasa alami:

1. Sering terjadi ambigu (makna ganda)
2. Kosa kata dalam bahasa alami sangat banyak karena selalu berkembang seiring dengan berjalannya waktu

Ada banyak contoh aplikasi yang dikembangkan menggunakan teknologi bahasa, diantaranya:

1. Alat bantu membaca untuk tunanetra, inputannya berupa text tercetak dan outputnya berupa ucapan dari text yang diinputkan
2. Alat bantu bicara untuk tunawicara, inputannya berupa posisi tangan yang dideteksi oleh sensor dan unit identifikasi, kemudian rangkaian huruf yang teridentifikasi akan disusun membentuk kata-kata kemudian akan dihasilkan suara.

C. Fonetik Otomatis (*Automatic Phonetics*)

Fonetik atau **fonetika** adalah bagian ilmu dalam linguistik yang mempelajari atau menyelidiki bunyi bahasa yang diproduksi oleh manusia tanpa melihat fungsi bunyi itu sebagai pembeda makna dalam suatu bahasa (*langue*). Ilmu **fonetik** menyelidiki bunyi dari sudut pandang tuturan atau ujaran (*parole*).^[1] Di sisi lain fonologi adalah ilmu yang berdasarkan fonetik dan mempelajari sistem fonetika.

International Phonetic Association (IPA) telah mengamati lebih dari 100 bunyi manusia yang berbeda dan mentranskripsikannya dengan *International Phonetic Alphabet* mereka.

Sejarah

Ilmu fonetika pertama kali dipelajari sekitar abad ke-5 SM di India Kuno oleh Pāṇini, sang resi yang mempelajari bahasa Sanskerta. Semua aksara yang berdasarkan aksara India sampai sekarang masih menggunakan klasifikasi Panini ini, termasuk beberapa aksara Nusantara. Tulisan Yunani Kuno dinobatkan sebagai dasar pertama penulisan lambang alfabet. Fonetika modern diawali oleh Alexander Melville Bell melalui bukunya *Visible Speech* (1867) yang memperkenalkan suatu sistem penulisan bunyi-bunyi bahasa secara teliti dan teratur. Ilmu fonetik kemudian berkembang dengan pesat di akhir abad ke-19 akibat ditemukannya fonograf, yang membantu perekaman bunyi-bunyi bahasa. Berkat alat tersebut, fonetisi dapat

mempelajari bunyi-bunyi bahasa dengan lebih baik, mudah, dan akurat dari sebelumnya karena alat tersebut dapat mengulang-ulang tuturan yang direkamnya sampai fonetisi dapat menganalisisnya dengan akurat. Dengan menggunakan fonograf Edison, Ludimar Hermann menyelidiki sifat-sifat spektral dalam bunyi vokoid dan kontoid. Dalam karya ilmiahnya istilah formant diperkenalkan. Hermann juga memutar-mutar bunyi-bunyi vokoid menggunakan fonograf Edison dalam berbagai kecepatan dalam rangka menguji teori Willis dan Wheatstone mengenai produksi bunyi vokoid.

Subbidang ilmu

Fonetika memiliki tiga cabang utama:

- fonetik organik atau artikulatoris ialah fonetik yang mempelajari bagaimana mekanisme alat-alat bicara yang ada dalam tubuh manusia menghasilkan suatu bunyi bahasa. Dalam fonetik ini yang dipelajari adalah posisi dan gerakan bibir, lidah dan organ-organ manusia lainnya yang memproduksi suara atau bunyi bahasa.
- fonetik akustik ialah fonetik yang mempelajari bunyi bahasa dari segi bunyi sebagai gejala fisik. Dalam fonetik ini yang dipelajari adalah gelombang suara dan bagaimana mereka didengarkan oleh telinga manusia.
- fonetik auditoris ialah fonetik yang mempelajari bagaimana mekanisme telinga menerima bunyi bahasa sebagai getaran udara. Dalam fonetik ini yang dipelajari adalah

proses resepsi bunyi dan terutama bagaimana otak mengolah data yang masuk sebagai suara.

Jenis-jenis fonetik

- Aksent nontirus
- Aksent nontirus ganda
- Aksent tirus
- Aksent tirus ganda
- Akustik fonetik
- Artikulasi fonetik
- Auditif Fonetik
- Cedilla
- Dot
- Ligatur
- Sirkumfleks
- Tanda gelombang
- Tirus
- Tirus panjang
- Umlaut/Dwinitik/Trema

BAB 9

TEKNIK MENGUBAH DARI FONEM KE UCAPAN

(TECHNIQUE CHANGES FROM FONEM TO RECREATION)

A. Strategi sintesa, Linear Prediction Synthesis, Algoritma Time domain (*Synthesis strategy, Linear Prediction Synthesis, Time domain algorithm*)

Pengenalan suara adalah teknologi masa depan yang menggantikan cara interaksi manusia dengan komputer dengan menggabungkan beberapa disiplin ilmu seperti pengenalan sinyal dan pengenalan pola. Dimana interaksi user dengan sistem dapat dilakukan dengan memberikan inputan suara.

Linier Predictive Coding (LPC) adalah salah satu metode parametrik yang digunakan untuk merepresentasikan sinyal. Umumnya LPC digunakan karena menyediakan pemodelan yang baik untuk sinyal suara, LPC dapat dengan mudah dan langsung diterapkan baik secara perangkat lunak maupun perangkat keras karena perhitungan matematis yang dilibatkan realtif lebih singkat dari metode-metode yang dikenal sebelumnya.

Fast Fourier transform (FFT) menjadi penting untuk bermacam-macam aplikasi, dari pengolahan sinyal digital dan memecahkan persamaan diferensial parsial menjadi algoritma-algoritma untuk penggandaan bilangan integer dalam jumlah yang banyak.

Secara garis besar, cara kerja sistem pengenalan suara ini ialah mula-mula sinyal suara manusia yang diterima dengan

menggunakan microphone (sinyal analog) dicuplik sehingga menjadi sinyal digital dengan bantuan sound card pada komputer. Sinyal digital hasil cuplikan ini terlebih dulu dinormalisasi kemudian diproses awal menggunakan metode LPC sehingga didapat beberapa koefisien LPC yang merupakan feature (ciri) dari suara pembicaraan. Kemudian koefisien LPC tersebut diproses dengan Fast Fourier Transform (FFT) untuk mendapatkan sinyal pada domain frekuensi. Hal ini bertujuan agar perbedaan antar pola kata yang satu dengan yang lain terlihat lebih jelas sehingga ekstraksi parameter sinyal memberikan hasil yang lebih baik. Hasil keluaran FFT ini akan di compare sebagai fungsi utama dari sistem untuk proses pengenalan.

B. Pemrosesan Bahasa Alamiah dan TTS Synthesis (*Processing natural languages and TTS Synthesis*)

Beberapa aplikasi text mining awal menggunakan penyajian sederhana yang disebut dengan 'bag-of-words' ketika mengenalkan struktur ke suatu kumpulan dokumen berbasis teks untuk mengklasifikasikan nya menjadi dua atau lebih kelas yang sudah ditentukan atau untuk meng-cluster nya menjadi pengelompokan-pengelompokan alami. Dalam model 'bag-of-words' tersebut, teks, misalnya suatu kalimat, paragraph, atau dokumen penuh, disajikan sebagai kumpulan kata, dengan mengabaikan tata bahasa atau urutan kata-kata yang akan muncul. Model 'bag-of-words' masih digunakan dalam beberapa tool klasifikasi dokumen yang sederhana. Contohnya, dalam

memfilter spam email, suatu pesan di email bisa dimodelkan sebagai suatu kumpulan kata-kata yang tak berurutan (a bag-of-words) yang dibandingkan dengan dua 'bags' berbeda yang sudah ditentukan sebelumnya. Satu 'bag' diisi dengan kata-kata yang ditemukan di dalam pesan-pesan spam dan yang satunya diisi dengan kata-kata yang ditemukan dalam email-email yang legitimate. Meskipun beberapa kata mungkin ditemukan dalam kedua 'bags', bag untuk spam akan berisi kata-kata yang terkait dengan spam seperti stock, Viagra, dan buy jauh lebih sering muncul dibanding dengan bag untuk email legitimate, yang berisi lebih banyak kata yang terkait dengan teman-teman dan tempat kerja pengguna. Tingkat kecocokan antara 'bag-of-words' dari email tertentu dan dua bags yang berisi 'descriptor' menentukan apakah email akan menjadi spam atau legitimate.

Secara alami, kita (manusia) tidak menggunakan kata-kata tanpa suatu urutan atau struktur. Kita menggunakan kata-kata dalam kalimat, yang memiliki semantik dan struktur sintaksis. Teknik-teknik otomatis (seperti teks mining) harus mencari cara untuk melebihi kemampuan interpretasi 'bag-of-words' dan menyatukan makin lama makin banyak struktur semantik dalam operasinya. Trend saat ini dalam text mining mengarah untuk selalu memasukkan banyak fitur-fitur canggih yang bisa diperoleh dengan menggunakan pemrosesan bahasa alami (natural language processing).

Sudah dibuktikan bahwa metode 'bag-of-word' mungkin tidak menghasilkan isi informasi yang cukup bagus untuk tugas-

tugas dalam text mining (misalnya, klasifikasi, clustering, asosiasi). Contoh yang bagus bisa ditemukan dalam evidence-based medicine (pengobatan/kedokteran berbasis bukti). Komponen penting dalam pengobatan berbasis bukti ini adalah menyatukan temuan-temuan penelitian terbaik yang ada ke proses pembuatan keputusan klinis, yang melibatkan penilaian terhadap informasi yang dikumpulkan dari media cetak untuk validasi dan relevansi.

Beberapa peneliti dari Univeristy of Maryland telah mengembangkan beberapa model penilaian bukti dengan menggunakan metode 'bag-of-words'. Mereka menerapkan beberapa metode machine-learning yang populer disertai dengan lebih dari setengah juta artikel penelitian yang dikumpulkan dari MEDLINE (Medical Literature Analysis and Retrieval System Online). Dalam model-model mereka, mereka menyajikan setiap abstrak sebagai 'bag-of-words', dimana setiap istilah menunjukkan satu feature. Meskipun menggunakan beberapa metode klasifikasi populer dengan berbagai metodologi design eksperimental, hasil-hasil prediksi dari eksperimen tersebut tidak jauh lebih baik dibandingkan dengan menebak secara sederhana, yang mungkin menunjukkan bahwa 'bag-of-words' tidak menghasilkan penyajian yang cukup bagus untuk artikel-artikel penelitian dalam domain ini; karena itu, teknik-teknik yang lebih handal seperti pemrosesan bahasa alami (Natural Language Processis/NLP) diperlukan.

Pemrosesan Bahasa Alami atau Natural Language Processing (NLP) adalah komponen penting dalam text mining dan subbidang dalam kecerdasan buatan (artificial intelligence/AI) dan komputasi linguistic. Dia mempelajari tentang bagaimana ‘memahami’ bahasa manusia alami, dengan cara mengubah gambaran bahasa manusia (seperti dokumen teks) menjadi penyajian yang lebih formal (dalam bentuk data numeric dan simbolik) yang lebih mudah untuk dimanipulasi oleh program computer. Tujuan NLP adalah untuk melangkah melebihi manipulasi teks berbasis sintaks (yang sering kali disebut dengan ‘word counting’) ke pemahaman yang benar dan memproses bahasa alami yang mempertimbangkan berbagai batasan semantik dan gramatikal dan juga konteks.

Definisi dan scope dari kata ‘memahami’ adalah salah satu topic utama dalam diskusi tentang NLP. Dengan mempertimbangkan bahwa bahasa alami manusia adalah kabur dan bahwa pemahaman yang benar terhadap suatu arti memerlukan pengetahuan yang luas terhadap suatu topic (jauh berada diluar kata, kalimat, dan paragraph), akankah komputer mampu memahami bahasa alami dengan cara yang sama dan akurasi yang sama dengan manusia? Barangkali tidak! NLP telah ada lama sejak era ‘word counting’ yang sederhana, namun dia masih perlu lebih lama lagi untuk menuju ke benar-benar mampu memahami bahasa alami manusia. Berikut adalah beberapa tantangan yang biasanya dikaitkan dengan implementasi NLP:

- **Part-of-speech tagging.** Sangatlah sulit untuk menandai istilah-istilah dalam suatu teks yang terkait dengan bagian tertentu dari suatu naskah (misalnya kata benda, kata kerja, kata sifat, kata keterangan, dst), karena bagian dari naskah tidak hanya bergantung pada definisi istilah tetapi juga pada konteks dimana teks digunakan.
- **Text segmentation.** Beberapa bahasa tulisan, seperti bahasa mandarin, jepang, dan thai, tidak memiliki batasan kata. Dalam contoh ini, tugas text-parsing memerlukan identifikasi terhadap batasan kata, yang seringkali merupakan tugas yang sangat sulit. Tantangan serupa dalam segmentasi naskah muncul ketika menganalisa bahasa verbal, karena suara menyajikan rangkaian huruf dan kata yang bercampur satu sama lain.
- **Word sense disambiguation.** Banyak kata yang memiliki lebih dari satu arti. Memilih arti yang paling masuk akan hanya bisa dicapai dengan mempertimbangkan konteks di mana kata digunakan.
- **Syntactic ambiguity.** Tata bahasa dalam bahasa alami seringkali ambigu; artinya, ada berbagai struktur kalimat yang memungkinkan yang perlu dipertimbangkan. Memilih struktur yang paling tepat biasanya memerlukan paduan informasi kontekstual dan semantik.
- **Imperfect or irregular input.** Aksentuasi asing atau lokal dan berbagai hambatan vokal dalam pidato dan kesalahan ketik

dan tata bahasa dalam teks-teks menyebabkan pengolahan bahasa bahkan lebih sulit.

- **Speech acts.** Suatu kalimat seringkali dianggap sebagai suatu aksi oleh si pembicara. Struktur kalimatnya sendiri mungkin tidak berisi cukup informasi untuk mendefinisikan tindakan ini. Contohnya, “Can you pass the class?” memerlukan jawaban singkat ya/tidak, padahal “Can you pass the salt?” adalah permintaan untuk melakukan suatu tindakan nyata untuk dilakukan.

Suatu mimpi panjang bagi komunitas kecerdasan buatan (AI) untuk memiliki algoritma yang mampu secara otomatis membaca dan memperoleh pengetahuan dari teks. Dengan menerapkan algoritma ‘learning’ terhadap teks yang diurai, para peneliti dari lab NLP Stanford University telah mengembangkan metode yang bisa secara otomatis mengidentifikasi konsep dan keterkaitan antara konsep-konsep tersebut di dalam teks. Dengan menerapkan prosedur unik terhadap sejumlah besar teks, algoritma-algoritma mereka secara otomatis memperoleh ratusan ribu item pengetahuan dunia dan menggunakan mereka untuk menghasilkan repository yang sudah ditingkatkan secara signifikan untuk WordNet. WordNet adalah database kosa kata bahasa Inggris yang dibuat dengan susah payah, definisi-definisinya, kumpulan sinonim, dan berbagai hubungan semantik antara kumpulan sinonim. Ini adalah sumberdaya utama bagi aplikasi NLP, tetapi telah terbukti sangat mahal untuk dibuat dan dimaintain secara manual. Dengan secara otomatis mendorong

pengetahuan ke dalam WordNet, ada potensi untuk membuat WordNet sumber daya yang lebih besar dan lebih komprehensif untuk NLP dengan sebagian kecil dari biaya itu.

Salah satu bidang terkenal dimana manfaat NLP sudah berhasil dipetik adalah di CRM (Customer Relationship Management). Secara umum, tujuan CRM adalah untuk memaksimalkan nilai pelanggan dengan memahami lebih baik dan merespon secara efektif terhadap kebutuhan actual mereka. Bidang penting dari CRM dimana NLP berpengaruh secara signifikan adalah analisa sentimen (sentiment analysis). Sentiment analysis adalah suatu teknik yang digunakan untuk mendeteksi opini positif dan negative terhadap suatu produk dan layanan tertentu dengan menggunakan sejumlah besar data teks (feedback pelanggan dalam bentuk posting web).

Sentiment analysis menawarkan peluang yang sangat besar untuk berbagai penerapan. Contohnya, dia akan memberikan fungsionalitas yang sangat dahsyat untuk analisa kompetitif, analisa marketing, dan deteksi rumor-rumor negatif terhadap manajemen risiko. Pendekatan sentiment analysis yang dikembangkan oleh para peneliti IBM berusaha untuk mengekstrak berbagai sentiment yang terkait dengan polaritas positif atau negative dari subyek tertentu (misalnya, produk atau layanan) dari sekumpulan dokumen. Isu utama dalam sentiment analysis adalah bagaimana mengidentifikasi sentiment yang dinyatakan dalam teks dan apakah pernyataan menunjukkan opini positif atau negative terhadap subyek. Untuk meningkatkan

akurasi sentiment analysis, hal yang penting untuk diperhatikan adalah mengidentifikasi keterkaitan semantic antara pernyataan sentiment dan subyek. Dengan menerapkan analisa semantic dengan menggunakan parser kalimat dan kamus sentiment, system dari IBM telah mencapai presisi yang tinggi (75-95%, bergantung pada data) dalam menemukan sentiment-sentimen dalam halaman-halaman web dan artikel-artikel berita.

