



Available online at : <http://bit.ly/InfoTekJar>

InfoTekJar : Jurnal Nasional Informatika dan Teknologi Jaringan

ISSN (Print) 2540-7597 | ISSN (Online) 2540-7600



Perbandingan Hasil Prediksi Durasi Fonem pada Bahasa Melayu Pontianak Berdasarkan Suku Kata Menggunakan Hidden Markov Model

Luthfia Justisia Loebis, Arif Bijaksana Putra Negara, Novi Safriadi

Jurusan Informatika, Fakultas Teknik, Universitas Tanjungpura, Jalan Prof. Dr. H. Hadari Nawawi, Pontianak, 78124, Indonesia

KATA KUNCI

Bahasa Melayu Pontianak, Durasi Fonem, Hidden Markov Model.

CORRESPONDENCE

E-mail: luthfia.justisia@gmail.com

A B S T R A C T

TTS is a technology that commonly used for preserving languages. One of the most important indicators to producing a good TTS speech is the right duration of phonemes. Phonemes duration can be predicted by HMM. The prediction is built based on syllables with the aim to reducing the size of the corpus. To predicting phonemes duration based on syllables, all the sentences must be converted into syllable codes. Also, the variations of phonemes duration are reduced by grouping them into 3 and 5 states. The tests carried out were comparison of phonemes and sentences duration test. Both tests were operated with two n-gram modelling, bigram and trigram. The test is preceded by determining the baseline score and then the next test was carried by k-fold cross validation in two types of corpus data dividing, a corpus training that contains 80% of corpus and a corpus training that contains 20% of corpus. The baseline score for the comparison of phonemes duration test using 3 states duration, carried out by the two n-gram models, both get 48%, while for the test using 5 states duration, each gets 48% and 69%. The baseline score of the comparison of sentences duration test using 3 states duration, also carried out by the two n-gram models, respectively gets 21% and 19%, while for the test using 5 states duration, gets 32% and 17% for each. Based on the test results using the k-fold cross validation method, it can be concluded for the used of 80% corpus training, the results are near the baseline score, while for the used of 20% corpus training the results are a little far from the baseline score. From the overall results of the test, it can be concluded that based on the use of the n-gram modelling, trigram get better results. Whereas, based on the use of the duration states, the used of 3 states got better results.

ABSTRAK

TTS merupakan teknologi yang kerap dimanfaatkan untuk melestarikan eksistensi bahasa. Salah satu indikator terpenting untuk menghasilkan ucapan TTS yang baik adalah durasi fonem yang tepat. Durasi fonem dapat diprediksi dengan HMM. Prediksi dilakukan berdasarkan suku kata untuk memperkecil korpus. Dalam melakukan prediksi durasi fonem berdasarkan suku kata, korpus dikonversi ke dalam bentuk kode suku kata. Selain itu dilakukan reduksi variasi fonem ke dalam pengelompokkan 3 dan 5 state. Pengujian yang dilakukan adalah pengujian perbandingan durasi fonem dan kalimat, dengan model n-gram bigram dan trigram. Pengujian didahului dengan penentuan nilai baseline, kemudian pengujian dengan k-fold cross validation untuk dua jenis pembagian data, yaitu korpus latih berisi 80% korpus dan korpus latih berisi 20% korpus. Nilai baseline pada pengujian perbandingan durasi fonem untuk durasi 3 state dengan kedua model n-gram mendapatkan hasil 48%, sedangkan untuk durasi 5 state masing-masing mendapatkan hasil 48% dan 69%. Nilai baseline pada pengujian perbandingan durasi kalimat untuk durasi 3 state dengan kedua model n-gram, masing-masing mendapatkan hasil 21% dan 19%, sedangkan untuk durasi 5 state masing-masing mendapatkan hasil 32% dan 17%. Dari hasil pengujian dengan k-fold cross validation diketahui bahwa dalam penggunaan korpus latih 80%, hasil yang didapatkan mendekati nilai baseline, sedangkan hasil pada penggunaan korpus latih 20% menjauhi nilai baseline. Dari hasil keseluruhan pengujian dapat disimpulkan berdasarkan model n-gram yang digunakan, trigram mendapatkan hasil yang lebih baik. Sedangkan berdasarkan state durasi yang digunakan, 3 state mendapatkan hasil lebih baik

PENDAHULUAN

Bahasa merupakan sistem lambang bunyi yang digunakan oleh para anggota suatu masyarakat untuk saling bekerja sama, berinteraksi, dan mengidentifikasi diri [1]. Dalam berkomunikasi setiap negara memiliki bahasa nasional dan bahasa-bahasa daerah. Berdasarkan hasil sensus penduduk pada tahun 2010 menunjukkan Indonesia memiliki kurang lebih 1.158 bahasa daerah [2].

