

IMPLEMENTASI DAN ANALISIS PENGOLAHAN KATA MENGGUNAKAN ALGORITMA HIDDEN MARKOV MODEL DENGAN POCKETSPHINX

IMPLEMENTATION AND ANALYSIS SPEECH RECOGNITION USING HIDDEN MARKOV MODEL ALGORITHM WITH POCKETSPHINX

Ayu Siti Aminah¹, Agung Nugroho Jati², Astri Novianty³

^{1,3} Prodi S1 Sistem Komputer, Fakultas Teknik, Telkom University
Bandung, Indonesia Ayusa42@gmail.com¹,
agungnj@telkomuniversity.ac.id²,
astrinov@telkomuniversity.ac.id³

Abstrak

Proses pengenalan suara dengan menggunakan algoritma HMM dan MFCC sebagai ekstraksi ciri akan dijelaskan pada paper ini. Pada tahun ini perkembangan aplikasi android sangat progresif. Semua aplikasi dapat diciptakan dengan mudah. Library pocketsphinx digunakan untuk mengembangkan pengenalan suara pada android. Karena library pocketsphinx telah menggunakan algoritma HMM untuk pengenalan suara. Secara default, library pocketsphinx digunakan untuk bahasa inggris. Pada paper ini, akan diciptakan pengenalan suara bahasa indonesia menggunakan library pocketsphinx. Dimana, algoritma HMM mempunyai tingkat akurasi sekitar 90%. Jadi, algoritma HMM cocok untuk pengenalan suara dengan minimal kesalahan.

Kata kunci: pengenalan suara, Hidden Markov Model (HMM), pocketsphinx

Abstract

Speech recognition that uses HMM algorithm and MFCC for extraction feature will explain in this paper. In this year, expansion of application on android are very progressive. All of application can create with easily. Pocketsphinx library is use to develop the speech recognition on android. Because, pocketsphinx library have used HMM algorithm to speech recognition. At default pocketsphinx library is used to english language. In this paper, will be create Indonesian language on pocketsphinx. In fact, HMM algorithm better than the other algorithm. HMM algorithm be obtained accuracy rate about 90%. So, hidden markov model is reasonable to speech recognition with the minimal error rate.

Keyword: speech recognition, Hidden Markov Model (HMM), pocketsphinx

1. Pendahuluan

Pengenalannya merupakan proses menginputkan suara manusia ke dalam suatu sistem. Dalam proses pengenalan suara ini membutuhkan berbagai macam perangkat seperti mikrofon, PC, dan alat pendukung lainnya. Mulai dari beberapa tahun yang lalu, sistem operasi berbasis Linux seperti Android menjadi sistem operasi favorit. Karena Android merupakan open source yang bisa dikembangkan oleh semua orang. Android mempunyai berbagai variasi aplikasi di smartphone. Berbagai macam aplikasi pengolahan teks, suara dan video.

Pocketsphinx berdasarkan HMM digunakan untuk mengembangkan aplikasi mobile. Dalam tulisan ini, akan menjelaskan HMM pada Android. Sehingga, dalam sidang ini akan membuat sebuah pengenalan suara pada smartphone terutama Android. Android memiliki catatan yang kompatibel di smartphone sehingga pembangunan Android harus easy, not hanya mudah tetapi juga lebih perustakaan yang mendukung untuk aplikasi di Android. Pocketsphinx adalah salah satu perpustakaan yang digunakan untuk pengenalan suara. Dengan Pocketsphinx, HMM algoritma dapat implementasi pada Android. Akurasi lebih dari 90%, sehingga Pocketsphinx yang lebih baik untuk menjadi mengembangkan aplikasi Android. Dalam tulisan akan menjelaskan, bagian pertama HMM yang digunakan di perpustakaan Pocketsphinx, bagian kedua menjelaskan beberapa perpustakaan untuk mendukung perpustakaan Pocketsphinx, pelaksanaan ketiga dan desain sistem pada perangkat Android.