NLP telah sukses diterapkan pada berbagai pekerjaan melalui program-program computer untuk memproses secara otomatis bahasa manusia yang sebelumnya hanya bisa dilakukan oleh manusia. Berikut ini adalah beberapa dari pekerjaan-pekerjaan yang paling populer dilakukan oleh NLP:

- **Information retrieval.** Ilmu untuk melakukan pencarian terhadap berbagai dokumen yang relevan, menemukan informasi tertentu didalamnya, dan menghasilkan metadata untuk isinya.
- **Information extraction.** Sejenis 'information retrieval' yang tujuannya adalah untuk mengekstrak secara otomatis informasi terstruktur, seperti data yang sudah terdefinisi dengan baik secara semantik dan secara kontekstual yang sudah terkelompok dari domain tertentu, dengan menggunakan berbagai dokumen tak-terstruktur yang bisa terbaca oleh mesin.
- **Named-entity recognition.** Juga dikenal sebagai 'entity identification' dan 'entity extraction', ini adalah sub-pekerjaan dari 'information extraction' yang bertujuan untuk

menemukan dan mengklasifikasikan elemen-elemen atomic di dalam teks menjadi berbagai kategori yang sudah ditentukan sebelumnya, seperti nama orang, organisasi, lokasi, pernyataan waktu, kuantitas, nilai uang, persentase, dan seterusnya.

- **Questioning answering.** Pekerjaan menjawab secara otomatis suatu pertanyaan yang diajukan dalam bahasa alami; yaitu, menghasilkan jawaban bahasa manusia ketika diberi pertanyaan bahasa manusia. Untuk mendapatkan jawaban terhadap pertanyaan, program computer bisa menggunakan baik database pra-terstruktur atau kumpulan dokumen bahasa alami (suatu 'text corpus' seperti world wide web).
- **Automatic summarization.** Pembuatan ringkasan terhadap dokumen teks yang dilakukan oleh program komputer yang berisi poin-poin paling penting dari dokumen asli.
- **Natural language generation.** Sistem mengubah informasi dari database komputer menjadi bahasa manusia yang bisa dibaca.
- **Natural language understanding.** Sistem mengubah contoh-contoh bahasa manusia menjadi penyajian yang lebih formal yang lebih mudah bagi program komputer untuk memanipulasi.
- **Machine translation.** Penerjemahan otomatis dari satu bahasa manusia ke bahasa yang lain.

- **Foreign language reading.** Program komputer yang membantu orang yang non-native untuk membaca bahasa asing dengan pengucapan benar dan aksen pada bagian yang berbeda dengan benar.
- **Foreign language writing.** Program komputer yang membantu orang yang non-native dalam menulis dalam bahasa asing.
- **Speech recognition.** Mengubah kata-kata verbal menjadi input yang bisa terbaca oleh mesin. Dengan adanya sound clip dari orang yang sedang berbicara, sistem menghasilkan dikte dari teks.
- **Text-to-speech.** Disebut juga 'speech synthesis', suatu program komputer yang mengubah secara otomatis teks menjadi suara manusia.
- **Text proofing.** Program komputer yang membaca bukti copy teks yang bertujuan untuk mendeteksi dan mengoreksi kesalahan.
- **Optical character recognition.** Penerjemahan otomatis dari gambar tulisan tangan, tulisan ketikan, atau teks cetak (biasanya diambil dengan scanner) menjadi dokumen-dokumen teks yang bisa diedit oleh mesin.

Keberhasilan dan popularitas text mining sangat bergantung pada perkembangan NLP baik dalam proses menghasilkan maupun memahami bahasa manusia. NLP memungkinkan ekstraksi berbagai features dari teks tak-terstruktur sehingga berbagai jenis teknik data mining bisa digunakan untuk mengekstrak pengetahuan

(pola dan keterkaitan yang baru dan bermanfaat) darinya. Dalam arti tersebut, gampangnya, text mining adalah kombinasi NLP dan data mining.

C. Pemrosesan sinyal digital dan TTS Synthesis (*Digital signal processing and TTS Synthesis*)

Di berbagai media, Anda mungkin pernah menyaksikan Stephen Hawking berbicara di depan mahasiswanya. Fisikawan yang terkenal dengan teori black hole-nya ini sudah tidak mampu lagi mengeluarkan suara dari lisannya, namun berkat teknologi speech synthesizer, dia masih bisa bercakap-cakap. Mesin speech synthesizer Hawking memang cukup kompleks. Alat ini tidak hanya memproduksi suara, tetapi juga menangkap input dari gerakan mata sang doktor. Demikian pula, misalnya, dengan aplikasi voice command yang banyak tertanam di smartphone mutakhir yang memadukan speech recognizer dengan speech synthesizer.

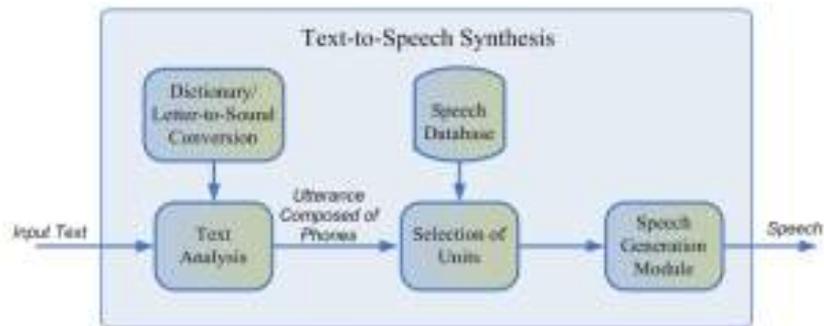


Aplikasi speech synthesizer yang paling sederhana sebenarnya ada pada setiap PC ber-OS Windows. Bila anda menekan tuts Winkey + U di keyboard, Windows akan mengaktifkan Utility Manager, yang di dalamnya terdapat aplikasi Microsoft Narrator. Aplikasi ini akan membaca setiap jendela yang anda aktifkan, termasuk tombol-tombol di dalamnya. Atau, mungkin anda pernah menginstal aplikasi microsoft reader di PC. Aplikasi yang diperuntukkan bagi file >LTT ini pun dilengkapi dengan kemampuan menerjemahkan teks menjadi suara (text to speech) yang merupakan contoh teknologi speech sybthesizer.

Speech synthesis adalah sebuah kemampuan bicara manusia yang dibuat oleh manusia (artificial). Sebuah sistem komputer digunakan untuk tujuan ini yang disebut sebagai speech synthesizer, dan dapat diimplementasikan ke dalam software atau hardware. Sebagai contoh sebuah sistem text-to-speech (TTS) yang dapat mengkonversikan teks dengan bahasa biasa menjadi suara.

Sebuah sistem komputer yang digunakan untuk tujuan ini disebut speech synthesizer, dan dapat diimplementasikan dalam perangkat lunak atau perangkat keras produk. Sebuah teks-to-speech (TTS) sistem mengkonversi teks bahasa normal menjadi berbicara; sistem lain membuat representasi linguistik simbolik seperti transkripsi fonetik ke dalam pidato, pidato disintesis dapat dibuat dengan menggabungkan potongan pidato direkam yang disimpan dalam database. Sistem berbeda dalam ukuran unit pidato disimpan, sebuah sistem yang menyimpan telepon atau

diphones menyediakan berbagai keluaran terbesar, tapi mungkin kurang jelas.



Speech synthesis adalah transformasi dari teks ke arah suara (speech). Transformasi ini mengkonversi teks ke pemadu suara (speech synthesis) yang sebisa mungkin dibuat menyerupai suara nyata, disesuaikan dengan aturan – aturan pengucapan bahasa. TTS (text to speech) dimaksudkan untuk membaca teks elektronik dalam bentuk buku, dan juga untuk menyuarakan teks dengan menggunakan pemaduan suara. Sistem ini dapat digunakan sebagai sistem komunikasi, pada sistem informasi referral, dapat diterapkan untuk membantu orang-orang yang kehilangan kemampuan melihat dan membaca.

Synthesized speech dapat diciptakan dengan menggabungkan beberapa potongan-potongan dari pembicaraan/pidato yang sudah direkam dalam sebuah basis data. Kualitas dari sebuah speech synthesizer dilihat dari kemiripannya dengan suara manusia dan kemampuannya untuk bisa dipahami. Program TTS yang jelas dapat membantu orang

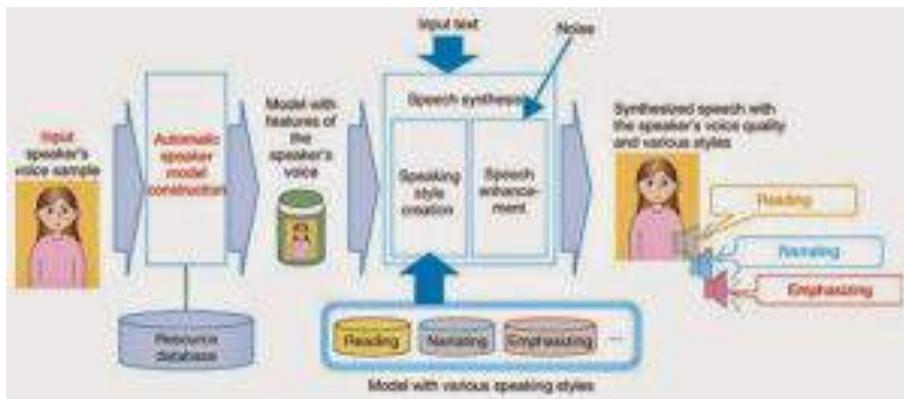
dengan gangguan visual atau ketidakmampuan membaca, untuk mendengarkan pada pekerjaan yang tertulis dalam komputer. Banyak Sistem Operasi komputer yang telah dimasukkan speech synthesizer sejak tahun 1980-an.

Sebuah sistem text-to-speech (atau “mesin”) terdiri dari dua bagian: front-end dan back-end . Front-end memiliki dua tugas utama. Pertama, mengubah teks mentah berisi simbol seperti angka dan singkatan menjadi setara dengan kata-kata tertulis-out. Proses ini sering disebut normalisasi teks, pra-pengolahan, atau tokenization . Front-end kemudian memberikan transkripsi fonetik untuk setiap kata, dan membagi dan menandai teks ke unit prosodi , seperti frase , klausa , dan kalimat . Proses menetapkan transkripsi fonetik untuk kata-kata ini disebut teks-ke-fonem atau grafem konversi -untuk-fonem. Transkripsi fonetik dan informasi prosodi bersama-sama membentuk representasi linguistik simbolik yang output dengan front-end. Back-end-sering disebut sebagai synthesizer-maka mengubah representasi linguistik simbolik menjadi suara. Dalam sistem tertentu, bagian ini meliputi perhitungan dari target prosodi (kontur pitch, durasi fonem), yang kemudian dikenakan pada pidato output.

Sejarah Speech Synthesis

Upaya yang paling awal untuk menghasilkan lahirnya pemandu suara, pada abad XVIII. Terlepas dari kenyataan bahwa upaya pertama adalah bentuk mesin mekanis, kita dapat mengatakan hari ini bahwa synthesizer sudah berkualitas tinggi.

Pada tahun 1779 di St Petersburg, Rusia Profesor Kratzenshtein Kristen fisiologis menjelaskan perbedaan antara lima vokal panjang (/ A /, / e /, / i /, / o /, dan / u /) dan membuat alat untuk menghasilkan mereka artifisial. Tahun 1791 di Wina, Wolfgang von Kempelen memperkenalkan nya “Akustik-Mekanik Mesin Speech”. Dalam sekitar pertengahan 1800-an Charles Wheatstone dibangun terkenal versi mesin berbicara von Kempelen’s.



Generasi dari sistem pemaduan suara ini dapat dibagi ke dalam 3 masa, yaitu:

1. Generasi pertama (1962-1977). Format sintesis dari fonem adalah teknologi dominan. Teknologi ini memanfaatkan aturan berdasarkan penguraian fonetik pada kalimat untuk kontur frekuensi formant. Beberapa sintesis masih miskin atau kurang dalam kejelasan dan kealamiannya.
2. Generasi kedua (1977-1992). Metode pemaduan suara adalah diphone diwakilkan dengan parameter LPC. Hal

tersebut menunjukkan bahwa kejelasan yang baik pada pemadu suara dapat diperoleh dengan andal dari input teks dengan menggabungkan diphone yang sesuai dengan unit. Kejelasan meningkat selama sintesis formant, tetapi kealamian dari pemadu suara masih tetap rendah.

3. Generasi ketiga (1992-sekarang). Generasi ini ditandai dengan metode ‘ sintesis pemilihan unit’ yang diperkenalkan dan disempurnakan oleh Sagisaka di Labs ATR Kyoto. Hasil dari pemadu suara pada periode ini sangat mendekati human-generated speech pada bagian kejelasan dan kealamian,

Teknologi pemadu suara modern melibatkan metode dan algoritma yang canggih dan rumit. alat pemadu suara dari keluarga “Infovox” mungkin menjadi salah satu multi bahasa TTS yang paling dikenal saat ini. Versi komersial pertamanya, Infovox-SA 101, dikembangkan pada tahun 1982 di Institute Teknologi Royal, Swedia dan didasarkan pada sintesis formant.

AT & T Bell Laboratories (Lucent Technologies) juga memiliki tradisi yang sangat panjang tentang pemadu suara (speech synthesis). TTS lengkap yang pertama didemostrasikan di Boston pada tahun 1972 dan diliris pada tahun 1973. Hal ini didasarkan pada model artikulatoris yang dikembangkan oleh Ceecil Coker (Klatt 1987). Pengembangan proses dari sistem penggabungan sintesis ini dimulai oleh Joseph Olive pada pertengahan tahun 1970-an (Bell Labs 1997). Sistem ini sekarang

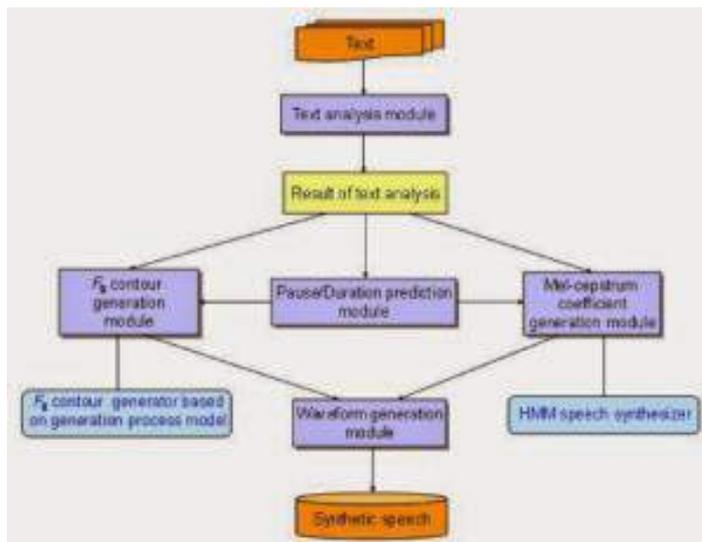
sudah tersedia untuk bahasa Inggris, Perancis, Spanyol, Italia, Jerman, Rusia, Rumania, Cina, dan Jepang (Mcbius et al 1996).

Perangkat Speech Synthesis

Pidato sistem sintesis berbasis komputer pertama diciptakan pada akhir 1950-an. Pertama umum Inggris sistem text-to-speech dikembangkan oleh Noriko Umeda et al. Pada tahun 1968 di Laboratorium Elektroteknik, Jepang. Pada tahun 1961, fisikawan John Larry Kelly, Jr dan Louis rekan Gerstman menggunakan IBM 704 komputer untuk mensintesis pidato, acara yang paling menonjol dalam sejarah Bell Labs . Kelly perekam suara synthesizer (vocoder) ulang lagu " Daisy Bell ", dengan iringan musik dari Max Mathews .Kebetulan, Arthur C. Clarke mengunjungi teman dan kolega John Pierce di fasilitas Bell Labs Murray Hill. Clarke begitu terkesan oleh demonstrasi bahwa ia digunakan dalam adegan klimaks dari skenario-Nya untuk novel nya 2001: A Space Odyssey, di mana HAL 9000 komputer menyanyikan lagu yang sama seperti yang sedang ditidurkan oleh astronot Dave Bowman. Meskipun keberhasilan pidato sintesis murni elektronik, penelitian masih terus dilakukan ke synthesizer pidato mekanis.

Handheld elektronik menampilkan sintesis pidato mulai muncul pada 1970-an. Salah satu yang pertama adalah Telesensory Systems Inc(TSI) Pidato + kalkulator portabel untuk orang buta pada tahun 1976. Perangkat lain yang diproduksi terutama untuk tujuan pendidikan, seperti Bicara & Eja , yang diproduksi oleh Texas Instruments pada tahun 1978. Fidelity

merilis versi berbicara komputer catur elektronik pada tahun 1979. Yang pertama video game yang memiliki fitur sintesis pidato adalah 1.980 shoot 'em up arcade game , Stratovox , dari Sun Electronics. Contoh lain awal adalah versi arcade dari Berzerk , dirilis pada tahun yang sama. Pertama multi-player permainan elektronik menggunakan sintesis suara adalah Milton dari Milton Bradley Company , yang memproduksi perangkat di tahun 1980.



Teknologi Speech Synthesis

Yang paling penting dalam kualitas sistem speech synthesis adalah kealamian dan kejelasannya. Kealamaan menjelaskan bagaimana dekatnya suara output dengan suara manusia, sementara kejelasan adalah dengan kemudahan di mana output tersebut dapat dipahami. Speech synthesizer yang

ideal adalah yang alami dan jelas. Sistem speech synthesis biasanya mencoba untuk memaksimalkan kedua karakteristik.

Kualitas terpenting dari sebuah aplikasi speech synthesizer adalah seberapa alami dan inteligibel output yang dihasilkannya. Alami, artinya seberapa dekat suara yang dihasilkan aplikasi speech synthesizer dengan suara manusia. Sedangkan inteligibel adalah seberapa mudah output tersebut dipahami oleh manusia. Semua aplikasi speech synthesizer berusaha untuk menghasilkan output yang alami dan inteligibel sekaligus.

Sampai saat ini, ada banyak teknologi untuk meng-generate gelombang suara sintesis ini. Dua teknologi yang paling banyak digunakan adalah concatenative synthesis dan formant synthesis. Keduanya memiliki keunggulan dan kekurangan sendiri-sendiri.