Masih berdasarkan hasil sensus penduduk yang sama, menunjukkan bahwa persentase penggunaan bahasa Melayu Pontianak di Kalimantan Barat mencapai 20,45% atau sejumlah 1.615.978 juta jiwa dari total seluruh penduduk Kalimantan Barat [2]. Sayangnya tuturan yang terjadi bukan lagi menggunakan bahasa Melayu Pontianak yang asli melainkan bahasa Indonesia dengan dialek Melayu Pontianak [3]. Hal ini dapat menyebabkan kepunahan bahasa Melayu Pontianak di kemudian hari. Oleh sebab itu perlu dilakukannya pelestarian terhadap Bahasa Melayu Pontianak.

Pelestarian bahasa Melayu Pontianak dapat dilakukan dengan memanfaatkan teknologi *Text to Speech* (TTS). TTS merupakan sebuah sistem yang mampu melakukan konversi dari teks menjadi sintesa ucapan. Hasil sintesa ucapan TTS dinilai dari dua parameter, yaitu *Intelligibility* dan *Naturalness* [4]. *Intelligibility* merupakan parameter yang menyatakan tingkat kejelasan sintesa ucapan, sedangkan *naturalness* merupakan parameter yang menyatakan kealamian suatu sintesa ucapan. Untuk menghasilkan sintesa ucapan yang baik harus dibangun dengan fonem yang memiliki durasi yang tepat, ini dapat dilakukan dengan memprediksi terlebih dahulu durasi untuk setiap fonem. Metode yang terus berkembang dalam bidang sintesa ucapan saat ini adalah metode *HMM-Based Speech Synthesis System* (HTS) sebuah implementasi metode statistik parametrik dengan menggunakan *Hidden Markov Model*. HMM adalah pengembangan model statistik dari model Markov. HMM bekerja dengan baik pada proses prediksi dan klasifikasi yang bersifat chain, sehingga model ini sangat cocok diterapkan untuk prediksi durasi fonem. Selain itu HMM merupakan proses stokastik ganda, sehingga cocok digunakan dalam prediksi durasi fonem. Namun untuk memprediksi durasi fonem menggunakan HMM dibutuhkan data korpus yang sangat besar. Ini dapat direduksi dengan memprediksi durasi fonem berdasarkan suku kata. Prediksi durasi fonem berdasarkan suku kata dilakukan dengan mengkonversi korpus ke dalam bentuk kode. Hal ini bertujuan agar setiap kode dapat mewakili banyak suku kata yang memiliki bentuk kode yang sama. Kode berdasarkan suku kata ini direpresentasikan oleh pola suku kata. Pola suku kata terbentuk berdasarkan jenis fonem. Jenis fonem terdiri dari fonem konsonan yang dilambangkan dengan huruf "k" dan fonem vokal yang dilambangkan dengan huruf "v".

Selain itu, meskipun HMM mampu melakukan prediksi terhadap data yang bervariasi, namun jumlah variasi yang digunakan terdapat batasan agar HMM mampu bekerja dengan baik. Maka dari itu dalam penelitian ini durasi fonem direduksi variasinya dengan cara mengelompokkan durasi ke dalam beberapa *state*. Cara ini umum digunakan dalam sejumlah penelitian prediksi dengan HMM.

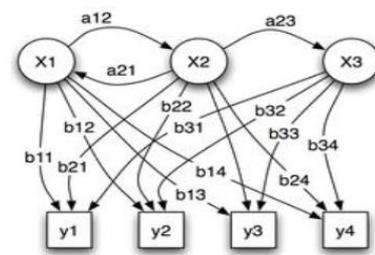
Berdasarkan penjelasan yang telah dipaparkan, maka dilakukan penelitian memprediksi durasi fonem bahasa Melayu Pontianak berdasarkan suku kata, dengan durasi fonem yang telah dikelompokkan dalam beberapa *state*, menggunakan *Hidden Markov Model*.