2. Dasar Teori dan Perancangan

2.1 Speech Recognition

Automatic speech recognition (ASR) merupakan pengembangan teknik dan sistem yang memungkinkan komputer untuk menerima masukan berupa kata yang diucapkan. Dalam kemajuan teknologi informasi ini memungkinkan perangkat yang dapat mengenali kata yang diucapkan. Proses pengenalan kata, dilakukan dengan merubah sinyal analog menjadi sinyal digital dengan pola tertentu yang tersimpan di suatu perangkat. Dalam proses pengenalan suara membutuhkan sebuah rangkaian Algoritma. Algoritma yang digunakan sesuai dengan kebutuhan pengguna. Proses pengenalan suara yang diucapkan dalam kata memiliki beberapa tahapan, Berikut ini tahapan proses pengenalan suara,

- a. Tahap Penerimaan masukan
Pada proses ini pengguna memasukkan kata-kata untuk dikenali melewati pengeras suara suatu perangkat.
- b. Tahap Ekstraksi
Pada proses ini dilakukan tahap penyimpanan masukan yang berupa suara dan pembuatan basis data sebagai pola. Dalam jurnal ini proses ekstraksi dilakukan berdasarkan Algoritma Hidden Markov Model (HMM). Sesuai dengan kata 'hidden', dalam algoritma ini perlu menentukan parameter-parameter tersembunyi dari parameter yang dapat diamati.

Tahapan ekstraksi pengenalan ucapan berdasarkan HMM :

1. Tahap ekstraksi tampilan
Penyaringan sinyal suara dan pengubahan sinyal suara analog ke digital
 2. Tahap tugas pemodelan
Pembuatan suatu model HMM dari data-data yang berupa sampel ucapan sebuah kata yang sudah berupa data digital
 3. Tahap sistem pengenalan HMM
Penemuan parameter-parameter yang dapat merepresentasikan sinyal suara untuk analisis lebih lanjut
- c. Tahap Pembandingan
Tahap pembandingan ini merupakan tahap pencocokkan data baru dengan data suara (pencocokkan tata bahasa) pada pola. Dari proses konversi sinyal suara digital hasil dari proses ekstraksi ke dalam bentuk spectrum suara yang akan dianalisa dengan membandingkan dengan pola berbasis data. Dari gelombang kontinu ke dalam bentuk diskrit. Proses selanjutnya adalah :

1. Transformasi gelombang diskrit menjadi data yang terurut. Dimana gelombang diskrit berbentuk masukan berukuran n yang menjadi objek yang akan dibagi pada proses konversi
2. Menghitung frekuensi pada tiap elemen data yang terurut

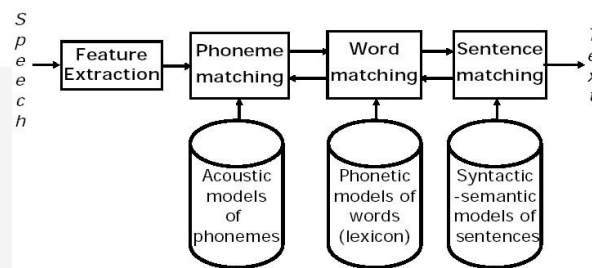
Kemudian dari tiap elemen data yang terurut akan dikonversi menjadi bentuk bilangan biner. Data biner nantinya akan dibandingkan dengan pola data suara dan kemudian diterjemahkan sebagai keluaran yang dapat berbentuk tulisan ataupun perintah pada perangkat.

d. Tahap validasi identitas pengguna

Pada sistem ini, alat pengenalan ucapan sudah memiliki verifikasi atau identifikasi suara. Suara akan melakukan identifikasi dengan orang yang berbicara berdasarkan kata yang diucapkan. Setelah itu menerjemahkan suara menjadi tulisan atau perintah.

2.1.1. Automatic Speech Recognition Process

Dalam pemrosesan mengenali sinyal suara secara umum terdapat empat tahapan yang harus dilakukan yaitu mengubah sinyal suara menjadi feature extraction untuk kemudian dilakukan pendeteksian phoneme yang cocok dengan sinyal suara tersebut. Gabungan dari phoneme-phoneme akan membentuk kata, dan banyak kata akan membentuk suatu kalimat yang prosesnya disebut dengan *sentence matching* seperti ditunjukkan pada gambar 2.6. Dari keempat proses tersebut akan didapatkan hasil akhir berupa teks yang berupa kalimat.



Gambar 2.1.1 Arsitektur umum ASR

2.2 Hidden Markov Model (HMM)

Hidden markov model mempunyai ciri khas yang mendasar didalam HMM suatu deretan *state* tidak dapat diaati secara langsung (tersembunyi), yang hanya dapat diamati hanyalah deretan simbol-simbol observasi. HMM adalah sebuah model statistik dari sebuah sistem yang diasumsikan sebuah Proses Markov dengan parameter yang tidak diketahui. dan tantangannya adalah menentukan parameter-parameter tersembunyi (state) dari parameter-parameter yang diamati (observer). Parameter-parameter yang ditentukan kemudian dapat digunakan untuk analisis yang lebih jauh. Untuk penjelasan HMM yang lebih detail bias dilihat pada [9].