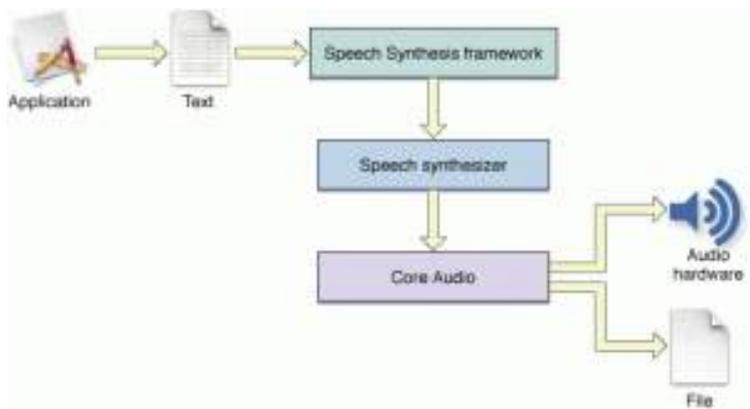
Teknologi pertama, concatenative synthesis, berbasis pada rangkaian (atau merangkai bersama) segmen-segmen dari suara yang direkam. Umumnya, teknologi ini menghasilkan suara sintesis yang terdengar paling alami. Namun, perbedaan antara suara alami yang direkam dengan segmentasi gelombang bunyi kadang menghasilkan suara yang mengganggu. Mirip seperti suara pemberitahuan nomor antrean di bank atau suara call center operator ponsel yang menyebutkan sisa pulsa dan masa berlaku kartu ponsel.

Teknologi kedua, formant synthesis, tidak menggunakan sampel suara manusia melainkan membuat suara sintesis menggunakan model akustik. Parameter-parameter seperti

frekuensi dasar, alunan suara, dan tingkat kebisingan bervariasi dari waktu ke waktu untuk menciptakan gelombang suara buatan. Kebanyakan aplikasi berbasis teknologi ini menghasilkan suara buatan (tidak alami) seperti suara robot. Melihat keterbatasan kedua teknologi ini dalam menghasilkan suara buatan, seperti kita harus sabar menunggu pengembangannya lebih lanjut dalam beberapa tahun atau dekade ke depan.

Kualitas yang paling penting dari sebuah sistem sintesis pidato kewajaran dan dimengerti. Kealamian menjelaskan seberapa dekat output terdengar seperti suara manusia, sementara kejelasan adalah kemudahan yang output dipahami. Speech synthesizer yang ideal adalah baik alami dan dimengerti. Sistem sintesis pidato biasanya mencoba untuk memaksimalkan kedua karakteristik.

Dua teknologi utama dalam pembuatan gelombang suara synthetic speech adalah Concatenative Synthesis dan Formant Synthesis. Setiap teknologi mempunyai kekuatan dan kelemahannya, dan penggunaan yang ditujukan dari sistem synthesis akan menentukan pendekatan mana yang digunakan.



1. Concatenative Synthesis :

Concatenative synthesis didasarkan dengan penggabungan dari segmen-segmen dari pembicaraan yang sudah direkam. Secara umum, concatenative synthesis memproduksi synthesized speech dengan suara yang paling alami. Tetapi, perbedaan antara variasi alami dalam pembicaraan dan sifat dari teknik otomatis untuk pensegmentasian gelombang suara terkadang menghasilkan kesalahan suara dalam output. Namun, perbedaan antara variasi alami dalam pidato dan sifat teknik otomatis untuk membagi bentuk gelombang kadang-kadang menyebabkan gangguan terdengar pada output. Ada tiga sub-jenis utama dari sintesis concatenative.

Sintesis Pemilihan unit

Sintesis Pemilihan unit menggunakan besar database pidato direkam. Selama pembuatan database, setiap ucapan tercatat tersegmentasi ke dalam beberapa atau semua hal berikut: individu telepon , diphones , setengah-telepon, suku kata , morfem, kata , frase , dan kalimat . Biasanya, pembagian ke dalam segmen dilakukan dengan menggunakan dimodifikasi khusus recognizer pidato disetel ke “keselarasan dipaksa” mode dengan beberapa koreksi manual setelah itu, dengan menggunakan representasi visual seperti yang gelombang dan spektogram . Sebuah indeks unit dalam database pidato kemudian dibuat berdasarkan segmentasi dan parameter akustik seperti frekuensi

dasar (lapangan), durasi, posisi dalam suku kata, dan telepon tetangga. Pada waktu berjalan , target ucapan yang diinginkan dibuat dengan menentukan rantai terbaik unit calon dari database (pemilihan unit). Proses ini biasanya dicapai dengan menggunakan khusus tertimbang pohon keputusan .

Pemilihan unit menyediakan kealamian terbesar, karena hanya berlaku sedikit pemrosesan sinyal digital (DSP) untuk pidato direkam. DSP sering membuat pidato yang direkam terdengar kurang alami, meskipun beberapa sistem menggunakan sejumlah kecil pengolahan sinyal pada titik Rangkaian untuk menghaluskan bentuk gelombang. Output dari yang terbaik unit-seleksi sistem sering dibedakan dari suara manusia nyata, terutama dalam konteks dimana sistem TTS telah disetel. Namun, kealamian maksimum biasanya membutuhkan unit-pilihan database pidato menjadi sangat besar, dalam beberapa sistem mulai ke gigabyte data dicatat, mewakili puluhan jam berbicara. Juga, pilihan algoritma Unit telah dikenal untuk memilih segmen dari Tempat yang menghasilkan kurang dari sintesis ideal (misalnya kata-kata kecil menjadi tidak jelas) bahkan ketika pilihan yang lebih baik ada dalam database. Baru-baru ini, peneliti telah mengusulkan berbagai metode otomatis untuk mendeteksi segmen alami di unit-pilihan sistem sintesis pidato.

Sintesis diphone

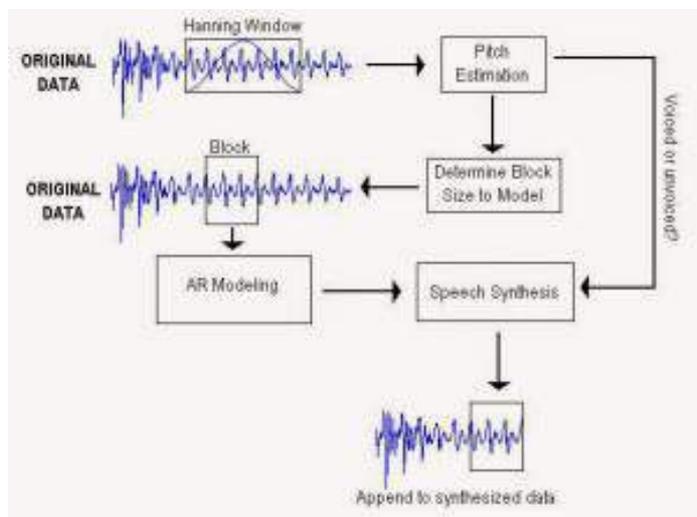
Sintesis diphone menggunakan database pidato minimal berisi semua diphones (suara-to-suara transisi) yang terjadi dalam

suatu bahasa. Jumlah diphones tergantung pada fonotaktik bahasa: misalnya, Spanyol memiliki sekitar 800 diphones, dan Jerman sekitar 2500. Dalam sintesis diphone, hanya satu contoh dari setiap diphone terkandung dalam database pidato. Pada saat runtime, target prosodi kalimat ditumpangkan pada unit-unit minimal dengan cara pemrosesan sinyal digital teknik seperti linear predictive coding, PSOLA atau MBROLA. Diphone sintesis menderita gangguan sonik sintesis concatenative dan robot-terdengar sifat sintesis formant, dan memiliki beberapa keuntungan baik pendekatan lain dari ukuran kecil. Dengan demikian, penggunaannya dalam aplikasi komersial menurun, meskipun terus digunakan dalam penelitian karena ada beberapa implementasi perangkat lunak tersedia secara bebas.

Domain-spesifik sintesis

Domain-spesifik sintesis concatenates direkam sebelumnya kata dan frase untuk menciptakan ucapan-ucapan yang lengkap. Hal ini digunakan dalam aplikasi di mana berbagai teks output sistem akan terbatas pada domain tertentu, seperti jadwal angkutan pengumuman atau laporan cuaca. Teknologi ini sangat sederhana untuk menerapkan, dan telah digunakan secara komersial untuk waktu yang lama, dalam perangkat seperti berbicara jam dan kalkulator. Tingkat kealamian sistem ini bisa sangat tinggi karena berbagai jenis kalimat terbatas, dan mereka cocok dengan prosodi dan intonasi dari rekaman asli.

Karena sistem ini dibatasi oleh kata-kata dan frasa dalam database mereka, mereka tidak tujuan umum dan hanya dapat mensintesis kombinasi kata dan frase yang mereka telah terprogram. Campuran kata-kata dalam bahasa alami diucapkan namun masih dapat menyebabkan masalah kecuali banyak variasi diperhitungkan. Misalnya, dalam non-rhotic dialek dari bahasa Inggris “r” dalam kata-kata seperti “jelas” / klɪə / biasanya hanya diucapkan ketika kata berikut memiliki vokal sebagai huruf pertama (misalnya “membersihkan” direalisasikan sebagai / ɪ klɪərvʌt /). Demikian juga di Perancis , banyak konsonan akhir menjadi tidak lagi diam jika diikuti oleh sebuah kata yang dimulai dengan vokal, efek yang disebut penghubung . Ini pergantian tidak bisa direproduksi oleh sistem kata-Rangkaian sederhana, yang akan membutuhkan kompleksitas tambahan untuk konteks-sensitif .



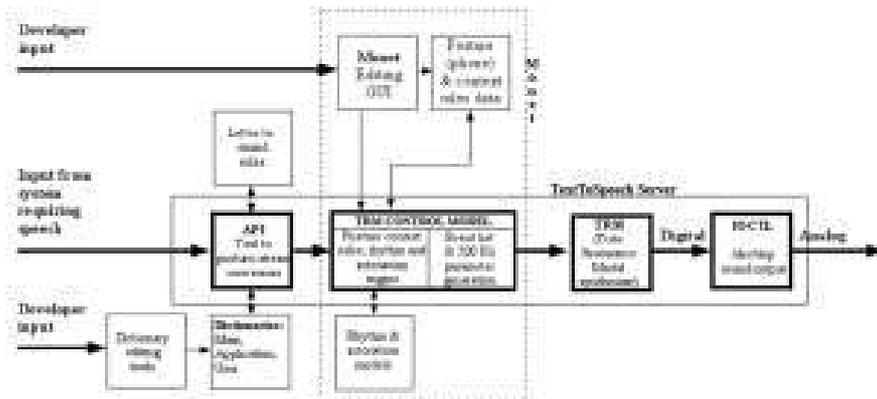
2. Formant Synthesis :

Formant synthesis tidak menggunakan pembicaraan manusia sebagai sample pada runtime. Daripada itu, synthesized speech yang dihasilkan dibuat dengan additive synthesis dan sebuah model akustik (physical modelling synthesis).

Parameter seperti frekuensi dasar, penyuaran, dan tingkat kebisingan di variasikan dari waktu ke waktu untuk menciptakan gelombang buatan (artificial) dari sebuah pembicaraan. Banyak sistem yang berdasarkan formant synthesis menciptakan pembicaraan yang seperti robot yang tidak mungkin dapat dikenal sebagai suara manusia. Tetapi, kealamian maksimum bukan selalu tujuan dari sebuah sistem speech synthesis, dan sistem formant synthesis mempunyai keuntungan dari sistem concatenative. Pembicaraan yang di-formant synthesis-kan dapat menjadi sangat jelas, bahkan dalam kecepatan yang tinggi, sehingga menghindari kesalahan suara yang sering dialami sistem concatenative.

Formant synthesis biasanya program yang lebih kecil dari concatenative sistem karena ia tidak menggunakan basis data dari sampel-sampel pembicaraan. Oleh karena itu formant synthesis dapat ditanamkan dalam sistem yang mempunyai memory dan microprosesor yang terbatas. Karena sistem yang berdasarkan formant mempunyai kendali penuh dari sluruh aspek dari hasil pembicaraan, variasi yang luas dari prosodi dan intonasi

dapat dihasilkan, menyampaikan tidak hanya pertanyaan dan pernyataan tetapi juga emosi dan nada suara.



Forman sintesis tidak menggunakan sampel suara manusia pada saat runtime. Sebaliknya, keluaran suara yang disintesis dibuat menggunakan aditif sintesis dan model akustik (sintesis pemodelan fisik). Parameter seperti frekuensi dasar , menyuarakan , dan kebisingan tingkat yang bervariasi dari waktu ke waktu untuk membuat gelombang pidato buatan. Metode ini kadang-kadang disebut aturan berbasis sintesis; Namun, banyak sistem concatenative juga memiliki aturan berbasis komponen. Banyak sistem yang didasarkan pada teknologi sintesis forman menghasilkan buatan, robot yang terdengar pidato yang tidak akan pernah salah untuk pidato manusia. Namun, kealamian maksimum tidak selalu tujuan sistem sintesis pidato, dan sistem sintesis forman memiliki keunggulan dibandingkan sistem concatenative. Pidato forman-disintesis dapat diandalkan dimengerti, bahkan pada kecepatan yang sangat tinggi, menghindari Glitches akustik yang biasanya wabah sistem

concatenative. Kecepatan tinggi disintesis pidato digunakan oleh tunanetra untuk navigasi cepat komputer menggunakan pembaca layar . Synthesizer formant adalah program biasanya lebih kecil dibandingkan dengan sistem concatenative karena mereka tidak memiliki database contoh pidato. Karena itu mereka dapat digunakan dalam embedded system , di mana memori dan mikroprosesor daya terutama terbatas. Karena sistem berbasis formant memiliki kontrol penuh dari semua aspek pidato output, berbagai prosodies dan intonasi dapat menjadi output, tidak hanya menyampaikan pertanyaan dan pernyataan, tetapi berbagai emosi dan nada suara.

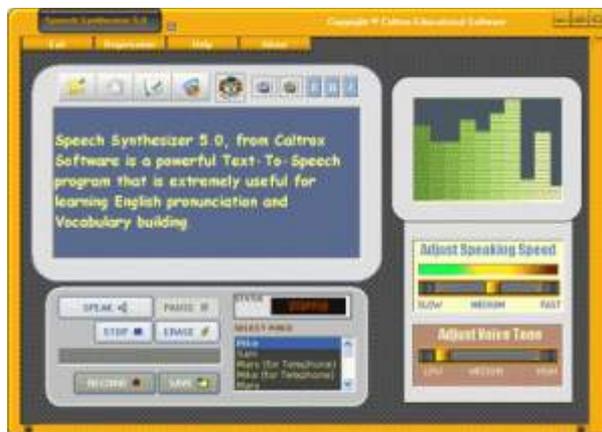
Contoh non-real-time tapi sangat akurat kontrol intonasi dalam sintesis formant meliputi pekerjaan yang dilakukan pada akhir tahun 1970 untuk Texas Instruments mainan Bicara & Eja , dan pada awal tahun 1980 Sega arcade mesin dan dalam banyak Atari, Inc. game arcade menggunakan TMS5220 LPC Chips . Menciptakan intonasi yang tepat untuk proyek ini adalah telaten, dan hasilnya masih harus dicocokkan dengan real-time text-to-speech interface.

3. Sintesis artikulatoris

Sintesis artikulatoris mengacu pada teknik komputasi untuk sintesis pidato berdasarkan model manusia saluran vokal dan artikulasi proses yang terjadi di sana. Synthesizer artikulatoris pertama teratur digunakan untuk percobaan laboratorium dikembangkan di Haskins Laboratories di pertengahan 1970-an

oleh Philip Rubin , Tom Baer, dan Paul Mermelstein. Synthesizer ini, yang dikenal sebagai ASY, didasarkan pada model saluran vokal dikembangkan di Bell Laboratories pada tahun 1960 dan 1970-an oleh Paul Mermelstein, Cecil Coker, dan rekan.

Sampai saat ini, model sintesis artikulatoris belum dimasukkan ke dalam sistem sintesis pidato komersial. Sebuah pengecualian adalah NeXT sistem berbasis awalnya dikembangkan dan dipasarkan oleh Trillium Suara Research, sebuah perusahaan spin-off dari University of Calgary , di mana banyak riset asli dilakukan. Setelah runtuhnya berbagai inkarnasi NeXT (dimulai oleh Steve Jobs pada akhir tahun 1980 dan bergabung dengan Apple Computer pada tahun 1997), perangkat lunak TRILLIUM diterbitkan di bawah GNU General Public License , dengan bekerja terus sebagai gnuttspeech . Sistem, pertama kali dipasarkan pada tahun 1994, memberikan penuh text-to-speech konversi berbasis artikulatoris menggunakan Waveguide atau transmisi-line analog dari saluran mulut dan hidung manusia dikendalikan oleh Carré ini “model daerah khas”.



4. Sintesis berbasis HMM

Sintesis berbasis HMM adalah metode sintesis berdasarkan model Markov tersembunyi , juga disebut statistik Parametrik Sintesis. Dalam sistem ini, spektrum frekuensi (vokal),frekuensi dasar (sumber vokal), dan durasi (prosodi) berbicara dimodelkan secara bersamaan oleh HMMs. Pidato bentuk gelombang yang dihasilkan dari HMMs sendiri berdasarkan maksimum kriteria.

5. Sintesis Sinewave

Sintesis sinewave adalah teknik untuk sintesis pidato dengan mengganti formant (band utama energi) dengan peluit nada murni.

Ada beberapa masalah yang terdapat pada pemaduan suara, yaitu:

1. User sangat sensitif terhadap variasi dan informasi suara. Oleh sebab itu, mereka tidak dapat memberikan toleransi atas ketidaksempurnaan pemadu suara.
2. Output dalam bentuk suara tidak dapat diulang atau dicari dengan mudah.
3. Meningkatkan keberisikan pada lingkungan kantor atau jika menggunakan handphone, maka akan meningkatkan biaya pengeluaran.