TINJAUAN PUSTAKA

A. Hidden Markov Model

Hidden Markov Model (HMM) adalah suatu proses stokastik ganda, di mana suatu sistem yang diasumsikan berupa Markov proses dengan kondisi yang tidak terobservasi (*hidden*) [5]. Pada *Markov Model* biasa, setiap *state* dapat dilihat secara langsung oleh pengamat. Oleh karena itu, kemungkinan dari transisi antar *state* (*state transition*) menjadi satu-satunya parameter yang dapat diamati. Padahal dalam suatu perhitungan atau pemrosesan urutan kejadian penting untuk diamati. Guna menyelesaikan masalah ini, dikembangkanlah model HMM.

Dalam HMM terdapat 2 *state*, yaitu *observed state* dan *hidden state*. *Observed state*, yaitu bagian yang berisi urutan kejadian yang dapat diamati secara langsung dan *hidden state*, yaitu bagian yang tersembunyi sehingga tidak dapat diamati secara langsung. Dalam penggunaannya pada NLP, *observed state* merupakan data yang telah dimiliki atau sudah diketahui, sedangkan *hidden state* merupakan data yang ingin dicari. Maka dari itu pada penelitian ini kode suku kata merupakan *observed state* dan durasi fonem merupakan *hidden state*. Proses probabilistik HMM sendiri dapat dilihat pada gambar 1.



Gambar 1. Proses probabilistik HMM [6]

Ket : x_i = keadaan tersembunyi (*hidden state*)
 y_i = keadaan terobservasi (*observed state*)
 a_{ij} = probabilitas transisi
 b_{ij} = probabilitas emisi

Dari proses probabilistik di atas sebuah keadaan N pada HMM dapat didefinisikan dengan persamaan sebagai berikut :

$$\text{HMM: } \lambda = \{A, B, \Pi\}$$

Ket : A = keadaan transisi ($\{a_{ij}\}_{i,j=1}^N$)
 B = keluaran probabilitas distribusi ($\{b_i(o)\}_{i=1}^N$)
 Π = probabilitas awal keadaan ($\{\pi_i\}_{i=1}^N$)

B. Bahasa Melayu Pontianak

Bahasa Melayu Pontianak merupakan satu di antara banyaknya bahasa daerah yang terdapat di Provinsi Kalimantan Barat. Bahasa ini dituturkan oleh masyarakat melayu di Kota Pontianak, Kabupaten Mempawah, dan Kabupaten Kubu Raya. Dalam banyak kosa kata, bahasa Melayu Pontianak hampir sama dengan bahasa Indonesia. Hal ini tidak mengherankan karena mengingat bahasa Indonesia memang berakar dari bahasa Melayu.

C. Fonem

Fonem merupakan satuan terkecil dalam bunyi pengucapan bahasa. Fonem mempunyai peran penting dalam memberi makna pada setiap kata dalam sebuah bahasa, maka dari itu perbedaan satu buah fonem saja dapat memberikan arti yang berbeda. Pada bahasa Indonesia 31 fonem. Terdiri dari 6 fonem vokal, yaitu /i/, /e/, /ə/, /a/, /u/, dan /o/ serta 18 fonem konsonan /p/, /t/, /c/, /k/, /b/, /d/, /j/, /g/, /m/, /n/, /ŋ/, /l/, /s/, /h/, /t/, /l/, /w/, dan /y/. Kemudian terdapat 4 fonem tambahan yang berasal dari bahasa

asing, yaitu /x/, /z/, /f/, dan /j/, serta 3 fonem diftong /ai/, /au/, dan /oi/ [7]. Perbedaan pengaturan fonem pada bahasa Melayu Pontianak adalah tidak adanya empat fonem tambahan yang berasal dari bahasa asing, yaitu /x/, /z/, /f/, dan /j/, selain itu terdapat perbedaan teratur yang meliputi akhiran kata pada sejumlah fonem, yaitu /a/-ə/, /i/-e/, /u/-o/, /r/-k/, dan /q/-k/ [8].

C. Suku Kata

Suku kata atau silabel adalah bagian kata yang diucapkan dalam satu hembusan napas. Guna memperoleh suku kata, perlu dilakukan pemenggalan terhadap kata. Jenis suku kata juga dapat dilihat dari pola suku kata yang terbentuk. Pola suku kata ini diperoleh dari jenis fonem pembangun suku kata yang terdiri dari fonem konsonan dan vokal.

