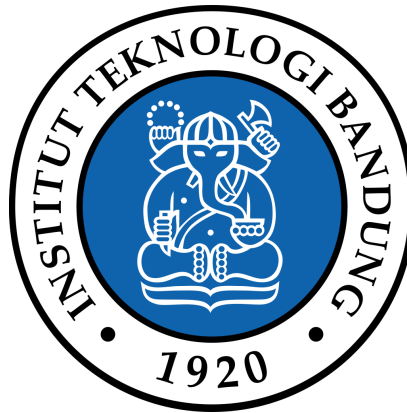


LAPORAN TUGAS BESAR 1
IF4071 Pemrosesan Ucapan
Small Vocabulary Speech Recognizer
Fase 2 : Data Preparation



Disusun oleh

Cynthia Athena Mahadewi S.	13518059
Annisa Rahim	13518089
Difa Habiba Rahman	13518098
Stefanus Gusega Gunawan	13518149

Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung
2021

1. Overview Aplikasi

Aplikasi yang kami buat adalah aplikasi untuk mendengarkan perintah terkait dengan alarm. Perintah yang dapat menjadi masukan, yaitu pasang alarm di setiap jam antara pagi/siang/sore/malam, matikan alarm, dan batalkan alarm.

2. Penjelasan Alur Data

a. Task Grammar

Grammar pada *Speech Recognition Engine (SRE)* akan mendefinisikan batasan pada input. Saat terdapat kata-kata yang telah didefinisikan pada grammar, didengar oleh SRE, SRE akan mengembalikan kata-kata yang didengar berdasarkan grammar tersebut. Kemudian, dialog manager akan melakukan processing dari hasil kata-kata yang dikembalikan. Pada julius, recognition grammar ini dipisah ke dalam dua file, yaitu grammar file dan voca file.

i. Grammar file

Grammar file pada julius akan mendefinisikan batasan input menggunakan format BNF dan tidak berisi list dari seluruh kata, namun hanya kategori kata. File tersebut tersusun dengan format :

Symbol: [Expression with Symbol]

Dimana :

Symbol = nonterminal

[Expression with Symbol] = ekspresi dari simbol yang berisi urutan dari beberapa simbol dan dapat berupa nonterminal dan terminal

S : NS_B COMMAND NS_E
COMMAND : LOOKUP OBJECT
COMMAND : SET OBJECT AT NUMBER TIME

ii. Voca file

Voca file berisi definisi kata dari setiap kategori kata yang ada pada file grammar. File ini tersusun dengan format :

```
%[Word Category]  
[Word Definition] [pronunciation...]
```

Berikut adalah file .voca yang kami gunakan.

```
% NS_B  
<s> sil  
  
% NS_E  
</s> sil  
  
% LOOKUP  
MATIKAN m aa t ih k aa n  
BATALKAN b aa t aa l k aa n  
  
% SET  
PASANG p aa s aa ng  
  
% OBJECT  
ALARM aa l aa r m  
  
% AT  
PUKUL p uw k uw l  
  
% NUMBER  
SATU s aa t uw  
DUA d uw aa  
TIGA t ih g aa  
EMPAT ah m p aa t  
LIMA l ih m aa  
ENAM ah n aa m  
TUJUH t uw j uw h  
DELAPAN d ah l aa p aa n  
SEMBILAN s ah m b ih l aa n  
SEPULUH s ah p uw l uw h  
SEBELAS s ah b ah l aa s  
DUABELAS d uw aa b ah l aa s  
  
% TIME  
PAGI p aa g ih  
SIANG s ih aa ng  
SORE s ao r eh  
MALAM m aa l aa m
```

b. Pronunciation Dictionary

Pronunciation dictionary berisi cara berbicara atau pronunciation untuk suatu kata. Untuk HTK dapat melakukan compile menjadi acoustic model, diperlukan phonetically balanced dictionary, dimana harus terdapat 30-40 kalimat dengan 8-10 kata tiap kalimat. Setiap phoneme perlu muncul 3-5 kali pada dictionary. Untuk membuat pronunciation dictionary ini, diperlukan file prompts, dan file indonesian dictionary.