Lingkungan dari aplikasi pemadu suara adalah:

Bagi tunanetra, pemadu suara menawarkan media komunikasi dimana mereka dapat memiliki akses yang tidak terbatas. Lingkungan dimana visual dan haptic skill user berfokus pada hal lain. Contohnya: sinyal bahaya pada kokpit pesawat udara.

BAB 10

DIPHONE STUDIO (*DIPHONE STUDIO*)

A. Diphone Studio (*Definitions of Diphone Studio*)

Diphone Studio adalah sekelompok beberapa aplikasi. Fungsi utamanya adalah untuk memberikan kontrol yang kuat untuk sintesis suara. Beberapa analisis / model sintesis dapat diakses. Mereka terintegrasi ke dalam Diphone Studio dalam bentuk program eksternal yang disebut model atau plugin. Ini adalah daftar mereka: • Analisis dan model sintesis aditif yang disebut Additive; • Model resonansi yang digunakan untuk analisis dan sintesis menggunakan fungsi gelombang formantik (FOF). Model ini disebut Chant; • Model terakhir yang digunakan untuk pemrosesan file suara eksternal yang lebih mudah, yang disebut Direct Signal. Aplikasi utama, Diphone, berkomunikasi dengan program analisis dan sintesis, Addan dan Resan, dan menggunakan data mereka untuk membuat interpolasi. data menjadi dapat diakses dalam bentuk representasi grafis (kurva Bpf, slotted "tiles" ...). Pengguna memanipulasi elemen-elemen ini saat ia ingin akhirnya mendapatkan urutan musik yang disimpan ke hard disk komputer. Manual ini sesuai dengan versi: • 2,8 dari aplikasi Diphone • 1,8 dari aplikasi Addan • 1.0 dari aplikasi Resan (pengiriman publik pertama dari aplikasi ini). Ini ditata dengan cara berikut. • Analisis menggunakan model aditif: program Addan. • Presentasi program Diphone: - Pembuatan skrip, kamus dan urutan- Model Sinyal Langsung- Model Nyanyian. • Program

Resan. • Menu dan preferensi. n Diphone, model mana pun yang digunakan, sebagian besar manipulasi tetap serupa untuk pengguna. Dalam manual ini, tinjauan terperinci atas fungsi-fungsi program dibuat untuk model Aditif. Poin yang spesifik untuk model Sinyal Langsung dan Nyanyian diungkapkan, dan rujukan silang diberikan untuk manipulasi umum.

Dalam pidato, diphone didefinisikan sebagai transisi antara dua fonem. Dalam konteks musik, transisi ini dapat dibuat dari dua suara vokal, tetapi lebih umum suara jenis apa pun, baik instrumental, vokal, atau rekaman «objek suara». Dengan demikian, definisi dari «diphone» musikal juga dapat diperluas untuk memasukkan suara atau kesunyian tunggal yang stabil. Gagasan sintesis menggunakan diphons dikandung pada akhir 1980-an oleh Xavier Rodet. frase menggunakan transient dan stablesound. Dengan menggunakan teknik studio analog tradisional, serangkaian suara sementara dan stabil dapat digabungkan, atau bersama-sama, dengan menyambung potongan-potongan kecil ujung ke ujung. Di studio digital saat ini, ini dilakukan oleh Cross-Fading. Dalam kedua kasus, hasilnya adalah bahwa itu bisa berakhir terdengar jauh dari meyakinkan karena konteks batin dari suara yang disambung atau pudar umumnya tidak cocok. Dalam sistem kontrol dan sintesis iPhone umum Rodet, suara dianalisis, dan dibagi dengan menginterpolasi nilai-nilai antara segmen-segmen analisis data yang berdekatan. Hasil penelitian ini pertama kali diterapkan pada workstation UNIX pada tahun 1988 oleh Xavier Rodet dan Philippe Depalle us-ing.

sintesis sumber-filter dan, kemudian, sintesis aditif. Hanya ada antarmuka pengguna grafis yang sangat sederhana yang diizinkan untuk menganalisis data secara berturut-turut di layar. Versi Diphone Macintosh yang pertama, diselesaikan pada tahun 1996 yang diprogram oleh AdrienLefevre

B. Perekaman Suara Dalam Bentuk Kata-Kata Yang Diucapkan (*Voice Recording In The Form Of Spoken Words*)

Google memperbarui (update) aplikasi perkantoran Docs, yang terintegrasi di layanan Google Drive. Dalam update terbaru ini, aplikasi tersebut hadir dengan berbagai fitur yang menarik. Salah satu fitur menarik adalah Voice Typing. Ya, sesuai dengan namanya, pengguna kini bisa mengetik hanya dengan menggunakan suara (dictation). Artinya, suara yang diucapkan akan langsung dibuah menjadi teks atau tulisan. Untuk menggunakan fitur tersebut, pengguna hanya perlu mengklik menu Tools, kemudian mengklik Voice Typing. Setelah itu, pengguna bisa langsung mengucapkan kata-kata yang ingin diketik oleh sistem. Menariknya, sistem dari Google Docs ini sudah mengenali berbagai macam bahasa. Sebagaimana KompasTekno rangkum dari blog resmi Google, Kamis (3/9/2015), total sudah ada 40 bahasa yang didukung, termasuk Indonesia. Saat KompasTekno mencobanya, sistem mampu mengubah suara yang diucapkan dengan cukup baik. Kesalahan yang terjadi cukup jarang. "Voice Typing tersedia lebih dari 40

bahasa, jadi kami bisa membantu Anda dengan PR bahasa Perancis. Voila," tulis Google. Fitur menarik lainnya disebut sebagai Research. Menggunakan fitur tersebut, pengguna bisa melakukan pencarian langsung dari Google Docs. Artinya, pengguna sudah tidak perlu lagi berpindah aplikasi ke layanan Google Search untuk melakukan hal tersebut. Hasil pencariannya pun bisa langsung dimasukkan ke Google Docs. Sebagai contoh, pengguna melakukan pencarian tentang mobil. Melalui menu Research, langsung saja masukkan kata mobil. Hasil pencariannya dapat berupa kata-kata dan juga gambar. Kemudian, tinggal pilih gambar yang diinginkan, kemudian tekan tombol Insert. Secara otomatis gambar tersebut akan masuk dalam dokumen yang dikerjakan. Fitur Insert tersebut juga berlaku untuk kata-kata. Pengguna tinggal blokir kata yang diinginkan, lalu menekan tombol Insert. Kata-kata tersebut secara otomatis akan dimasukkan ke dalam dokumen. Tombol Research sendiri dapat ditemukan di menu Tools. "Anda bisa melakukan pencarian tanpa harus meninggalkan Docs, dan saat telah menemukan quote, fakta atau gambar, bisa menambahkannya (ke dalam dokumen) hanya dengan beberapa sentuhan," ujar Google. Update ini sendiri bisa ditemukan di Google Drive versi desktop dan perangkat berbasis Android.

C. Penentuan Diphone Dalam Suatu Kata (*Determination of Diphone In A Word*)

Penelitian di bidang pensintesa ucapan mengalami perjalanan yang sangat panjang dan telah dimulai sejak lama. Salah satu catatan literatur awal yang berhubungan dengan sintesa ucapan adalah pernyataan seorang ahli matematika dan engineer terkenal yang bernama Leonhard Euler pada tahun 1761. Euler menyatakan “It would be a considerable invention indeed, that of a machine able to mimic speech, with its sounds and articulations. I think it is not imposible”.

Berdasarkan hasil studi literatur dari berbagai sumber bacaan, perkembangan teknologi pensintesa ucapan dapat dibagi menjadi tiga kurun waktu. Kurun waktu pertama adalah sebelum 1930. Pada masa ini penelitian-penelitian lebih banyak ditujukan untuk memahami karakteristik sinyal ucapan serta pengembangan pensintesa ucapan berbasis mekanik atau elektromekanik. Kurun waktu kedua dimulai sejak tahun 1930-an sampai dengan ditemukannya komputer digital. Masa ini ditandai dengan pengembangan berbagai alat pensintesa ucapan menggunakan teknologi elektronik analog. Kurun waktu ketiga dimulai sejak ditemukannya komputer digital hingga sekarang. Pada masa ini, sintesa ucapan dilakukan menggunakan pendekatan-pendekatan pemrosesan sinyal digital.

Kurun Waktu Pertama

Penelitian tentang ucapan dimulai dengan penelitian-penelitian untuk melakukan pemahaman tentang sinyal ucapan. Pada tahun 1779, Imperial Academy of St. Petersburg menyelenggarakan suatu kompetisi dengan tujuan untuk mengetahui hal-hal berikut [Pel93].

1. "What is the nature and character of the sounds of the vowels a, e, i, o, u that make them different from one another?"
2. "Can an instrument be constructed like the vox humana pipes of an organ, which shall accurately express the sounds of the vowels?"

Seorang peneliti dari Rusia yang bernama Christian Gottlieb Kratzenstein memenangkan kompetisi tersebut dengan membuat satu set resonator akustik yang mensimulasikan mulut manusia. Resonator Kratzenstein terdiri dari 5 bentuk tabung, masing-masing untuk mensimulasikan satu bunyi vokal.

Robert Willis, pada tahun 1829 melakukan penelitian yang berhasil memperlihatkan bahwa sintesa ucapan yang dihasilkan oleh Kratzenstein dapat pula dicapai dengan hasil yang sama menggunakan tabung tunggal yang dapat diatur panjangnya. Selama dua dekade, antara tahun 1769 sampai dengan 1790, Wolfgang Ritter von Kempelen telah menghasilkan speaking machine yang lengkap. Pada prakteknya, Wolfgang telah membuat 3 model yang berbeda, semuanya dioperasikan dengan

tangan. Penemuannya dipublikasikan dalam bentuk buku pada tahun 1791.

Wolfgang von Kempelen berpendapat bahwa untuk membuat mesin yang dapat berbicara, pertama-tama harus dapat menghasilkan suara vokal. Wolfgang mulai dengan mencari sumber bunyi yang sesuai, yaitu suatu substitusi mekanik dari suara vokal. Dia mencoba menggunakan reed bergetar yang biasa digunakan dalam instrumen musik, walaupun hasilnya kurang memuaskan. Suara dari reed disalurkan melalui suatu alat berbentuk bel yang dilengkapi baffle pada mulut yang dapat digerakan untuk menghasilkan bunyi vokal yang berbeda. Tidak puas dengan hasil percobaannya yang pertama, von Kempelen menggunakan tangannya untuk menggantikan baffle. Meskipun hasilnya menjadi lebih baik, tetapi suara yang dihasilkan masih belum memuaskan.

Model yang kedua dirancang untuk memenuhi kebutuhan akan perlunya beberapa resonansi pada beberapa frekuensi yang berbeda untuk mencapai berbagai suara berlainan yang diinginkan. Versi ini bersifat modular, berupa tiga belas buah resonator yang masing-masing dilengkapi dengan reed dan bersifat dapat dibongkar pasang, sehingga dapat saling dipertukarkan. Gambar 2.2 memperlihatkan model tersebut.

Dengan mesin tersebut, von Kempelen mengklaim bahwa dia telah mampu menghasilkan suara vokal a, o dan u serta suara p, m dan l yang dapat diterima. Secara monotonik, mesin buatannya dapat mengucapkan suara seperti “mama” dan “papa”, tetapi

masih menghadapi dua masalah utama. Pertama, suara vokal yang dihasilkan mengandung bunyi yang sifatnya eksplosif yang mirip bunyi “k”. Masalah lain yang dihadapi adalah transisi antara dua bunyi yang berdekatan yang tidak smooth seperti suara alami. Satu bunyi dengan bunyi berikutnya masih terasa sebagai dua bunyi yang terpisah. Untuk mengatasi masalah tersebut, dia menambahkan kulit halus pada reed, juga menggunakan reed tunggal sebagai pengganti dari sejumlah reed yang sebelumnya digunakan pada setiap resonator.

Mesin ketiga buatan von Kempelen secara fisik sangat berbeda dari mesin-mesin sebelumnya (lihat Gambar 2.3). Paru-paru disimulasikan dengan pompa yang digerakan dengan bahu yang secara kontinyu dapat menghembuskan udara. Vokal dapat dihasilkan dengan cara menutup “nostrils” mesin tersebut dengan tangan kanan sambil menghembuskan udara dari simulator paru-paru. Sementara itu, tangan kiri harus mengatur resonansi melalui alat berbentuk bel. Hanya orang yang terlatih memainkannya yang dapat menghasilkan bunyi-bunyi yang diharapkan. Suara seperti F, H, V, W dan beberapa lainnya adalah suara-suara yang juga dapat dihasilkan dengan mesin tersebut. Wolfgang mengklaim bahwa mesin ketiga buatannya dapat menghasilkan semua suara vokal serta sembilan belas konsonan. Meskipun mesin tersebut memiliki kapasitas menghasilkan udara sekitar enam kali lebih besar dari kapasitas paru-paru manusia, tetapi mesin ini hanya mampu mengucapkan kalimat yang pendek sebelum kehabisan udara. Pada tahun 1791 von Kempelen

mempublikasikan hasil penelitiannya dalam bahasa Jerman dan Perancis dengan judul “Mechanismus der menschlichen Sprache nebst der Beschreibung seiner sprechenden Maschine”.

Di Perancis, pada waktu yang hampir bersamaan dengan von Kempelen, Abbe' Mical mengembangkan mesin lain yang dikenal sebagai “two talking head”. Mesin ini terdiri dari dua silinder yang mirip dengan silinder yang biasa kita lihat pada instrumen musik. Satu silinder disediakan untuk memainkan sejumlah ucapan tertentu dengan prosodinya. Silinder lainnya digunakan untuk menghasilkan semua bunyi dalam bahasa Perancis. Tidak diketahui dengan pasti otentikasi mesin buatannya tersebut.

Hermann Helmholtz, seorang perintis peneliti akustik, pada pertengahan abad ke-19 membuat perangkat elektro-mekanik yang terdiri dari sejumlah garpu yang dapat ditala, kumparan elektrik, dan sejumlah resonator yang dapat mensintesa suara komposit yang sangat mirip suara vokal manusia. Perangkat ini mungkin tidak memperlihatkan hubungan langsung dengan berbagai penemuan alat-alat lainnya yang berhubungan dengan aplikasi suara, tetapi keberadaan mesin tersebut memberikan ilham bagi Alexander Graham Bell yang menghasilkan beberapa penemuan di bidang aplikasi mesin yang berhubungan dengan suara manusia. Pada saat yang bersamaan juga, Hermann Helmholtz telah melakukan berbagai penelitian yang memberikan pemahaman yang lebih mendalam tentang akustik.

Peranan Sir Charles Wheatstone yang lebih dikenal dengan “Jembatan Wheatstone”-nya tidak dapat diabaikan dalam perkembangan alat pensintesa ucapan manusia. Wheatstone tumbuh besar dalam keluarga yang melakukan bisnis perangkat musik di London. Tahun 1821, pada usia sembilan belas tahun ia mendemonstrasikan alat ciptaannya yang dapat menggetarkan batang logam yang dieksitasi oleh suatu sumber yang vibrasinya dirambatkan melalui konduktor yang padat. Pada tahun 1835, Wheatstone mendemonstrasikan ciptaannya kepada Dublin Association.

Alexander Graham Bell yang lahir di Edinburg pada tahun 1846 dikenal sebagai penemu telpon. Berdasarkan buku yang ditulis oleh Kempelen, Bell beserta dua saudaranya (Melly dan Ted) pernah melakukan pengembangan mesin yang dapat menirukan ucapan-ucapan manusia. Pengembangan tersebut dilakukan di Edinburg sekitar tahun 1863. Pada usia 19 tahun, Bell mencoba mengulangi penelitian akustik Helmholtz. Bell mengira bahwa garpu tala dapat mentransmisikan bunyi vokal secara elektrik. Untuk memperbaiki kesalahan dugaan tersebut, akhirnya dia menemukan suatu keyakinan bahwa suara apapun dapat ditransmisikan secara elektrik. Pada akhirnya, Bell berhasil menemukan telpon.

Pada awal tahun 1990-an, J. L. Flanagan melaporkan hasil kerjanya yang merupakan kelanjutan dari pemikiran Helmholtz dan menguji berbagai alat yang dapat melakukan sintesa suara vokal. Penelitian ini meliputi penggunaan pipa organ, multiple

sirens, garpu vibrasi yang dapat ditala, serta ide R. R. Riesz yang pada tahun 1937 mengusulkan alat bicara mekanik yang dapat dioperasikan dengan jari-jari tangan.

BAB 11

PROYEK AKHIR PENGENALAN UCAPAN DAN TEKS KE UCAPAN

(FINAL RECOGNITION AND TEXT PROJECT TO RECEIVE)

A. Proyek Akhir (*Final Project*)

Komunikasi bahasa antar manusia dengan manusia merupakan salah satu cara yang paling efektif untuk menyampaikan maksud dan tujuan seseorang dalam menyampaikan informasi untuk memudahkan seseorang dalam menyelesaikan pekerjaan. Informasi dapat disampaikan dalam bentuk teks, citra, suara, dan lain-lain. Pemakaian teknologi komputer untuk menjalin komunikasi manusia dengan komputer sudah menjadi suatu kebutuhan, jika komputer mengerti ucapan yang dimaksudkan manusia bisamenjadi suatu kemudahan dalam pengoperasian komputer, seperti voice command, akses kontrol sistem berbasis suara, dan identifikasi suara untuk keamanan sistem.