D. Durasi

Durasi adalah suatu jenis suprasegmental yang ditandai oleh panjang-pendeknya waktu yang diperlukan untuk pengucapan sebuah bunyi. Durasi sering juga disebut dengan kuantitas karena menyangkut jumlah atau lamanya waktu yang diperlukan untuk pengucapan sebuah bunyi [9]. Sehingga durasi menjadi tolak ukur untuk lama atau tidaknya sebuah bunyi diucapkan. Setiap pengucap dapat memiliki durasi yang berbeda-beda walaupun bunyi yang diucapkan sama.

E. Python

Python merupakan bahasa pemrograman yang bersifat *open source multi-platform*, sehingga dapat digunakan pada berbagai macam sistem operasi. *Python* merupakan bahasa pemrograman tingkat tinggi yang dewasa ini telah menjadi standar dunia komputasi ilmiah [10]. Pada penelitian ini bahasa pemrograman *python* digunakan untuk mengkonversi korpus menjadi sejumlah kode berdasarkan suku kata dengan aturan pada setiap kode berupa pola suku kata yang terdiri dari 2 jenis fonem, yaitu konsonan (k) dan vokal (v).

F. K-Fold Cross Validation

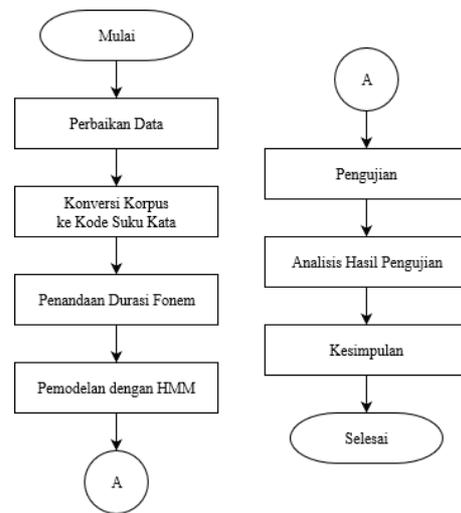
K-fold adalah salah satu metode cross validation yang populer dengan melipat data sebanyak K dan melakukan iterasi sebanyak K juga [11]. Ilmuwan membuat set data K training-validation dengan memilih salah satu kelompok sebagai validation set dan menggabungkan semua kelompok yang tersisa ke dalam training set seperti yang diilustrasikan pada gambar 2 [11].



Gambar 2. Ilustrasi pembuatan *k-fold cross validation set*

METODE PENELITIAN

Langkah-langkah penelitian yang dilakukan pada penelitian ini dapat dilihat pada gambar 3 sebagai berikut:



Gambar 3. Diagram alir penelitian

Berdasarkan gambar 3 dapat dijelaskan metodologi penelitian sebagai berikut:

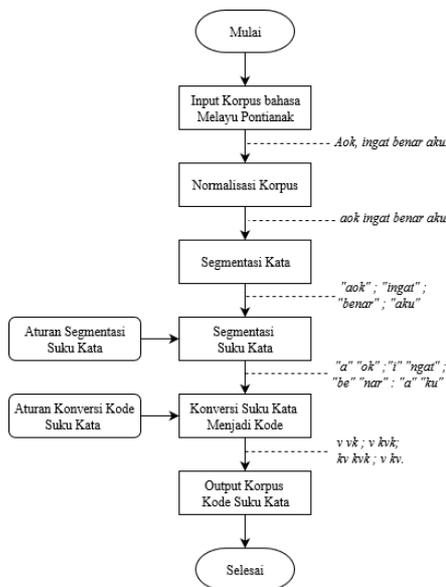
A. Perbaikan Data

Data yang digunakan merupakan data yang telah terkumpul sebelumnya. Terdiri dari korpus teks dan korpus suara. Kedua korpus berisi 526 kalimat bahasa Melayu Pontianak yang berasal dari sebuah buku berjudul “Sepok Satu” karangan Pay Jarot Sujarwo. Sedangkan korpus suara merupakan hasil rekaman suara seorang penutur bahasa Melayu Pontianak. Selain itu juga terdapat sejumlah data dalam format *excel* yang merupakan durasi pada tiap fonem pada korpus yang digunakan. Adapun terhadap data yang ada dilakukan sejumlah perbaikan guna untuk memaksimalkan kualitas hasil penelitian. Pada kedua korpus dilakukan seleksi terhadap kalimat yang ada, sehingga tidak terdapat kalimat yang mengandung kata atau istilah asing. Kalimat yang digunakan menjadi berjumlah 500 kalimat. Sedangkan untuk data durasi fonem dilakukan pengecekan ulang pada penandaan durasi untuk setiap fonem, hal ini dilakukan agar durasi yang digunakan lebih tepat. Pengecekan ulang dilakukan dengan menggunakan aplikasi *WaveSurfer 1.8.8p4*.