2.3 Training Acoustic Model

Algoritma forward backward umumnya digunakan untuk melatih sebuah HMM yang menggunakan fungsi Gaussian sebagai penentu peluang observasi. algoritma maju dan mundur digunakan untuk mencari peluang suatu observasi jika diketahui model. Prinsip dari algoritma mundur hampir sama dengan algoritma maju. Jika Algoritma maju menghitung probabilitas berada pada suatu keadaan tertentu setelah melihat t observasi pertama, maka algoritma mundur menghitung peluang melihat observasi dari waktu $t + 1$ sampai akhir jika diketahui berada pada suatu state tertentu.

a. Baum Welch

Algoritma Baum Welch merupakan algoritma yang digunakan pada proses training dengan mengoptimalkan nilai parameter Model Markov Tersembunyi atau Hidden Markov Model Method

(HMM). Dalam menentukan rangkaian yang belum diketahui, algoritma Baum Welch mengoptimalkan nilai parameter yang di berikan. Algoritma Baum Welch digunakan untuk menentukan parameter yang tidak diketahui HMM.

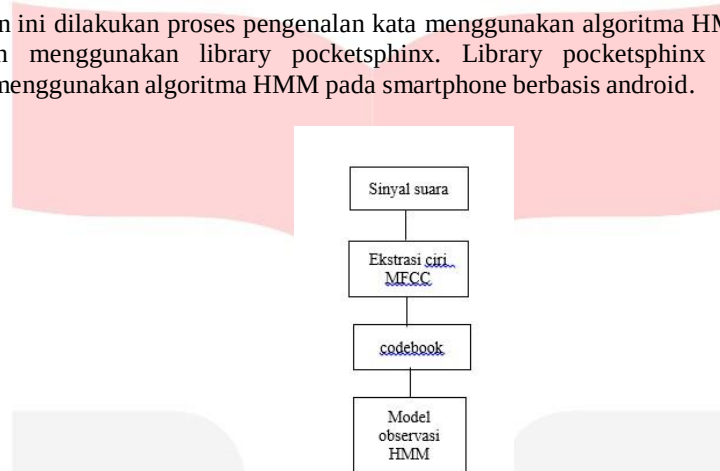
b. Normalisasi

Normalisasi adalah proses mengubah relasi menjadi bentuk normal atau proses untuk mengidentifikasi dan mengeliminasi anomali.

3. Pembahasan

3.1 Alur Speech recognition

Pada penelitian ini dilakukan proses pengenalan kata menggunakan algoritma HMM. Pada aplikasi ni akan dibangun dengan menggunakan library pocketsphinx. Library pocketsphinx digunakan sebagai tool pengolahan kata menggunakan algoritma HMM pada smartphone berbasis android.



Gambar 3.1 Proses speech recognition menggunakan HMM

3.2 Perancangan

3.2.1 Language Model Training

Data yang digunakan untuk membangun language model pada tugas akhir ini menggunakan kosa kata bahasa inggris sesuai dengan kamus yang ada pada pocketsphinx tersebut.

3.2.2 Grammar

Grammar di sini berfungsi sama seperti language model yaitu untuk membangun kalimat yang benar. Cmusphinx tidak menyediakan tools untuk membangun grammar karena grammar memang harus didefinisikan sendiri oleh developer ASR sesuai dengan pola kalimat yang ingin dikenali. Pada tugas akhir ini digunakan penulisan grammar dengan format *Java Speech Grammar Format* (JSGF).

3.2.3 Dictionary

Satu komponen lagi yang diperlukan dalam pembangunan aplikasi ini adalah dictionary. Dictionary di sini berfungsi sebagai vocabulary yang berisi daftar kata dan phoneme pembentuknya. Seluruh kata yang ada pada kalimat di language model maupun grammar harus didefinisikan phoneme pembentuknya di sini.

Daftar Pustaka :

- [1] Seltzer, Michael L. *Microphone Array Processing for Robust Speech Recognition*. Pittsburgh: Carnegie Mellon University, 2003.
- [2] Juang, B. H. and L. R. Rabiner. *Hidden Markov Models for Speech Recognition*. Speech Research Department AT&T Bell Laboratories Murray Hill, 2007
- [3] Abdulla, Waleed and Nikola Kasabov. *The Concepts of Hidden Markov Model in Speech Recognition*. University of Otago New Zealand, 1999.
- [4] Siivola, Vesa and Bryan L. Pellom. *Growing an n-gram Language Model*. University of Colorado at Boulder, USA.







3.3

3.4

3.2

4. ANALYSIS

5. CONCLUSION