Perkembangan speech recognition (speech to text) berjalan cukup pesat pada saat ini dilihat dari banyaknya jurnal yang membahas mengenai speech to text. Suara manusia mempunyai karakteristik yang sangat kompleks, satu kata yang diucapkan oleh orang yang berbeda bisamenghasilkan karakteristik suara yang berbeda, namun suatu sistem diharuskan dapat mengenali sebagai suatu kata yang sama. Selain itu faktor yang mempengaruhi suara adalah kesehatan, psikologi, umur, dan

jenis kelamin seseorang. Speech to text memungkinkan suatu perangkat untuk mengenali dan memahami kata-kata yang diucapkan dengan cara digitalisasi kata dan mencocokkannya dengan suatu pola tertentu yang tersimpan dalam suatu perangkat.

Kata-kata yang diucapkan diubah bentuknya menjadi sinyal digital dengan cara mengubah gelombang suara menjadi sekumpulan angka yang kemudian disesuaikan dengan kode-kode tertentu untuk mengidentifikasi kata-kata tersebut. Hasil dari identifikasi kata yang diucapkan dapat ditampilkan dalam bentuk tulisan. Di Indonesia penelitian mengenai speech to text sudah mulai banyak dilakukan dilihat dari bermunculannya jurnal-jurnal mengenai speech to text bahasa Indonesia.

Dari penelitian sebelumnya ada yang menggunakan metode ekstraksi ciri Mel Frequency Cepstral Coefficients (MFCC) dan metode pengenalan pola Hidden Markov Model (HMM) (Fawziah, et al., 2013), namun masih terbatas untuk speaker laki-laki saja dan hanya beberapa kata saja yang menjadi data pengujian. Dalam penelitian kali ini dicoba untuk menggunakan metode yang sama dengan menggunakan speaker laki-laki dan perempuan dengan data pengujian yang lebih banyak.

Mel Frequency Cepstrum Coefficients (MFCC) feature extraction mengkonversikan sinyal suara ke dalam beberapa vektor data berguna bagi proses pengenalan suara. Terdapat 7 tahapan dalam MFCC yaitu Pre Emphasize, Frame Blocking, Windowing, Fast Fourier Transform, Mel Frequency Wrapping,

Discrete Cosine Transform, dan Cepstral Liftering. Metode ini memiliki beberapa kelebihan diantaranya adalah mampu menangkap informasi penting dalam sinyal suara, menghasilkan ciri sinyal ucapan seminimal mungkin tanpa menghilangkan informasi-informasi yang ada, dan mereplikasikan organ pendengaran manusia dalam melakukan persepsi terhadap sinyal suara (Andriana, 2011).

Hidden Markov Model (HMM) merupakan suatu metode pendekatan yang dapat mengelompokkan sifat-sifat spectral dari tiap bagian suara dengan beberapa pola. HMM memiliki 5 proses dasar dalam melakukan pengenalan suara, yaitu: Feature Analysis, Unit Matching System, Lexical Decoding, Syntactic Analysis, and Semantic Analysis. Proses-proses itulah yang menyebabkan HMM mempunyai tingkat akurasi yang lebih tinggi dibanding metode lain, terbukti dengan banyaknya penelitian mengenai speech recognition yang menggunakan metode ini untuk pendekatan dalam mengenali pola suara (Rabiner & Juang, 1993)

B. Pengenalan Ucapan dan Teks (*Introduction to Speech and Text*)

Proses pengenalan suara dilakukan oleh komponen perangkat lunak yang dikenal sebagai mesin pengenalan ucapan. Fungsi utama dari mesin pengenalan ucapan adalah mengolah kata-kata input dan menerjemahkannya ke dalam teks yang dimengerti oleh sebuah aplikasi. Aplikasi bisa menafsirkan hasil pengakuan

sebagai perintah. Dalam kasus ini, aplikasi adalah aplikasi perintah dan kontrol. Jika aplikasi menangani teks yang dikenali sebagai teks, maka aplikasi ini dianggap sebagai aplikasi pendiktean. Pengguna berbicara kepada komputer melalui mikrofon, yang pada gilirannya mengidentifikasi makna kata-kata dan mengirimkannya ke perangkat NLP untuk diproses lebih lanjut. Setelah dikenali, kata-kata itu bisa digunakan dalam berbagai aplikasi seperti display, robotika, perintah ke komputer, dan dikte.

Pada dasarnya, pendekatan ini berurusan dengan konversi kata-kata yang diucapkan menjadi teks. Pengenalan ucapan juga disebut ASR (otomatis speech recognition), STT (speech to text) atau hanya pengenalan suara komputer. Pengenalan ucapan juga dapat dipahami sebagai bidang ilmu komputer, yang berhubungan dengan perancangan dan pengembangan sistem komputer, untuk mengenali kata-kata yang diucapkan.

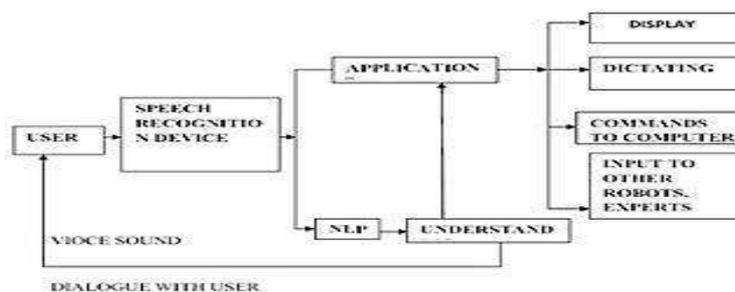
Metode pengenalan ucapan lainnya, yaitu digunakan untuk menguraikan kata adalah Kecerdasan buatan (AI) bisa jadi dipahami sebagai kombinasi dari pengenalan pola pendekatan dan pendekatan fonetis akustik. kecerdasan buatan adalah teknik yang paling berkembang dan efektif, yang mendukung pengenalan suara tanpa cela dan akurat. Ini karena; kecerdasan buatan menggabungkan pendekatan algoritmik tertentu, yang mendorong konversi koheren dan transformasi ucapan ke dalam pola yang mudah dibaca, dan sebaliknya.

Speech Recognition

1. Pengenalan Speech Recognition

Pengguna berbicara ke komputer melalui mikrofon, yang pada gilirannya, mengidentifikasi themdaing kata-kata dan mengirimkannya ke perangkat NLP untuk diproses lebih lanjut. Setelah dikenali, kata-kata tersebut dapat digunakan dalam berbagai aplikasi seperti display, robotika, perintah ke komputer, dan pendiktean. Kata recognizer adalah sistem pengenalan ucapan yang mengidentifikasi kata-kata individual. Sistem perintis awal hanya bisa mengenali huruf dan angka individual. Saat ini, mayoritas sistem pengenalan kata adalah pengenal kata dan memiliki lebih dari 95% pengakuan. Sistem seperti itu mampu mengenali kosa kata kecil dari kata-kata tunggal atau ungkapan-ungkapan sederhana. Seseorang harus mengucapkan informasi masukan dengan kata-kata lisan yang jelas, dengan jeda di antara kata-kata, untuk memasukkan data di komputer. Pengenal ucapan terus menerus jauh lebih sulit dibangun daripada pengenal kata. Anda berbicara kalimat lengkap ke komputer. Masukan akan dikenali dan kemudian diproses oleh NLP.

2. Proses Speech Recognition



Gambar 3.1 Proses pengenalan suara

Setelah proses pelatihan, kata-kata yang diucapkan pengguna akan menghasilkan teks; Keakuratan ini akan membaik dengan dikte lebih lanjut dan penggunaan prosedur koreksi secara teliti. Dengan sistem yang terlatih, sekitar 95% kata yang diucapkan bisa ditafsirkan dengan benar. Sistem dapat dilatih untuk mengidentifikasi kata-kata dan ungkapan tertentu dan memeriksa dokumen standar pengguna untuk mengembangkan file suara yang akurat bagi individu. Namun, masih banyak faktor lain yang perlu dipertimbangkan untuk mencapai tingkat pengenalan yang tinggi. Tidak ada keraguan bahwa perangkat lunak itu bekerja dan dapat membebaskan banyak peserta didik, namun prosesnya bisa jauh lebih memakan waktu daripada yang dapat dihargai pengguna pertama kali dan hasilnya seringkali bisa menjadi miskin. Ini bisa sangat mendemotivasi, dan banyak pengguna menyerah pada tahap ini. Dukungan kualitas dari seseorang yang mampu menunjukkan kepada pengguna cara paling efektif menggunakan perangkat lunak sangat penting.

Model Statistik Speech Recognition

1. Model Akustik (AM)

Salah satu model yang paling menonjol dan diadopsi secara luas Pengenalan suara adalah model akustik (AM). Telah didirikan bahwa model akustik menangkap ucapan pengakuan karakteristik unit pengakuan dasar. Menurut unit pengakuan dapat berada pada tingkat fonem, tingkat suku kata, dan pada tingkat kata.

Beberapa kekurangan dan kendala menjadi pertimbangan dengan pemilihan masing-masing dari unit ini Referensi telah mengklaim bahwa untuk LVCSR (kosakata besar pengenalan suara terus-menerus) sistem, fonem adalah unit yang paling menguntungkan. Model Markov tersembunyi dan jaringan syaraf tiruan (NN) adalah pendekatan yang diadopsi secara luas, yang sedang digunakan untuk pemodelan akustik pidato sistem pengenalan

2. Model Bahasa (LM)

Model bahasa adalah statistik lain yang paling signifikan model pengenalan suara Salah satu tujuan utama Model bahasa adalah untuk menyampaikan atau mentransmisikan tingkah laku bahasa. Hal ini disebabkan oleh fakta bahwa ia bermaksud untuk meramalkan adanya urutan kata tertentu dalam target pidato. Dari aspek mesin, model pengenalan suara ini membantu meminimalkan ruang pencarian yang handal dan kredibel kombinasi kata-kata Penting untuk memperhatikan bahasa itu model dikembangkan dengan bantuan CMU statistik LM toolkit.

3. Model Leksikon

Telah diklaim oleh model leksikon yang menyediakan pengucapan kata-kata dalam pidato target, yang telah untuk diakui Sesuai dengan persepsi, Model leksikon memainkan peran yang tak terelakkan dan tak terpisahkan pengenalan ucapan otomatis Hal ini disebabkan fakta bahwa operasi model leksikal

didasarkan pada dua parameter, yaitu, akses keseluruhan kata, dan dekomposisi keseluruhan pidato ke dalam potongan kecil. Proses ini akhirnya menghasilkan sesuai pengakuan pidato. Misalnya, jika speech recognition Model dalam bahasa asli, model leksikon harus ada dirumuskan dalam bahasa asli, untuk mendapatkan hasil yang berharga dan bermanfaat. Dalam hal ini, buatan netral Fonem jaringan dapat dianggap sebagai salah satu yang terbesar pendekatan, karena membantu mengembangkan leksikon asli dari leksikon asing; Oleh karena itu, menghasilkan pemetaan telepon Bahasa Inggris ke telepon bahasa asli. Itu penting untuk memperhatikan bahwa keseluruhan proses dilakukan, sementara mengingat informasi kontekstual.

4. Hidden Markov Model (HMM)

Markov adalah alat statistik yang paling populer, yang digunakan untuk pemodelan data telah dianalisis bahwa HMM telah memainkan peran terpuji dalam mengurangi masalah klasifikasi ucapan, yang merupakan salah satu isu utama, dalam pendekatan pengenalan suara. HMM memasukkan berbagai isu, yang digunakan untuk mempengaruhi keakuratan pengenalan suara. Untuk menyelesaikan masalah tersebut, algoritma proyeksi subruang dan tertimbang tersembunyi model Markov diusulkan.

Tools

1. Pencarian dan optimasi

Banyak masalah dalam AI dapat dipecahkan secara teori dengan mencari secara cerdas melalui banyak solusi yang mungkin. Penalaran dapat dikurangi untuk melakukan pencarian. Misalnya, bukti logis dapat dilihat sebagai mencari jalan yang mengarah dari premis ke kesimpulan, di mana setiap langkah adalah penerapan aturan inferensi. Algoritma perencanaan mencari melalui pohon tujuan dan sasaran, mencoba menemukan jalan menuju sasaran, sebuah proses yang disebut analisis berarti-akhir. Algoritma robot untuk memindahkan anggota badan dan benda tangkapan menggunakan penelusuran lokal di ruang konfigurasi. Banyak algoritma pembelajaran menggunakan algoritma pencarian berdasarkan optimasi.

2. Logika

Beberapa bentuk logika yang berbeda digunakan dalam penelitian AI. Logika proposisional atau sentimental adalah logika pernyataan yang benar atau salah. Logika orde pertama juga memungkinkan penggunaan quantifier dan predikat, dan dapat mengungkapkan fakta tentang objek, sifat mereka, dan hubungannya satu sama lain. Logika fuzzy, adalah versi logika orde pertama yang memungkinkan kebenaran astatement diwakili sebagai nilai antara 0 dan 1, bukan hanya True (1) atau False (0). Sistem fuzzy dapat digunakan untuk penalaran yang tidak pasti

dan telah banyak digunakan dalam sistem kontrol produk industri modern dan konsumen.

3. Metode probabilistik untuk penalaran yang tidak pasti

Banyak masalah di AI (dalam penalaran, perencanaan, pembelajaran, persepsi dan robotika) mewajibkan agen untuk beroperasi dengan informasi yang tidak lengkap atau tidak pasti. Dimulai pada akhir tahun 80an dan awal tahun 90an, Judea Pearl dan yang lainnya memperjuangkan penggunaan metode yang diambil dari teori probabilitas dan ekonomi untuk merancang sejumlah alat canggih untuk memecahkan masalah ini.

4. Klasifikasi dan metode pembelajaran statistik

Aplikasi AI yang paling sederhana dapat dibagi menjadi dua jenis: pengklasifikasi ("jika thendiamond mengkilap") dan pengendali ("jika mengkilap lalu mengambil"). Controller bagaimanapun juga mengklasifikasikan kondisi sebelum mengambil tindakan, dan karena itu klasifikasi merupakan bagian utama dari banyak sistem AI. Classifier adalah fungsi yang menggunakan pencocokan pola untuk menentukan kecocokan terdekat. Mereka dapat disetel sesuai contoh, membuatnya sangat menarik untuk digunakan di AI. Ini examplesare dikenal sebagai pengamatan atau pola. Dalam pembelajaran yang diawasi, setiap pola termasuk dalam kelas yang telah ditentukan sebelumnya. Sebuah kelas bisa dilihat sebagai keputusan yang harus dibuat. Semua pengamatan yang disertakan dengan label

kelas mereka dikenal sebagai kumpulan data. Bila pengamatan baru diterima, pengamatan tersebut dikelompokkan berdasarkan pengalaman sebelumnya.

5. Jaringan syaraf tiruan

Jaringan syaraf tiruan adalah kelompok node yang saling berhubungan, mirip dengan jaringan neuron yang luas di otak manusia. Studi jaringan syaraf tiruan dimulai pada dekade sebelum penelitian AI dimulai, dalam karya Walter Pitts dan Warren McCullough. Periset penting lainnya adalah Frank Rosenblatt, yang menemukan perceptron dan Paul Werbos yang mengembangkan algoritma propagasi balik.

Metode Speech Recognition

1. Natural Language Processing

Natural Language Processing memberi mesin kemampuan untuk membaca dan memahami bahasa yang digunakan manusia untuk berbicara. Banyak peneliti berharap bahwa sistem pemrosesan natural language yang cukup kuat akan dapat memperoleh pengetahuan dengan sendirinya, dengan membaca teks yang ada yang tersedia melalui internet. Beberapa aplikasi pemrosesan bahasa alami yang mudah dilakukan mencakup pencarian informasi (atau penambahan teks) dan terjemahan mesin.

2. Artificial Intelligence

Teknologi telah dikembangkan oleh para peneliti, yang telah memungkinkan untuk mencapai ketepatan kata-kata yang masuk akal. Tepatnya, muncul pendekatan dan paradigma teknologi memainkan peran terpuji untuk terus meningkatkan integritas pengenalan ucapan. Sebaliknya telah menyatakan fakta bahwa teknologi ini tidak mampu cukup bersaing dengan akurasi pendengar manusia. Oleh karena itu, ini adalah salah satu tugas yang paling menantang bagi peneliti merancang dan mengembangkan tanpa cacat dan sangat efisien teknik pengenalan suara. Dalam keadaan seperti itu, Pendekatan kecerdasan buatan bisa dianggap sebagai salah satu kesempatan terbesar, dalam hal mengenali pola pidato, secara akurat. Hal ini disebabkan fakta bahwa buatan kecerdasan mengubah pidato menjadi terstruktur dengan baik algoritma, dengan tepat mengikuti semua tahap. Tahapan yang paling penting, yang terlibat dalam pidato Pengakuan melalui kecerdasan buatan mencakup representasi unit pidato, formulasi dan pengembangan dari algoritma pengenalan, serta demonstrasi yang benar masukan (ucapan).

3. Pattern Recognition

Pengenalan pola ditemukan paling umum dan teknik pengenalan ucapan yang banyak diadopsi. Metode ini terutama menggabungkan dua langkah penting, termasuk pola perbandingan dan latihan pola. Sejak didirikan studi bahwa

karakteristik utama ini Metode yang digunakan adalah terstruktur dengan baik dan terpadu kerangka matematis. Kerangka matematika ini membantu dalam merumuskan representasi pidato yang konsisten pola; maka berakibat pada perolehan yang lebih akurat hasil.

4. Acoustic Phonetic

Pendekatan paling primitif Speech Recognition terutama didasarkan pada proses pencarian suara dan kata. Salah satu tujuan utama semacam itu kegiatan adalah untuk memberikan label yang memadai untuk sampel suara, untuk mengenali pola suara. Ini penting untuk memperhatikan bahwa metode tersebut ditemukan sebagai dasar dari pendekatan fonetis akustik. Sesuai dengan konsep akustik Pendekatan fonetik, ada fonem (unit fonetik) dan unit terbatas dalam bahasa lisan. Unit akustik ini Pendekatan fonetik secara luas dikategorikan oleh koleksi sifat akustik yang biasanya terlihat di sinyal ucapan.