B. Konversi Korpus ke Kode Suku Kata

Korpus bahasa Melayu Pontianak yang sudah ada kemudian dikonversi dalam bentuk kode suku kata. Kode suku kata ini dipresentasikan dalam pola suku kata yang terbentuk dari jenis fonem pada suku kata. Terdapat 2 jenis fonem, yaitu fonem konsonan yang dilambangkan dengan “k” dan fonem vokal yang dilambangkan dengan “v”. Fonem konsonan terdiri dari /t/, /c/, /k/, /b/, /d/, /j/, /g/, /m/, /n/, /ng/, /ny/, /s/, /h/, /r/, /l/, /w/, dan /y/, sedangkan fonem vokal terdiri dari /i/, /e/, /ə/, /a/, /u/, /o/, /ai/, /au/, dan /oi/.

Konversi suku kata dilakukan dengan bantuan pengkodean algoritma dalam bahasa pemrograman *python*. Algoritma bekerja dengan cara menarik *file* korpus bahasa Melayu Pontianak sebagai *input* dan menghasilkan *file* data korpus dalam bentuk hasil konversi kode suku kata sebagai *output*. Adapun proses konversi korpus ke bentuk kode suku kata dapat dilihat pada gambar 4..



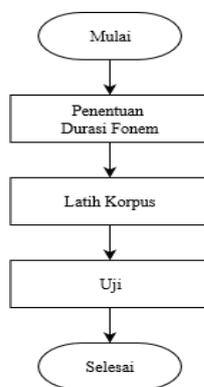
Gambar 4. Diagram alir proses konversi kode suku kata

C. Penandaan Durasi Fonem

Penandaan durasi fonem dilakukan pada korpus suku kata yang telah dihasilkan pada proses sebelumnya. Durasi yang digunakan merupakan durasi hasil rekaman suara penutur bahasa Melayu Pontianak. Penandaan dilakukan untuk keperluan proses latihan pada pemodelan HMM nantinya.

D. Pemodelan dengan HMM

Proses pemodelan dengan *Hidden Markov Model* (HMM) bertujuan untuk mendapatkan hasil prediksi durasi fonem pada bahasa Melayu Pontianak. Pada proses ini terdapat beberapa tahapan yang dapat dilihat pada gambar 5.



Gambar 5. Diagram alir pemodelan dengan HMM

E. Pengujian

Pengujian awal dilakukan untuk menentukan standar kualitas durasi fonem yang dihasilkan, yaitu pengujian penentuan standar kualitas ucapan. Sedangkan pengujian yang digunakan pada hasil prediksi durasi terdapat dua pengujian, yaitu pengujian perbandingan durasi durasi fonem dan pengujian perbandingan durasi kalimat.

F. Analisis Pengujian

Pada tahap ini, hasil pengujian dari prediksi durasi fonem pada bahasa Melayu Pontianak menggunakan HMM dianalisis secara keseluruhan untuk mempermudah dalam penarikan kesimpulan.

G. Kesimpulan

Kesimpulan dirumuskan berdasarkan tahapan-tahapan yang telah dilakukan sebelumnya yaitu untuk mengetahui tingkat keberhasilan prediksi durasi fonem menggunakan HMM pada bahasa Melayu Pontianak yang telah dilakukan.

HASIL DAN PEMBAHASAN

A. Hasil Konversi Korpus ke Kode Suku Kata

Dari hasil konversi kode suku kata didapatkan enam macam kode suku kata, yaitu kkv (konsonan-konsonan-vokal), kkvk (konsonan-konsonan-vokal-konsonan), kv (konsonan-vokal), kvk (konsonan-vokal-konsonan), v (vokal), dan vk (vokal-konsonan).

B. Implementasi Model HMM

Pada penelitian ini implementasi model HMM dilakukan dengan *tool IPOSTagger* karya Alfian Rizki Wicaksono (2010) [12]. Prediksi durasi fonem pada bahasa Melayu Pontianak menggunakan HMM dilakukan dalam dua proses, yaitu proses latihan korpus dan uji. Pada penelitian ini juga dilakukan proses penentuan durasi fonem yang digunakan terlebih dahulu.