i. Prompts

Prompts file berisi kata-kata dari grammar dan tambahan kata agar tercipta phonetically balanced dictionary. Setiap list kata yang perlu direkam, akan dicantumkan, dengan format setiap satu rekaman satu baris. Total samples pada eksperimen ini berisi ada 60 sampel dengan 9:1 = train:test.

```
*/sample1 PASANG ALARM PUKUL SATU PAGI DUA SIANG SEMBILAN MALAM
*/sample2 MATIKAN ALARM PASANG PUKUL TIGA SIANG SEBELAS MALAM
*/sample3 BATALKAN ALARM MATIKAN ALARM PASANG ALARM PUKUL EMPAT
MALAM
*/sample4 PASANG ALARM DUA PAGI EMPAT SORE TUJUH MALAM
*/sample5 AKU PASANG ALARM PUKUL SEPULUH PAGI DAN ENAM SORE
*/sample6 PAGI TIDUR SIANG MAKAN SORE KERJA MALAM BANGUN
*/sample7 BATALKAN PUKUL SEPULUH MALAM PASANG JAM LIMA PAGI
*/sample8 TULIS EMPAT ENAM DUABELAS SEBELAS KERJA SORE BATALKAN
*/sample9 MALAS BANGUN MAKAN KERJA DUABELAS SEPULUH DELAPAN
LIMA
*/sample10 MATIKAN ALARM AKU DAN TIGA ENAM LIMA TUJUH
*/sample11 BANGUN TIDUR AKU MATIKAN ALARM PASANG PUKUL SEBELAS
PAGI
*/sample12 MAKAN PUKUL SATU PAGI DUA SIANG EMPAT SORE TUJUH
MALAM
*/sample13 SENIN PASANG ALARM JAM DUA SIANG BATALKAN PUKUL TIGA
*/sample14 BATALKAN PAGI MATIKAN SIANG PASANG SEPULUH ALARM
MALAM
*/sample15 MALAS BANGUN MATIKAN ALARM PUKUL LIMA AKU TIDUR
*/sample16 PASANG ALARM MATA SEMBILAN PAGI TUJUH MALAM SIANG
SORE
*/sample17 MATIKAN SEMBILAN RIBU PAGI SORE MALAM TUJUH DELAPAN
*/sample18 SEMBILAN SIANG TUJUH PAGI SORE ENAM BATALKAN SIKAT
*/sample19 BATALKAN SATU DUA TIGA EMPAT LIMA MATI TUJUH
*/sample20 PUKUL ALARM DUA KAKI AKU TIGA SIANG SEMBILAN
*/sample21 PASANG GIGI DUA MALAM TUJUH SORE SIANG EMPAT
*/sample22 BUANG PASANG BATALKAN MATIKAN DUA TIGA SORE SIANG
*/sample23 PASANG ALARM PUKUL LIMA PAGI TAPI SEBELAS SIANG
*/sample24 TADI PAGI PUKUL LIMA PASANG ALARM DUA SIANG
*/sample25 TIBA SORE PUKUL DUABELAS SIANG SEPULUH TUJUH ENAM
*/sample26 PASANG ALARM PUKUL LIMA BANGUN SEBELAS SIANG DUA
```

```

*/sample27 BATU ALARM HABIS BELI PUKUL TUJUH MALAM PASANG
*/sample28 NOTA SAPI ALARM DUA PAGI TUJUH LIMA NASI KAMBING
*/sample29 PASANG ALARM SATU DUA SEMBILAN MALAM TUJUH DUABELAS
ENAM
*/sample30 LALU AKU TIDUR BANGUN PUKUL LIMA PAGI LAGI
*/sample31 DUABELAS INGIN MATIKAN ALARM JAM DELAPAN DAN PERGI
MAKAN
*/sample32 PASANG LAGI JAM MAU MANDI BOLA SEKARANG ALARM
SEBELAS
*/sample33 PADA HARI MINGGU AKU IKUT AYAH KE KOTA SEPULUH
*/sample34 PASANG ALARM PUKUL TIGA EMPAT LIMA JAM BUAH
*/sample35 PUKUL SEMBILAN PASANG ALARM REM MOBIL JAM TIGA
*/sample36 AKU PERGI MAKAN PUKUL TIGA SORE BATALKAN ALARM
*/sample37 BERANGKAT PADA PUKUL TUJUH PAGI DAN TUJUH MALAM
*/sample38 ATUR UNTUK PASANG ALARM PUKUL LIMA PAGI NANTI
*/sample39 PAGI SIANG SORE MALAM BACA BUKU PASANG ALARM
*/sample40 BATALKAN ALARM PASANG ALARM PUKUL SEPULUH PAGI INI
*/sample41 TAKUT TIDAK BANGUN PASANG ALARM PUKUL TIGA PAGI
*/sample42 NANTI ADA RAPAT JADI PASANG ALARM PUKUL EMPAT
*/sample43 TANGGAL DUABELAS TADI BILANG MATIKAN ALARM ITU HARUS
*/sample44 ADA ANGIN DINGIN AKU BATALKAN ALARM PUKUL LIMA PAGI
*/sample45 BATALKAN ALARM PUKUL SEPULUH SIANG UNTUK ACARA
SEKOLAH
*/sample46 COBA PASANG DELAPAN SEMBILAN SEPULUH SEBELAS
DUABELAS AYAM
*/sample47 PUKUL ALARM DELAPAN TIGA DUA MATIKAN TERUS DIAM
*/sample48 BATALKAN ALARM SEBELAS SEBELAS TERUS BANYAK DELAPAN
KALI
*/sample49 ALARM PASANG BATALKAN PUKUL MATIKAN SATU SATU SEMUA
*/sample50 DELAPAN KALI COBA PUKUL BATALKAN ALARM PUKUL SATU
MALAM
*/sample51 DELAPAN LAGI SUPAYA SENYUM COBA SATU JUGA SATU
*/sample52 ENAM KURANG ENAM KAMBING MATIKAN AYAM DUABELAS KALI
*/sample53 BATALKAN PASANG MATIKAN ENAM DELAPAN SEMBILAN
SEPULUH SEBELAS
*/sample54 DUABELAS MALAM LEPAS SEPATU SORE SIANG MATIKAN SATU
*/sample55 SATU TAMBAH SATU BAGI ENAM SORE SIANG MATIKAN
*/sample56 EMPAT ENAM DELAPAN SEMBILAN SEPULUH SEBELAS DUABELAS
TIGA
*/sample57 CARI KUCING SATU TIGA EMPAT ENAM DELAPAN SEMBILAN
*/sample58 SEPULUH SEBELAS DUABELAS SORE MALAM BATALKAN EMPAT
ENAM
*/sample59 DELAPAN SEMBILAN SEPULUH SEBELAS DUABELAS SATU TIGA
EMPAT
*/sample60 SENYUM KUNYAH BAU BAUT SATE SATU DUA TIGA