5. Aplikasi

Telah diamati dari evaluasi penelitian, yang disajikan bahwa kecerdasan buatan saat ini sedang digunakan di berbagai bidang kehidupan, termasuk penemuan ilmiah, penginderaan jarak jauh, transportasi, penerbangan, hukum, kontrol robot, perdagangan saham, diagnosa medis, dan bahkan mainan. Namun, satu dari aplikasi kecerdasan buatan yang paling menonjol adalah pengenalan suara. Studi tentang pengenalan

suara menunjukkan bahwa pendekatannya secara luas digunakan pada mesin penjawab pusat layanan pelanggan dan call center. Dalam akun ini, mengklaim bahwa perangkat lunak pengenalan suara memungkinkan komputer untuk menangani tingkat pertama pengolahan bahasa alami, teks pertambangan, dan dukungan pelanggan, untuk mendorong peningkatan dan penanganan pelanggan yang lebih baik; maka menghasilkan pelanggan kepuasan. Pengenalan ucapan adalah salah satu masalah yang sulit, karena dibutuhkan teknik yang terintegrasi. Pengenalan suara, sering terjadi masalah terjadi karena kurangnya kosa kata yang cukup banyak. Pada saat ini era, pendekatan pengenalan suara telah digunakan di berbagai wilayah yang berbeda, termasuk sistem telepon otomatis, ponsel, dll. Namun, pencapaian pengenalan suara yang bebas dari gangguan, khususnya untuk kemampuan pembicaraan secara terus-menerus, tetap merupakan masalah yang belum terpecahkan dan sulit.

6.Hasil dan pembahasan

Salah satu manfaat utama sistem pengenalan ucapan (Speech recognition) adalah memungkinkan pengguna melakukan pekerjaan lain secara bersamaan. Pengguna dapat berkonsentrasi pada operasi observasi dan manual, dan masih mengendalikan mesin dengan perintah input suara. Dengan mempelajari dan menggabungkan 2 metode pengenalan suara yang berbeda, dapat disajikan sebuah studi retrospektif tentang

sistem pengenalan suara dan kecerdasan buatan dan membantu dalam memahami semua model statistik pada speech recognition (pengenalan suara). Juga telah merumuskan metode decoding yang berbeda, yang digunakan untuk tugas decoding yang realistis dan bahasa buatan terbatas.

Keuntungan dan kerugian

a. Artificial intelligence

Metode ini merupakan salah satu metode terbaik pada pengenalan suara dalam hal menguraikan kata, karena kecerdasan buatan dapat melakukan tugasnya dengan baik dan realistis. kecerdasan buatan menggabungkan pendekatan algoritmik tertentu, yang mendorong konversi koheren dan transformasi ucapan ke dalam pola yang mudah dibaca, dan sebaliknya. Kelemahan atau keterbatasannya adalah Speech Recognition adalah salah satu masalah yang sulit, karena dibutuhkan memiliki teknik yang sangat terintegrasi dan perhatian tinggi. Dalam pengenalan suara, sering terjadi masalah terjadi karena kurangnya kosa kata yang cukup banyak.

b. Natural Language Processing

Kemampuan mesin untuk membaca dan memahami pembicaraan yang dilakukan manusia secara alami dan mampu memperjemahkan kata kedalam teks ataupun membaca teks yang tersedia melalui berbagai media yang dapat dimengerti oleh mesin. Keterbatasan dari metode ini yaitu, sebagian besar sistem

pengenalan adalah pembicara independen, perlu untuk melatih sistem untuk mengenali dialek setiap pengguna. Selain itu faktor tata bahasa yang digunakan oleh pembicara dan diterima oleh sistem, tingkat kebisingan, tipe kebisingan, posisi mikrofon, dan kecepatan bicara juga mempengaruhi kualitas penerimaan suara.

Pengembangan Speech Recognition

Speech Recognition saat ini sangat populer untuk dikembangkan, saat ini banyak peneliti yang mengembangkan berbagai metode dan teknik untuk proses pengenalan suara, menguraikan kata, menterjemahkan kata maupun suara dan melakukan kegiatan percakapan secara terus menerus secara berulang hingga dapat membantu semua pekerjaan manusia. Dari kedua metode diatas dapat dibuat inovasi baru terkait dengan pengembangan atau proses speech recognition menjadi lebih efektif dan akurat yaitu dengan menggabungkan artificial intelligence dan Natural language processing.

Proses menggabungkan kedua metode yang berberda dapat menghasilkan sebuah inovasi baru yaitu menguraikan suatu kata dengan kosakata yang ada dipadukan dengan kemampuan mesin dalam mengenali dan memahami suara maupun kata, kemudian dapat dihasilkan suatu metode baru yang dapat mengoptimalkan Speech Recognition sehingga aplikasi atau dukungan terhadap pengenalan suara dapat berjalan luas dan merevolusi bagaimana cara orang untuk berkomunikasi dan mampu

menyediakan teknologi dimana speech recognition dapat dipergunakan secara global meliputi bisnis, ekonomi, kesehatan, pendidikan, dll.

C. Ketepatan Dalam Mempresentasikan Proyek Akhir (*Accuracy in Presenting Final Projects*)

1. Pengertian dan Tujuan Presentasi Ilmiah

Presentasi ilmiah adalah “suatu kegiatan yang lazim dilakukan dalam dunia ilmiah. Kegiatan tersebut berfungsi untuk menyebarkan informasi ilmiah. Karena mahasiswa merupakan intelektual yang berkewajiban menyebarkan ilmu yang dimilikinya, kemahiran untuk melakukan presentasi ilmiah merupakan suatu kebutuhan” (Liani, 2012).

Presentasi dapat didefinisikan sebagai kegiatan berbicara dihadapan publik untuk mengkomunikasikan suatu pokok *bahasan yang merupakan informasi mengenai suatu gagasan atau objek. Sedangkan presentasi ilmiah adalah presentasi yang disampaikan oleh seorang ilmuwan mengenai suatu gagasan atau objek ilmiah dihadapan khalayak ilmiah.*

Adapun tujuan dari presentasi secara umum menurut Jay dan Jay (2004) adalah sebagai berikut :

- a. Untuk memberikan informasi
- b. Untuk mempengaruhi orang lain
- c. Untuk memberikan kesan yang baik kepada orang lain

2. Tata Cara Presentasi Ilmiah yang Baik dan Benar

Presentasi ilmiah akan berhasil jika penyaji memenuhi tata cara yang lazim. Menurut Boeditama (2008) tata cara presentasi ilmiah yang lazim adalah sebagai berikut :

- a. Penyaji perlu memberi informasi kepada peserta secara memadai Informasi akan dipahami dengan baik jika peserta memperoleh bahan tertulis, baik bahan lengkap maupun bahasan presentasi powerpoint.
- b. Penyaji menyajikan bahan dalam waktu yang tersedia Penyaji perlu merencanakan penggunaan waktu dan menaati panduan yang diberikan oleh moderator.
- c. Penyaji menaati etika yang berlaku di forum ilmiah

Dalam forum ilmiah ada beberapa peran yang dimainkan oleh aktor yang berbeda, yakni penyaji, pemandu (moderator), notulis, peserta, dan teknisi. Semua pihak wajib melakukan tugasnya dan menjaga agar jalannya presentasi ilmiah dapat berjalan dengan lancar sesuai dengan aturan main yang telah ditetapkan. Forum ilmiah adalah wahana bagi ilmuwan dan akademi dari berbagai disiplin ilmu untuk saling bertukar pendapat atau informasi sebagai hasil penelitian.

3. Langkah-langkah dalam Presentasi Ilmiah

Langkah-langkah dalam melakukan presentasi ilmiah antara lain :

a. Tahap persiapan

Dalam tahap persiapan ada beberapa hal yang harus dilakukan, yaitu :

1) Menentukan Topik

Hal-hal yang dapat dianjurkan dalam memilih topik menurut Kasali (2001:68) adalah sebagai berikut :

- a) Tolak topik-topik yang tidak termasuk dalam bidang keahlian
- b) Bila topik berhubungan dengan keahlian ide namun bersinggungan dengan bidang lain yang bersifat agak teknis, lakukan tandem (terjun bersama-sama) dengan orang lain yang dianggap ahli
- c) Topik yang baik haruslah yang relevan dengan kebutuhan audiens, aktual, berisi wawasan teoritis (mampu menjelaskan suatu fenomena), praktis sekaligus mudah dipahami, terfokus, dan dapat disajikan dengan menarik
- d) Adakalanya kita diminta untuk memberikan wawasan praktik saja, tanpa perlu memberikan warna teoretis sama sekali

2) Menentukan tujuan

Menyusun tujuan yang tepat untuk menjadi pedoman saat kita memilih segala sesuatu yang pantas dimasukkan atau dikeluarkan dari presentasi (Jay dan Jay, 2004:3).

3) Mengenal Situasi dan Audiens

Mengenal situasi atau suasana yang dimaksud adalah suasana lingkungan apakah itu di perkotaan, pedesaan atau di tengah kota yang bising. Suasana lingkungan ini harus kita pikirkan ketika kita mempersiapkan presentasi. Logikanya suasana lingkungan yang dipilih untuk suatu presentasi akan mempengaruhi presentasi dan emosi dari peserta presentasi (Darmastuti, 2006:102-103).

Faktor lain yang harus diperhatikan yaitu adalah dengan mengenal siapa audiens kita, agar kita dapat menyesuaikan materi seperti apa yang tepat buat mereka. Menurut Darmastuti (2006:103) mengenal audiens dibedakan berdasarkan latar belakang mereka yang terdiri dari :

a) Tingkat Pendidikan

Data tingkat pendidikan sangat diperlukan untuk memperhitungkan kemampuan daya serap peserta presentasi.

b) Profesi

Data profesi diperlukan agar penyaji memahami dengan siapa kita berhadapan.

c) Umur dan Pengalaman Hidup

Data umur dari calon audiens diperlukan agar penyaji mengerti dengan pasti kecenderungan seseorang pada umur tertentu.

4) Menyusun Materi

Menurut Darmastuti (2006:106) dalam penyusunan materi perlu diperhatikan faktor-faktor :

a) Aktualitas

Kemampuan mengemukakan pendapat pada hal-hal yang baru dan tidak ketinggalan jaman.

b) Konsepsional

Memberi keluasan wawasan, jangkauan pandangan, visi ke depan dan konsep-konsep yang bersumber pada gagasan segar.

c) Faktual

Menyajikan hal-hal yang sesungguhnya, kebenaran, data, bukan merupakan fiksi semata.

d) Kontekstual

Hendaknya memperhatikan relevansi antara tema yang dipilih dan materi yang disajikan.

5) Menentukan Pendekatan yang Digunakan

Setiap presentasi membutuhkan metode yaitu sebuah pendekatan yang digunakan dalam menyampaikan materi. Metode yang baik adalah metode yang sesuai dengan topik yang disampaikan.

6) Menyusun Slide Presentasi

Jika presentasi menggunakan slide, maka slide disusun semenarik mungkin. Ada beberapa prinsip yang harus dipahami dalam membuat slide, yaitu :

a) Sederhana

- b) Konten yang kuat
- c) Font yang indah
- d) Gambar yang menarik dan sesuai
- e) Penggunaan warna yang tepat

7) Latihan presentasi

Latihan presentasi memiliki dua tujuan menurut Jay dan Jay (2004:164), yaitu:

- a) Untuk melatih presentasi yang dipersiapkan dengan baik
- b) Untuk memberikan pengalaman dan dukungan kepada mereka yang terlibat dalam presentasi

b. Tahap Membawakan

Ada tiga struktur dalam membawakan presentasi, yaitu :

1) Membuka Presentasi

Presentasi yang baik harus dimulai dengan pembukaan yang baik. Pembukaan yang baik sangat menentukan keberhasilan sebuah presentasi.

Terdapat lima poin untuk membuka presentasi yang lengkap menurut Jay dan Jay (2004:10-11) :

a) Ucapan selamat datang

Mengucapkan terima kasih kepada audiens yang telah meluangkan waktu dan berharap waktu mereka tidak sia-sia.

b) Identifikasi diri

Memperkenalkan nama dan pekerjaan kita, latar belakang anda secara relevan.

- c) Tujuan
Menjelaskan rencana apa saja yang akan disampaikan di dalam presentasi.
- d) Pemetaan
Menjelaskan berapa lama presentasi yang akan berlangsung.
- e) Aturan main
- f) Menjelaskan apakah anda mengizinkan orang-orang melakukan interupsi saat mereka mempunyai pertanyaan, atau menunggu sampai akhir sesi, atau menyimpan semua pertanyaan sampai presentasi berakhir.

2) Pembahasan

Dalam menyampaikan materi harus secara teratur, dan materi disampaikan secara keseluruhan. Penyaji perlu memahami beberapa hal yang perlu diperhatikan saat menyampaikan materi menurut Pujiono (2013:87-91), diantaranya sebagai berikut :

- a) Ketepatan ucapan
Dalam hubungannya dengan olah suara atau tata bunyi ini, hal-hal yang harus diperhatikan menurut Pujiono (2013:87)., berikut :
 - (1) Logat baku tidak tercampur dengan dialek tak baku
 - (2) Lafal harus jelas dan tegas

- (3) Nafas yang kuat agar dapat menguraikan kalimat yang cukup panjang atau tidak terputus dalam wicara
 - (4) Tempo (cepat lambat suara) dan dinamik (intonasi, tekanan, aksen) suara
 - (5) Penghayatan, berbicara memerlukan penjiwaan agar sesuai dengan tuntutan situasi dan kondisi
- b) Penempatan tekanan, nada, sandi, dan durasi
- Tekanan berhubungan dengan keras lemahnya suara, nada berhubungan dengan tinggi-rendahnya suara, sandi atau tempo berhubungan dengan cepat-lambatnya berbicara, dan durasi atau jeda menyangkut perhentian (Pujiono, 2013:88).
- c) Pilihan kata (diksi)
- Untuk memperoleh ketepatan dalam penggunaan kata-kata, pembicara perlu memperhatikan prinsip-prinsip menurut Pujiono (2013:88-89) berikut :
- (1) Hindari kata-kata klise
 - (2) Hati-hati dalam penggunaan kata-kata
 - (3) Gunakan bahasa pasaran secara hati-hati
 - (4) Jangan menggunakan penjulukan
 - (5) Hindari fulgarisme dan kata-kata yang tidak sopan
 - (6) Jangan menggunakan *eufemisme* yang berlebihan
- d) Menggunakan kalimat efektif untuk berbicara
- Kalimat yang benar dan jelas yang dapat dengan mudah dipahami pendengar sesuai dengan maksud

pembicara disebut kalimat efektif. Kalimat efektif memiliki ciri-ciri keutuhan, perpautan, pemusatan perhatian, dan kehematan (Pujiono, 2013:89).

e) Sikap yang wajar dan tenang

Kesan pertama dalam presentasi sangat menentukan keberhasilan dalam proses presentasi berikutnya. Untuk itu, dalam presentasi seorang pembicara harus dapat bersikap yang wajar, tenang, dan tidak kaku agar dapat menambah kepercayaan pendengar kepada pembicara (Pujiono, 2013:89).

f) Kontak mata dengan audiens

Melihat audiens secara sekilas sangat penting saat presentasi. Pandangan kita terhadap audiens harus merata ke seluruh ruangan. Berikan pandangan positif dan penuh semangat agar audiens konsentrasi terhadap apa yang kita presentasikan (Pujiono, 2013:90).

g) Gerak dan mimik

h) Kenyaringan suara

i) Kelancaran

j) Penalaran

3) Menutup Presentasi

Tutup presentasi dengan menyimpulkan inti dari presentasi yang disampaikan. Jangan tinggalkan audiens dengan kebingungan. Tinggalkan audiens dengan sesuatu yang bermakna.

c. Tahap Evaluasi

Tahap evaluasi sebaiknya harus dilakukan agar menjadi satu kebiasaan yang positif. Tahap evaluasi bertujuan untuk mengevaluasi bagian mana saja yang tidak sesuai dengan rancana, mana saja yang perlu mendapat perbaikan dan mana saja yang perlu dipertahankan agar memberikan dampak positif untuk presentasi berikutnya.

4. Metode Presentasi

a. Metode impromptu

Presentasi impromptu adalah presentasi tanpa persiapan, latihan, riset, dan dilakukan secara mendadak (Kasali, 2001:63).

b. Metode Menghafal

Presentasi dengan metode menghafal adalah presentasi dengan menggunakan naskah yang telah dipersiapkan kemudian dihafal kata demi kata. Dengan metode menghafal penyaji cenderung berbicara cepat dan tidak menghayati maknanya. Selain itu, pembicara juga sulit menyesuaikan diri dengan situasi dan reaksi-reaksi pendengar selagi presentasi (Santoso, 2013).

c. Metode Naskah

Metode naskah sering dipakai dalam pidato resmi atau pidato radio. Metode naskah masih bersifat agak kaku kecuali banyak latihan, karena mata pembicara selalu ditunjukkan pada naskah dan intonasi monoton (Santoso, 2013).

d. Metode Ekstemporan

Metode ekstemporan adalah metode presentasi dengan menggunakan naskah yang dipersiapkan tetapi pembicara tidak menghafal isinya, hanya menggunakan catatan untuk mengingat urutan-urutan idenya (Santoso, 2013).

5. Presentasi yang Efektif dan Efisien

Ada beberapa hal agar presentasi dapat berlangsung secara efektif dan efisien menurut Santoso (2013), diantaranya yaitu :

a. Kunci dari presentasi

1) Komunikasi

Penyaji harus bisa semenarik mungkin dan harus mampu membuat audiens mengerti tentang apa yang dipresentasikannya.

2) Kemampuan berbicara di depan umum

Kemampuan berbicara di depan umum harus dilatih agar menumbuhkan rasa percaya diri dan tidak gugup.