1) Penentuan Durasi Fonem

Penentuan durasi fonem dilakukan guna memperkecil variasi durasi fonem, sehingga HMM dapat melakukan proses perhitungan. Ini dilakukan dengan mengelompokkan durasi dalam 3 dan 5 *state*. Pengelompokkan durasi dilakukan pada masing-masing jenis fonem, yaitu vokal dan konsonan. Pengelompokkan dilakukan dengan menarik tiga nilai tengah untuk durasi 3 *state* dan lima nilai tengah untuk durasi 5 *state*. Hasil dari penentuan durasi fonem untuk setiap fonem dalam 3 *state* dapat dilihat pada c dan dalam 5 *state* dapat dilihat pada Tabel 2.

Tabel 1. Hasil penentuan durasi fonem 3 state

State	Durasi Fonem Vokal (ms)	Durasi Fonem Konsonan (ms)
1	69	54
2	199	153
3	337	260

Tabel 2. Hasil penentuan durasi fonem 5 state

State	Durasi Fonem Vokal (ms)	Durasi Fonem Konsonan (ms)
1	43	35
2	121	94
3	199	153
4	280	212
5	377	302

2) Latih Korpus

Proses latihan korpus dilakukan dengan membuat sebuah korpus sebagai data latihan. Pada penelitian ini korpus yang digunakan berisi kalimat bahasa Melayu Pontianak yang telah dikonversi ke dalam kode suku kata dengan penandaan durasi fonem. Terdapat dua macam korpus latihan yang digunakan pada penelitian ini, yaitu korpus dengan penandaan durasi hasil pengelompokkan 3 *state* dan 5 *state*. Korpus yang telah dibuat disimpan sebagai file bernama "Corpus.crp". File ini tersimpan dalam subfolder "resource" pada folder tool *IPOSTagger*. Adapun contoh perintah proses latihan korpus dapat dilihat pada gambar 6.

```
Microsoft Windows [Version 10.0.17134.950]
(c) 2018 Microsoft Corporation. All rights reserved.

C:\Users\luthf>D:
D:\>cd IPOSTagger_v1.1
D:\IPOSTagger_v1.1>java ipostagger train_
```

Gambar 6. Contoh proses latihan korpus pada IPOSTagger

Dari proses latihan ini didapatkan dua file baru yang secara otomatis tersimpan pula pada subfolder "resource". File pertama bernama "Lexicon.trn", merupakan kamus data dari korpus latihan. Hasil dari perhitungan prediksi durasi yang dilakukan dapat diketahui kode suku kata beserta dengan penandaan durasi, serta frekuensi kemunculan masing-masing dursi tersebut. File kedua adalah "Ngram.trn" yang berisi jumlah frekuensi kemunculan setiap durasi fonem beserta kemungkinan pertemuan-pertemuan durasi tersebut sebagai hasil penandaan durasi fonem pada korpus latihan. Dari file ini dapat diketahui probabilitas pertemuan antar setiap durasi.

3) Uji

Proses uji dilakukan untuk mendapatkan durasi pada setiap fonem berdasarkan suku kata. Sama halnya dengan proses latihan untuk melakukan proses uji perlu dibuat sebuah korpus. Korpus uji berisi kalimat bahasa Melayu Pontianak yang telah dikonversi ke dalam bentuk kode suku kata. Proses uji ini dilakukan dengan menggunakan dua macam pemodelan n-gram, yaitu bigram dan trigram.

C. Hasil dan Analisis Pengujian

1) Pengujian Penentuan Standar Kecepatan Ucapan

Pengujian penentuan standar kecepatan ucapan bertujuan untuk menentukan standar perubahan kecepatan ucapan yang mampu diterima oleh pendengar. Pengujian ini juga dilakukan pada penelitian terhadap prediksi durasi fonem pada bahasa Melayu Pontianak oleh Numrotul Hatimah (2019) [13]. Penentuan standar dilakukan dengan pengujian terhadap satu kalimat yang digunakan sebagai standarisasi. Audio rekaman kalimat dimodifikasi durasinya dengan dipercepat dan diperlambat sebesar 10% hingga 50%, dengan perbedaan 10%. Sehingga dari satu audio rekaman tadi, menjadi sepuluh file suara dengan kecepatan durasi berbeda. Semakin tinggi rata-rata hasil pengujian didapat oleh suatu file suara, maka semakin mudah responden menerimanya. Hasil pengujian penentuan standar kecepatan ucapan dapat dilihat pada Tabel 3. Kemudian analisis terhadap hasil pengujian dapat dilihat pada tabel 4.