```

ii. Lexicon (Indonesian dictionary)

Dictionary file berisi informasi pronunciation, termasuk kata-kata yang dibutuhkan aplikasi, serta kata-kata yang tidak dibutuhkan aplikasi dan tercantum pada file prompts. Untuk aplikasi ini,

dictionary yang digunakan adalah indonesian dictionary karena kata perintah atau kata masukan menggunakan bahasa Indonesia.

```
...
AB      aa b
ABA      aa b aa
ABA-ABA    aa b aa aa b aa
ABAD      aa b aa d
ABADI      aa b aa d ih
ABADIAH    aa b aa d ih aa hh
ABAH      aa b aa hh
ABAH-ABAH  aa b aa hh aa b aa hh
...
```

c. Transcription file

Words level Transcription

HTK tidak dapat langsung melakukan compile terhadap file prompts, diperlukan adanya file master label file yang berisi label entry dari setiap line di file prompts. Ini dihasilkan dengan code dibawah ini yang akan menghasilkan **words.mlf** :

```
julia ../bin/prompts2mlf.jl prompts.txt words.mlf
```

Phone level Transcription

Dengan menggunakan HLEd command, word level transcription akan ditransformasi menjadi phone level transcription.

d. MFCC Feature

HTK hanya dapat memproses dalam bentuk MFCC feature (berbentuk *file* dengan *extension* **.mfc**, maka dari itu hasil rekaman dalam format wav, perlu dikonversi dengan daftar *file* yang akan dikonversi terdapat di *file* **codetrain.scp** menggunakan command:

```
HCopy -A -D -T 1 -C wav_config -S codetrain.scp
```

Untuk menjalankan *command* tersebut, digunakan **wav_config** dengan isi sebagai berikut.

```
SOURCEFORMAT = WAV
TARGETKIND = MFCC_0_D
TARGETRATE = 100000.0
SAVECOMPRESSED = T
SAVEWITHCRC = T
WINDOWSIZE = 250000.0
USEHAMMING = T
PREEMCOEF = 0.97
NUMCHANS = 26
CEPLIFTER = 22
NUMCEPS = 12
```

3. Tautan Terkait

Github :

<https://github.com/stefanusgusega/tubes-speech>

Recording Samples :

<https://drive.google.com/drive/folders/1pYCLQggMh-ZHQbbbC4hFT2WcvF53tkS?usp=sharing>