3) Artikulasi

Artikulasi atau pengucapan harus jelas.

4) Bahasa

Bahasa yang digunakan sebaiknya menggunakan Bahasa Indonesia yang baik dan benar.

5) Media

Media merupakan salah satu kunci presentasi, karena dapat membantu saat presentasi berlangsung. Dengan adanya media presentasi akan menjadi semakin hidup.

b. Presentasi yang baik dan buruk

1) Presentasi yang baik menurut Santoso (2013) adalah:

- a) Energik dan penuh semangat
- b) Kontak mata dengan audiens
- c) Berbicara jelas dan cukup keras
- d) Seseekali bergerak saat berbicara
- e) Menggunakan anekdot dan humor yang sesuai
- f) Argumen-argumen terstruktur dengan baik
- g) Slide dapat dibaca
- h) Tipe slide bervariasi
- i) Tidak lebih dari 1 slide per menit
- j) Variasi teknologi lain, misalnya video
- k) Selesai tepat waktu dan sediakan waktu untuk tanya jawab

2) Presentasi yang buruk menurut Santoso (2013)

adalah sebagai berikut :

- a) Tujuan tidak jelas
- b) Pandangan tidak fokus tidak ada kontak dengan audiens
- c) Pengulangan yang tidak perlu
- d) Kurang persiapan
- e) Terlalu rumit atau sederhana bagi audiens

- f) Terlalu banyak slide
- g) Slide tidak dapat dibaca
- h) Penggunaan efek-efek teknis PowerPoint yang berlebihan
- i) Penggunaan warna yang buruk pada slide
- j) Penggunaan peralatan teknis yang salah
- k) Melebihi waktu yang dialokasikan untuk presentasi

6. Tips Presentasi yang Baik

Berikut ini adalah tips presentasi yang baik baik menurut Kasali (2001:1-4), adalah sebagai berikut :

- a. Jangan membiasakan diri tergantung pada teks. Teks dapat membunuh bakat, merusak flow, dan menciptakan jarak.
- b. Pelajari dulu siapa audiens kita, latar belakang, jalan pikiran, pendidikan, dan jabatan mereka.
- c. Jangan bicarakan hal yang sudah mereka ketahui atau yang tidak ingin mereka dengar. Selalu sajikan hal-hal yang orisinal, jangan merusak *mood* audiens dengan pernyataan yang tidak mereka sukai.
- d. Humor tidak boleh berlebihan.
- e. Perikasa ruangan dan fasilitas presentasi termasuk mikrofon sebelum presentasi dimulai.
- f. Biasakan interaktif. Jangan asyik bicara sendiri.
- g. Selalu berikan contoh dan ilustrasi.
- h. Jangan merendahkan mutu.
- i. Latihan yang cukup. Selalu mintalah umpan balik.

- j. Perhatikan bahasa yang tubuh. Jangan melakukan gerakan yang merusak penampilan.
- k. Berpakaianlah agak cerah agar anda menciptakan kesegaran di dalam ruangan.
- l. Jangan berbicara seperti sedang ngobrol dengan seseorang.

7. Etika Presentasi Ilmiah

Etika dalam presentasi ilmiah berkaitan dengan keyakinan dan prinsip mengenai mana yang benar mana yang salah serta mana yang patut dan mana yang tidak patut. Salah satu nilai yang harus dipegang dalam menjaga etika adalah menjaga perilaku agar tidak merugikan orang lain. Etika dalam forum ilmiah harus dijaga agar tujuan forum ilmiah tercapai dengan baik.

Dalam presentasi ilmiah terdapat lima peran penting agar kegiatan tersebut dapat berlangsung dengan baik. Etika dalam presentasi ilmiah berdasarkan perannya masing-masing menurut Sundusiah (2012) adalah :

a. Etika Penyaji

- 1) Perlu memberikan informasi secara memadai kepada peserta
- 2) Menyajikan bahan dalam waktu yang tersedia
- 3) Menjaga perilaku agar tidak merugikan orang lain
- 4) Kerugian berupa hak dan kesempatan berbicara, tersinggung, atau kehilangan harga diri
- 5) Mengutamakan kejujuran

- b. Etika Peserta
 - 1) Menjadi penyimak yang baik
 - 2) Mengungkapkan apresiasi positif terhadap penyaji
 - 3) Tidak meninggalkan forum sebelum jawaban disampaikan
 - 4) Jika akan meninggalkan forum, peserta meminta izin kepada moderator atau penyelenggara
- c. Etika Moderator
 - 1) Adil
 - 2) Taat jadwal
- d. Etika Notulis
 - 1) Mencatat dengan rapi dan teliti semua hal yang terungkap dalam forum
 - 2) Meringkas hasil catatan dan disebarikan kepada forum
- e. Etika Teknisi
 - 1) Memastikan bahwa peralatan teknologi bekerja dengan baik
 - 2) Mengontrol peralatan selama presentasi

Pidato Ilmiah

1. Pengertian dan Tujuan Pidato Ilmiah

Pidato adalah “penyamapaian gagasan, pikiran, informasi, dan tujuan dari pembicara kepada orang lain secara lisan. Berpidato merupakan salah satu kegiatan menyampaikan gagasan secara lisan dengan menggunakan penalaran yang

tepat serta memanfaatkan aspek-aspek non-kebahasaan (seperti ekspresi wajah, kontak pandang, gerak tangan) yang dapat mendukung efisiensi dan efektivitas pengungkapan gagasan kepada orang banyak dalam suatu acara tertentu” (Amedia, 2013).

Tujuan penyajian pidato ada tiga menurut Amedia (2013), yaitu :

- a. Menyampaikan informasi
- b. Meyakinkan dan mempengaruhi sikap pendengar
- c. Menghibur pendengar

2. Kriteria Pidato yang Baik

Pidato yang baik menurut Venus (2013) ditandai dengan kriteria sebagai berikut :

- a. Isinya sesuai dengan kegiatan yang berlangsung
- b. Isinya menggugah dan memiliki manfaat bagi pendengar
- c. Isinya tidak menimbulkan pertentangan SARA
- d. Isinya jelas, benar, objektif
- e. Menggunakan bahasa yang mudah dipahami
- f. Menyampaikan secara santun, rendah hati, dan bersahabat

3. Persiapan Sebelum Pidato

Menurut Keraf (1980:317-318) ada tujuh langkah yang perlu diperhatikan dalam mempersiapkan pidato yang baik, yaitu :

- a. Menentukan topik dan tujuan
- b. Menganalisis pendengar dan situasi

- c. Memilih dan menyempitkan topik
 - d. Mengumpulkan bahan
 - e. Membuat kerangka uraian
 - f. Menguraikan secara mendetail
 - g. Melatih dengan suara nyaring
4. Tata Cara Pidato yang Baik dan Benar
- “Tata cara berpidato merujuk pada langkah-langkah dan urutan untuk memulai, mengembangkan, dan mengakhiri pidato. Langkah-langkah dan urutan berpidato secara umum diawali dengan pembukaan, isi, dan penutup”.
- a. Pembukaan
Biasanya berisi sapaan kepada pihak-pihak yang diundang atau hadir dalam suatu acara, ucapan rasa syukur kepada Tuhan Yang Maha Esa.
 - b. Isi
Biasanya berisi hasil penjabaran gagasan pokok.
 - c. Penutup
Biasanya berisi penegasan kembali gagasan pokok yang telah dipaparkan dan disajikan isi, berisi harapan dan ucapan terima kasih atas partisipasi semua pihak dalam acara yang sedang berlangsung (Amedia, 2013).
5. Metode Berpidato
- Pada umumnya metode berpidato ada empat menurut Amedia (2013), yaitu :
- a. Metode Naskah

Metode naskah adalah metode berpidato yang digunakan dalam pidato resmi dan dibacakan secara langsung agar tidak terjadi kekeliruan.

b. Metode Menghafal

Metode menghafal adalah metode berpidato dengan naskah yang telah dipersiapkan sebelumnya bukan untuk dibaca, melainkan untuk dihafal.

c. Metode Impromptu

Metode impromptu adalah metode berpidato yang tidak dilakukan persiapan atau pembuatan naskah tertulis terlebih dahulu.

d. Metode Ekstemporan

Metode ekstemporan atau penjabaran kerangka adalah metode berpidato dengan menjabarkan materi pidato yang terpola secara lengkap. Maksud dari terpola yaitu materi yang akan disampaikan harus disiapkan garis-garis besar isinya dengan menuliskan hal-hal yang dianggap paling penting untuk disampaikan

6. Etika dalam Berpidato

Etika dalam berpidato akan menjadi pegangan bagi siapa saja yang akan berpidato. Ketika berpidato janganlah menyinggung perasaan orang lain. Sebaiknya berupaya menghargai dan membangun optimisme bagi pendengarnya. Selain itu keterbukaan, kejujuran, empati, dan persahabatan perlu diusahakan dalam berpidato.

BAB 12

PENGENALAN POLA ERA DIGITALISASI

(RECOGNITION OF DIGITALIZED ERA PATTERNS)

A. Online Pendidikan Jarak Jauh (*Online Distance Education*)

Pendidikan jarak jauh (bahasa Inggris: *distance education*) adalah pendidikan formal berbasis lembaga yang peserta didik dan instrukturanya berada di lokasi terpisah sehingga memerlukan sistem telekomunikasi interaktif untuk menghubungkan keduanya dan berbagai sumber daya yang diperlukan di dalamnya. Pembelajaran elektronik (*e-learning*) atau pembelajaran daring (*online*) merupakan bagian dari pendidikan jarak jauh yang secara khusus menggabungkan teknologi elektronika dan teknologi berbasis internet.

Kemajuan yang terjadi dalam dunia teknologi komunikasi dan informasi memunculkan peluang maupun tantangan baru dalam dunia pendidikan. Peluang baru yang muncul termasuk akses yang lebih luas terhadap konten multimedia yang lebih kaya, dan berkembangnya metode pembelajaran baru yang tidak lagi dibatasi oleh ruang dan waktu. Di sisi lain kemajuan teknologi dengan beragam inovasi digital yang terus berkembang juga menghadirkan tantangan baru bagi penyelenggara pendidikan untuk terus menyesuaikan infrastruktur pendidikan dengan teknologi baru tersebut.

Pendidikan jarak jauh bukan metode baru dalam sistem pendidikan. Metode pembelajaran ini telah digunakan di Amerika Serikat sejak tahun 1892 ketika Universitas Chicago meluncurkan program pembelajaran jarak jauh pertamanya untuk tingkat pendidikan tinggi. Metode pembelajaran jarak jauh terus berkembang dengan menggunakan beragam teknologi komunikasi dan informasi termasuk radio, televisi, satelit, dan internet.^[3] Meluasnya penggunaan internet oleh publik di berbagai negara pada tahun 1996 menjadi suatu fenomena yang berkembang dan diikuti oleh kemunculan beragam konten digital di dalamnya.^[4] Pada tahun yang sama, John Bourne mengembangkan *Asynchronous Learning Network Web* yang merujuk kepada kemampuan untuk memberikan pendidikan kapan saja dan di mana saja melalui internet.^[3]

Sistem pendidikan konvensional

Pendidikan merupakan suatu proses akademis yang bertujuan untuk meningkatkan nilai-nilai moral, sosial, budaya dan agama sekaligus mempersiapkan peserta didik untuk mampu menghadapi berbagai tantangan dalam proses kehidupan. Dalam pendidikan terjadi proses komunikasi yang terorganisasi dan berkelanjutan untuk menumbuhkan aktivitas belajar dalam diri pembelajar sehingga pembelajar dapat mengembangkan kemampuannya dalam menemukan, mengolah, dan mengevaluasi berbagai informasi dan pengetahuan untuk kemudian berkontribusi dalam pencarian solusi atas masalah

yang ada dan berpartisipasi aktif di masyarakat.^[5] Dalam sistem pendidikan konvensional, metode yang digunakan adalah melalui pertemuan tatap muka antara pengajar dan peserta didik. Dengan kata lain, pengajar dan pembelajar berada di ruang yang sama pada waktu yang bersamaan juga untuk saling berkomunikasi dan berinteraksi.

Karakteristik Pendidikan Konvensional

Pendidikan konvensional adalah pendidikan formal yang menggunakan sistem klasikal dalam menyampaikan materi ajar baik di sekolah, akademi, universitas, dan sejenisnya.^[6] Beberapa karakteristik dasar dari sistem pendidikan konvensional antara lain:

- Pengajar dan peserta didik berada dalam ruang yang sama pada waktu yang sama untuk melakukan kegiatan belajar-mengajar.
- Kegiatan belajar-mengajar dilakukan dalam bentuk pertemuan tatap muka.
- Pengajar menentukan tujuan belajar, materi ajar, dan evaluasi proses belajar dari peserta didiknya.
- Proses komunikasi antara pengajar dan peserta didik dilakukan secara langsung atau bersifat analog.
- Menitikberatkan pada peran pengajar sebagai sumber informasi dan dalam pengelolaan kelas selama proses belajar-mengajar berlangsung.

Teknologi Komunikasi Dalam Pendidikan Konvensional

Dalam sistem pendidikan konvensional, teknologi komunikasi yang digunakan untuk mendukung proses belajar-mengajar mencakup baik teknologi analog maupun teknologi digital dengan titik berat pada proses penyampaian informasi secara analog. Teknologi analog merujuk kepada segala bentuk teknologi yang dibuat untuk menyerupai bentuk asli dan yang dapat ditangkap oleh panca indra manusia, sedangkan teknologi digital merupakan teknologi berbasis komputerisasi yang basis datanya terdiri dari bilangan nol dan satu.

Proses komunikasi yang berlangsung dalam pertemuan tatap muka menekankan pada interaksi dan relasi sosial yang terbangun di antara pengajar dan peserta didik. Pengajar menyampaikan materi ajar kepada peserta didik secara langsung dengan berbagai ekspresi dan gerakan (*gesture*) yang mendukungnya, dan peserta didik dapat memberikan tanggapan serta mengekspresikan diri secara langsung juga. Teknologi komunikasi penunjang yang biasanya digunakan dalam pendidikan konvensional mencakup antara lain kertas, buku, papan tulis, spidol, televisi, radio, alat peraga, komputer, proyektor, dan lain-lain.

Keunggulan dan kelemahan pendidikan konvensional

Keunggulan dari metode pendidikan konvensional antara lain:

- Tingginya tingkat interaksi langsung antara pengajar dan peserta didik yang akan mempercepat terbentuknya relasi dan nilai-nilai dalam proses belajar-mengajar.
- Pertemuan tatap muka antara pengajar dan peserta didik mendukung terselenggaranya proses belajar-mengajar yang terfokus dan terkontrol sehingga pembelajaran dapat dioptimalkan.

Kelemahan dari metode pendidikan konvensional antara lain:

- Kebergantungan kepada lokasi, tempat, dan kehadiran dari para peserta proses belajar-mengajar baik pengajar maupun peserta didik.
- Biaya pendidikan relatif tinggi sebagai akibat dari timbulnya berbagai biaya untuk menunjang operasional sehari-hari pada lembaga penyelenggara pendidikan konvensional.

Sistem pendidikan jarak jauh

Pendidikan jarak jauh adalah suatu kajian kependidikan yang terus berkembang seiring dengan pesatnya kemajuan teknologi komunikasi dan informasi. Karena itu juga pendidikan jarak jauh sering dipersepsikan sebagai suatu inovasi dalam metode pembelajaran abad 21 yang memiliki daya jangkauan lintas ruang, waktu, dan sosioekonomi. Dengan adanya inovasi ini, masyarakat memiliki pilihan alternatif untuk mengakses pendidikan. Secara umum, pendidikan jarak jauh memiliki prinsip yang mencakup antara lain:

- Akses, yakni terkait dengan keinginan untuk memperluas akses masyarakat terhadap pendidikan melalui penyelenggaraan pendidikan yang berbasis teknologi komunikasi dan informasi, bersifat massal, ekonomis, serta meminimalkan kendala jarak dan waktu.
- Pemerataan yang merujuk kepada asas keadilan dan persamaan hak bagi siapa saja untuk mengenyam pendidikan tanpa dibatasi oleh berbagai kendala.
- Kualitas, yaitu berkenaan dengan jaminan standar pengajar, materi bahan ajar dan ujian, dan proses pembelajaran interaktif yang berbasis teknologi komunikasi dan informasi.

Karakteristik pendidikan jarak jauh

Pendidikan jarak jauh memiliki beberapa karakteristik dasar, yaitu:

- Pengajar dan peserta didik tidak berada dalam satu ruang yang sama saat proses belajar-mengajar berlangsung.
- Penyampaian materi ajar dan proses pembelajaran dilakukan dengan memanfaatkan media komunikasi dan informasi.
- Menekankan pada cara belajar mandiri namun ada lembaga yang mengaturnya.
- Keterbatasan pada pertemuan tatap muka. Biasanya pertemuan tatap muka dilakukan secara periodik antara peserta didik dengan pengajar atau tutor.
- Fleksibilitas dalam proses pembelajaran. Dengan kata lain masing-masing peserta didik dapat mengatur waktu belajarnya sendiri sesuai dengan ketersediaan waktu dan kesiapannya.

Teknologi komunikasi dalam pendidikan jarak jauh

Penyelenggaraan pendidikan jarak jauh tidak dapat dilepaskan dari penggunaan teknologi. Hal ini dikarenakan dalam pendidikan jarak jauh tidak terjadi kontak secara langsung antara pengajar dan peserta didik. Proses komunikasi antara keduanya dilakukan melalui pemanfaatan teknologi komunikasi dan informasi. Walau demikian, pertemuan tatap muka tetap dapat dilakukan dengan frekuensi yang terbatas. Teknologi komunikasi dan informasi yang banyak digunakan dalam pendidikan jarak jauh adalah komputer dan internet.