Tabel 3. Hasil Pengujian Standar Kecepatan Ucapan

File Suara	Modifikasi Durasi	Hasil Pengujian
1	30%	47%
2	-50%	13%
3	10%	100%
4	-20%	93%
5	40%	0%
6	-10%	100%
7	50%	0%
8	-30%	63%
9	20%	83%
10	-40%	47%

Tabel 4. Analisis Hasil Pengujian Standar Kecepatan Ucapan

Kualitas	Rentang Modifikasi Kecepatan Ucapan (%)
Sangat Baik	(-20%) s.d. 20%
Baik	(-30%) s.d. (-21%)
Buruk	(-40%) s.d. 30%
Sangat Buruk	(-50%) s.d. 50%

Hasil pada Tabel 3 kemudian dianalisis dengan mengelompokkan file suara berdasarkan hasil pengujian yang didapatkan. Berdasarkan analisis yang dilakukan diketahui bahwa perubahan kecepatan ucapan dalam rentang (-20%) sampai 20% termasuk kualitas "sangat baik". Kemudian untuk kualitas "baik" berada pada rentang perubahan kecepatan dari (-30%) sampai (-21%). Sedangkan perubahan kecepatan ucapan dalam rentang (-40%) sampai 30% termasuk kualitas "buruk". Terakhir untuk perubahan kecepatan berkualitas "sangat buruk" terjadi pada rentang (-50%) sampai 50%. Kualitas yang dapat diterima adalah "sangat baik" dan "baik".

2) Pengujian Perbandingan Durasi Fonem

Pengujian dilakukan dengan cara membandingkan durasi fonem hasil prediksi HMM dengan durasi fonem hasil rekaman suara penutur bahasa Melayu Pontianak pada kalimat yang sama. Perbandingan durasi fonem dinyatakan sebagai selisih perbedaan dalam bentuk persentase. Perbandingan ini dilakukan menggunakan perhitungan berikut.

$$\text{Persentase perbandingan perbedaan} = \frac{\text{Durasi HMM} - \text{Durasi Rekaman}}{\text{Durasi Rekaman}} \times 100$$

Pengujian dilakukan dengan menentukan nilai *baseline* terlebih dahulu. Kemudian dilakukan pengujian dengan menggunakan *k-fold cross validation* pada dua macam pembagian data korpus bahasa Mealyu Pontianak.

Nilai Baseline Durasi Fonem 3 State

Nilai *baseline* didapatkan dari rata-rata hasil pengujian. Hasil pengujian dapat dilihat pada Tabel 5. Hasil pengujian yang ada kemudian diolah kembali berdasarkan kualitas durasi pada Tabel 6.

Tabel 5. Hasil Nilai Baseline Durasi Fenom 3 State

Model n-gram	Nilai Baseline
Bigram	47,79%
Trigram	48,04%

Tabel 6. Hasil Nilai Baseline Dutsi Fenom 3 State Berdasarkan Kualitas

Kualitas	Jumlah Fonem		Persentase Fonem	
	Bigram	Trigram	Bigram	Trigram
Sangat Baik	3073	3098	14,81%	14,93%
Baik	4761	4799	22,94%	23,13%
Buruk	2886	2912	13,91%	14,03%
Sangat Buruk	10030	9941	48,34%	47,91%
Total	20750	20750	100%	100%

Analisis terhadap nilai *baseline* hasil pengujian perbandingan durasi fonem pada durasi 3 state dapat diketahui bahwa kedua model n-gram mendapatkan nilai *baseline* yang hampir sama. Hal ini terjadi karena sedikitnya variasi durasi fonem yang digunakan, sehingga mendorong terpilihnya durasi fonem yang sama. Dapat terlihat bahwa perbedaan terhadap durasi fonem hasil rekaman penutur yang terjadi cukup besar, hal ini dikarenakan penggunaan durasi fonem dengan tiga nilai tengah, sehingga menciptakan selisih yang besar terhadap durasi fonem asli yang sangat

beragam. Selain itu penggunaan kode suku membuat HMM hanya mampu memilih durasi berdasarkan informasi tersebut. Sedangkan berdasarkan kualitas durasi yang dihasilkan, trigram lebih banyak menghasilkan durasi dengan kualitas yang dapat diterima oleh pendengar. Ini dikarenakan dengan model trigram perhitungan yang dilakukan cakupannya lebih luas, sehingga presisi prediksi yang dilakukan meningkat. Pada kedua model perhitungan, hasil yang didapat didominasi kualitas "sangat buruk". Penyebabnya terjadi hal ini sejalan dengan penyebab besarnya perbedaan hasil pengujian perbandingan durasi fonem.