Pemanfaatan komputer dan internet memberikan kesempatan kepada peserta didik untuk dapat mengakses materi-materi ajar yang juga sudah dikemas dalam bentuk digital di mana pun dan kapan pun. Dengan menggunakan komputer dan internet juga, pengajar dan peserta didik dapat melakukan interaksi baik menggunakan aplikasi surat elektronik, video konferensi, atau forum diskusi dalam jaringan.^[12] Meski penggunaan berbagai teknologi digital dalam pendidikan jarak jauh membuat batas-batas geografis seakan lenyap, namun proses komunikasi yang dimediasi oleh komputer dan internet memiliki keterbatasan dalam menangkap ekspresi dan gerakan (*gesture*) dari pengajar dan peserta didik. Teknologi komunikasi pendukung lainnya yang digunakan untuk menunjang penyelenggaraan pendidikan jarak jauh antara lain buku elektronik, *compact disc* (CD) atau *digital versatile disc* (DVD) untuk rekaman audio dan video, perangkat pengolah informasi seperti tablet atau laptop.

Keunggulan dan kelemahan pendidikan jarak jauh

Keunggulan dari metode pendidikan jarak jauh antara lain:

- Proses pembelajaran dapat dilakukan tanpa dibatasi oleh keharusan pengajar dan peserta didik untuk berada di ruang dan waktu yang sama.
- Penggunaan teknologi komunikasi dan informasi sebagai media pembelajaran menimbulkan biaya yang lebih rendah baik bagi penyelenggara pendidikan jarak jauh maupun peserta didik.
- Materi ajar dan berbagai interaksi dalam bentuk tulisan yang dikemas secara digital memungkinkan peserta didik untuk dapat membaca kembali berulang-ulang informasi yang tercatat di dalamnya.

Kelemahan dari metode pendidikan jarak jauh antara lain:

- Minimnya kontak langsung antara pengajar dan peserta didik memperlambat proses terbangunnya relasi sosial dan nilai-nilai yang menjadi tujuan dasar dari pendidikan.
- Rendahnya kontrol terhadap proses pembelajaran sebagai implikasi dari cara belajar mandiri yang menjadi titik berat dari pendidikan jarak jauh.
- Keterbatasan teknologi komunikasi dan informasi yang tidak dapat menggantikan sepenuhnya proses komunikasi dan interaksi secara langsung yang terjadi dalam pendidikan konvensional.

B. Robotisasi (*Robotization*)

Kemajuan teknologi digital saat ini bisa menjadi berkah sekaligus musibah bagi kehidupan manusia. Di satu sisi, kehadiran teknologi memudahkan pekerjaan manusia, namun di sisi lain menciptakan masalah.

Cukup banyak sebutan untuk era sekarang ini. Disebut "Abad ke-21" karena memang memasuki tahun 2100 (2018). "Era milenial" karena masuk di milenium ketiga. "Revolusi Industri Keempat" sebab ketiga industri sebelumnya telah dilewati (mesin uap, minyak dan batubara, komputer dan internet).

Dinamai "Internet of Thing" dikarenakan segala aspek kehidupan kita menggunakan internet. Ada lagi "Artificial Intelligence" atau kecerdasan buatan, disebabkan mesin-mesin baru (robot) memiliki kecerdasan tertentu, bisa melakukan pembelajaran sendiri, atau terkadang memiliki kecerdasan yang tak dimiliki oleh manusia.

Di era milenial ini semuanya mengalami digitalisasi atau otomatisasi, dalam arti banyak pekerjaan manusia yang diambil alih oleh mesin-mesin baru yang bekerja secara digital atau otomatis. Melakukan berbagai transaksi keuangan atau perbankan, jual-beli, pembayaran dan tagihan, mengirim aplikasi, dan berbagai aktivitas rutin lainnya. Bahkan ke depan, pengobatan modern, kendaraan mobil juga mengalami digitalisasi.

Dengan mengalami digitalisasi, akan ada banyak orang kehilangan pekerjaan. Tapi di sisi lain, akan banyak pula

pekerjaan-pekerjaan baru bermunculan. Tidak sedikit orang yang merasa khawatir (pesimis), tapi tidak sedikit juga orang yang merasa penuh harap (optimis). Namun pada intinya, bagi mereka yang tidak kreatif dan mampu beradaptasi, akan terkalahkan bahkan punah.

Dengan digitalisasi, semua pekerjaan manusia menjadi lebih mudah, cepat, dan simpel. Belanja sesuatu tidak harus pergi jauh (ke kota) yang harus menanggung biaya transportasi, jajan, waktu, dan tenaga. Transaksi keuangan tidak harus pergi ke bank, antre lama, mengisi banyak formulir, dan prosedur yang rumit.

Urusan pemerintahan pun tidak harus ke ibukota provinsi atau kabupaten, antrean panjang, dan berkas setumpuk. Membeli tiket tidak harus ke bandara, stasiun, terminal dll. Dan berbagai kemudahan hidup lainnya.

Dengan digitalisasi pula, semua masuk ke dalam lingkaran jejaring, tidak hanya lokal dan nasional, tapi global. Terkoneksi, terhubung, terintegrasi satu sama lain secara menyeluruh. Kita telah menjadi penduduk global, sehingga batasan waktu dan letak geografis tak lagi menjadi penghalang. Segalanya telah realtime, 24 jam 7 hari, saat ini dan sekarang.

Mesin-mesin baru alis robot, makhluk yang tak bernyawa, bisa melakukan banyak hal seperti yang dilakukan oleh makhluk bernyawa (manusia). Ini terkait dengan kecerdasan buatan, algoritma, dan big data. Mesin tahu apa yang mesti dilakukan sesuai dengan keinginan manusia (input).

Dalam kasus belanja online misalnya, mesin tahu produk apa yang kita suka, yang sering kita lihat, kemampuan keuangan kita, sebulan belanja berapa dan apa saja. Dia tahu preferensi kita, kebiasaan kita, gaya dan perilaku kita.

Kalau dulu yang disebut modal perusahaan itu seperti tanah, gedung, kendaraan, maupun uang. Tapi sekarang, DATA adalah modal terpenting, terutama bagi perusahaan digital startup. Banyak perusahaan online yang tak memiliki kantor atau aset tertentu, tapi bisa menjadi besar dan sukses. Perusahaan taksi tak harus memiliki armada taksi, perusahaan penginapan tak harus memiliki hotel.

Banyak hal yang ingin dicapai oleh manusia di era kecerdasan buatan ini, seperti peningkatan taraf hidup, kelimpahan, kemakmuran, dan kualitas kesehatan yang lebih baik. Bahkan, ada sebagian yang memimpikan adanya imortalitas (keabadian, tidak mengalami mati) pada manusia. Sekiranya jantung bisa dibuat tetap berdetak, jika sel bisa diregenerasi, dan seandainya darah bisa terus mengalir; maka manusia akan hidup terus.

C. Kolaborasi (*Kolaborasi*)

Di era digital, perusahaan membutuhkan sistem komunikasi yang singkat, dimana, kapan saja, dan dengan perangkat komunikasi apa saja. Dengan komunikasi model ini perusahaan dapat memperoleh efisiensi bisnis dan produktivitas karyawan.

Kini Cisco memiliki Collaboration Solution yang telah dipakai oleh operator telekomunikasi dalam menjual layanan. Sementara Microsoft memiliki solusi yang disebut dengan Microsoft Teams. Kedua solusi dapat menyatukan orang, melalui percakapan, dan konten dengan perangkat yang dibutuhkan oleh tim.

PT Lintasarta menawarkan solusi kolaboratif dari Cisco. Anak usaha Indosat Ooredoo ini memiliki platform Lintasarta Collaboration Solution (LCS), yang dapat memberikan layanan collaboration dengan sistem sewa.

Sistem dari Cisco ini menawarkan *internet protocol phone*, *video telepresence*, *instance messaging* dan *voice gateway*. Layanan ini dapat digabungkan dengan layanan connectivity Lintasarta seperti MPLS, Metro-E dan Internet.

IT Services Director Lintasarta, Arya N Soemali menuturkan, keuntungan yang akan didapatkan oleh pelaku bisnis diantaranya selalu mendapatkan feature terdepan untuk collaboration technology, serta dapat mengintegrasikan beragam perangkat komunikasi dan video telepresence dengan perangkat dari berbagai macam merek.

“Melalui skema opex model pelaku bisnis juga tidak perlu investasi besar untuk menerapkan solusi collaboration dan sumber daya tenaga IT sehingga dapat fokus di bisnis utamanya.”

LCS merupakan platform komunikasi untuk memberikan layanan aplikasi bisnis dan aplikasi komunikasi terpadu. Proses ini juga menggunakan Cloud Infrastructure dari

Lintasarta untuk dapat meningkatkan customer experience dalam hal collaboration communication.

Bag perusahaan pelanggan, LCS dapat mempercepat proses bisnis para pelaku usaha dari berbagai sektor industri dan membantu mencapai target bisnis perusahaan dengan lebih cepat. Solusi ini diklaim dapat memenuhi standar kualitas layanan tinggi berkat dukungan infrastruktur TI kelas bisnis yang tersertifikasi.

Sistem LCS tidak membebani perusahaan karena menerapkan sistem sewa. Dengan pola operation expenditure (opex) model, pelanggan hanya membayar untuk fitur yang dibutuhkan dan digunakan sehingga komunikasi kolaboratif terjangkau dengan biaya lebih dapat diprediksi.

Selain komunikasi kolaboratif, bekerja secara bersama secara kolaboratif, di mana saja dan kapan saja saat ini sedang menjadi tren. Karyawan dapat menghasilkan pekerjaan yang membutuhkan banyak tenaga ahli tanpa harus bertemu dalam sebuah ruangan.

Microsoft misalnya baru saja meluncurkan Microsoft Teams. Solusi ini dapat menyatukan orang, percakapan, dan konten dengan perangkat yang dibutuhkan oleh tim.

Microsoft Teams terintegrasi dengan aplikasi Office yang berbasis pada Office 365 dan jaringan cloud global Microsoft yang aman. Microsoft Teams memiliki empat pilar yaitu Chat for Today's Team, A hub for teamwork, Customization for Every Team, dan Security teams trust. A hub for teamwork misalnya

adalah aplikasi dan layanan Office yang digunakan oleh tim setiap hari - Word, Excel, PowerPoint, SharePoint, OneNote dan PowerBI.

Customization for Every Team menawarkan kemampuan untuk menyesuaikan ruang kerja dengan tab, konektor dan bot dari mitra pihak ketiga serta layanan Microsoft lainnya seperti Microsoft Planner dan Visual Studio Team Services. Hari ini, lebih dari 150 integrasi telah tersedia atau akan segera hadir melalui perusahaan seperti SAP, Trello, Hipmunk, Growbot dan ModuleQ.

Sementara Security teams trust merupakan fitur keamanan yang dibuat dengan standar sangat tinggi, enterprise-grade cloud Office 365, memberikan kemampuan keamanan dan penyesuaian keamanan yang canggih seperti harapan pengguna. Teams mendukung standar global termasuk SOC 1, SOC 2, EU Model Clauses, ISO27001 dan HIPAA.

“Manfaat yang kita lihat dari Microsoft Teams yang terhubung dengan Office 365 cukup besar. Kami mampu bekerja sebagai satu tim, secara real time, membantu kami membuat keputusan lebih cepat, mendorong keterlibatan dan koneksi dalam bisnis.

DAFTAR PUSTAKA

(BIBLIOGRAPHY)

Aida-Zade, K.R., C. Ardil dan A.M. Sharifova. 2014. The Main Principles of Text-to-Speech Synresearch System. International Journal of Signal Processing 6.1

Apolloni, Bruno, Harpri (2015). Neural Nets. Springer-Verlag. hlm. 3. ISBN 3540202277.

Application in Speech Recognition. vol. 77. no. 2. pp. 257-286. IEEE.

Arman, A.A. tahun 2014. Teknologi Kunci untuk Meningkatkan Cara Interaksi antara Manusia dengan Mesin.

Away, Gunaidi Abdia. 2016. MATLAB programming, Bandung:Informatika

Aarne Ranta (Eds.), Advances in Natural Language Processing. Jurafsky D. dan Martin, J.H. (2018).

Baecker, Ronald M, Jonathan Grudin, William A. X. Buxton, Saul Greenberg (2015). Human-Computer Interaction : Toward the Year 2000 (Second Edition). Morgan Kauffman Publishers, Inc. hlm. 546. ISBN 1558602461.

Becchetti, Claudio dan Ricotti, Lucio Prina. 2017. Speech Recognition Theory and C++ Implementation. Chichester : John Willey & Sons.

Benesty, Jacob, M. Mohan Sondhi, dan Yiteng Huang (2018). Handbook of Speech Processing. Springer-Verlag. hlm. 6. ISBN 159904840X.

Bradbury, Jeremy, Linear Predictive Coding, paper, 2016

- Chou, Wu dan Juang, B. H., 2015. *Pattern Recognition in Speech and Language Processing*. Washington DC : CRC PRESS.
- D. T. Larose, *An Introduction to Data Mining*, Hoboken: John Wiley & Sons Inc., 2015.
- Fairus N.H, 2017, *Mahir Menggunakan Teknologi Informasi dan Komunikasi Microsoft Word 2003*, Ganeca Exact, Bandung.
- Google Now, Wikipedia Bahasa Indonesia, ensiklopedia bebas.
Available : (https://id.wikipedia.org/wiki/Google_Now)
- <http://alialbuni-aa.blogspot.co.id/2016/10/microsoft-xp.html>
<http://ghalbi201520169f.blogspot.co.id/2016/02/perbedaan-microsoft-office-word-excel.html>
- <http://jennysarumaha.blogspot.co.id/2016/04/makalah-tentang-microsoft-office.html>
- Huang, X., Acero, A., Hon, H.-W. 2015. *Spoken Language processing*. New Jersey: Prentice-Hall
- Ingason, Helgadóttir, Loftsson, Rögnvaldsson. (2018). A Mixed Method Lemmatization Algorithm Using a Hierarchy of Linguistic Identities (HOLI).
- Jelinek, Frederick (2017). *Statistical Methods For Speech Recognition*. Massachusetts Institute of Technology. hlm. 17. ISBN 0262100665.
- John. G, and Dimitris G. Manolakis, *Pemrosesan Sinyal Digital : Prinsip- prinsip, Algoritmadan Aplikasi*, PT Prenhallindo, Jakarta, 2015.
- Jurafsky, D., Martin, J.H. 2018. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition, 2nd Edition*. New Jersey: Prentice-Hall

- Kapetanios, E., Tatar, D., and Sacarea, C. 2015. *Natural Language Processing: Semantic Aspects*. Boca Raton: CRC Press
- Kowalski, M. (2014). *Information Retrieval Architecture and Algorithms*. New York : Springer.
- Kusrini dan Luthfi, E. T. 2017. *Algoritma Data Mining*. Edisi 1. Andi Offset.
- Kusumadewi, Sri. 2015. *Artificial Intelligence (Teknik dan Aplikasinya)*. Graha Ilmu. Yogyakarta.
- Levison, M., Lessard, G., Thomas, C., and Donald, M. 2014. *The Semantic Representation of Natural Language*. London: Bloomsbury
- M. J. Islam, Q. M. J. Wu, M. Ahmadi and M. A. SidAhmed, "Investigating the Performance of Naive- Bayes Classifiers and K- Nearest Neighbor Classifiers," International Conference on Convergence Information Technology, pp. 1541-1546, 2017.
- Manunggal, H. S. (2015). *Perancangan dan Pembuatan Perangkat Lunak Pengenalan Suara Pembicara Dengan Menggunakan Analisa MFCC Extraction*. Universitas Kristen Petra, Fakultas Teknologi IndustriJurusan Teknik Informatika. Surabaya, Indonesia: Tugas Akhir Strata 1.
- May, Ignatius Leo, *Pengenalan Vokal Bahasa Indonesia dengan Jaringan Saraf Tiruan melalui Transformasi Wavelet Diskret*, Jurusan Teknik Elektro Universitas Diponegoro, Tugas Akhir, 2017.
- Mudyahardjo, Redja. 2016. *Pengantar Pendidikan Komputer*. Jakarta: Rajawali Pers Tirtarahardja, U. Dan La Sula. 2014.

- Nirenburg, S. (2017). *Language Engineering for Lesser-Studied Languages*. Amsterdam : IOS Press.
- Pitowarno, Endro, *Robotika, Desain, Kontrol, dan Kecerdasan Buatan*, Penerbit Andi, 2016.
- Pustejovsky, J., Stubbs A. (2017). *Natural Language Annotation for Machine Learning*. Beijing : O'Reilly.
- Rabiner, Lawrence and Biing-Hwang Juang, *Fundamentals of Speech Recognition*, PTR Prentice-Hall Inc., New Jersey, 2014.
- Russel, S. J., Norvig, P. (2015). *Artificial Intelligence A Modern Approach*. New Jersey : Pearson Education Inc.
- Sadiman, 2016, *Teknologi Informasi dan Komunikasi untuk SMA kelas X*, Erlangga, Jakarta.
- Suhartono Derwin, (2015). *Natural Language Processing*.
- _____ Christiandy D., Rolando (2014). *Lemmatization Technique in Bahasa: Indonesian Language*. Kuwait: *Journal of Software* (unpublished)
- Supriyanto, 2015, *Pengantar Teknologi Informasi*, Salemba Infotek, Jakarta.
- Taylor, P. 2017. *Text-to-Speech Synthesis*. Cambridge: Cambridge University Press
- Tim Dosen. 2015. *Pemrosesan Bahasa Alami*. Bandar Lampung: STMIK Teknokrat. University, Yokohama.
- Waibel, Alex dan Kai-Fu Lee (2018). *Readings in Speech Recognition*. Morgan Kauffman Publishers, Inc. hlm. 1 & 267. ISBN 1558601244.

Zitouni, I. 2014. *Natural Language Processing of Semantic Languages*. Berlin: Springer