Pengujian K-Fold Cross Validation Durasi Fonem 3 State

Hasil pengujian dengan penggunaan pembagian data korpus 80%-20% dapat dilihat pada Tabel 7. Sedangkan penggunaan pembagian data korpus 20%-80 dapat dilihat pada Tabel 8.

Tabel 7. Hasil Pengujian K-Fold Cross Validation 80%-20% Durasi Fonem 3 State

Fold	Korpus Uji	Bigram	Trigram
1	1-100	46,65%	49,16%
2	101-200	49,50%	50,49%
3	201-300	45,87%	46,89%
4	301-400	45,14%	45,84%
5	401-500	53,05%	53,85%
Rata-Rata		48,04%	49,25%

Tabel 8. Hasil Pengujian K-Fold Cross Validation 20%-80% Durasi Fonem 3 State

Fold	Korpus Uji	Bigram	Trigram
1	1-400	47,69%	51,72%
2	1-300, 401-500	50,38%	52,44%
3	1-200, 301-500	48,99%	51,59%
4	1-100, 201-500	48,49%	50,19%
5	101-500	48,60%	50,49%
Rata-Rata		48,83%	51,29%

Rata-rata hasil yang telah didapatkan kemudian dianalisis dengan cara dibandingkan terhadap nilai *baseline*. Pada hasil dengan penggunaan korpus latihan 80% hasil yang didapat mendekati nilai *baseline*, karena perubahan probabilitas yang terjadi di antara penggunaan korpus latihan 500 kalimat dan 400 kalimat, tidak terdapat perubahan yang signifikan.

Sedangkan hasil dengan penggunaan korpus latihan 20% diketahui terdapat sedikit perbedaan hasil dibandingkan dengan nilai *baseline*. Ini dikarenakan perbedaan yang jauh antara korpus latihan yang digunakan. Akibatnya probabilitas yang terbentuk juga berbeda jauh. Adapun perbedaan hasil yang terjadi pada tiap *fold* terjadi karena perbedaan korpus latihan dan uji yang digunakan.

Penentuan Nilai Baseline Durasi Fonem 5 State

Nilai *baseline* didapatkan dari rata-rata hasil pengujian. Hasil pengujian dapat dilihat pada Tabel 9. Hasil pengujian yang ada kemudian diolah kembali berdasarkan kualitas durasi pada Tabel 10.

Tabel 9. Hasil Nilai Baseline Durasi Fonem 5 State

Model n-gram	Nilai Baseline
Bigram	48,03%
Trigram	68,81%

Tabel 10. Hasil Nilai Baseline Durasi Fonem 5 State Berdasarkan Kualitas

Kualitas	Jumlah Fonem		Persentase Fonem	
	Bigram	Trigram	Bigram	Trigram
Sangat Baik	2239	2498	10,79%	12,04%
Baik	4746	4522	22,87%	21,79%
Buruk	2544	2344	12,26%	11,30%
Sangat Buruk	11221	11386	54,08%	54,87%
Total	20750	20750	100%	100%

Analisis terhadap nilai *baseline* hasil pengujian perbandingan durasi fonem dapat diketahui bahwa pada penggunaan durasi 5 *state*, bigram mendapatkan hasil lebih baik, yaitu sebesar 48,03% dan trigram mendapatkan hasil sebesar 68,81%. Hal ini terjadi karena dengan perhitungan trigram mendorong HMM untuk memilih durasi fonem berdasarkan tiga suku kata, sehingga durasi fonem yang terpilih merupakan durasi untuk kata dengan tiga suku kata. Padahal mayoritas kata yang ada memiliki dua suku kata. Kemudian penggunaan durasi 5 *state* ternyata belum mampu menutupi selisih besar yang terjadi dengan durasi fonem hasil rekaman.

Sedangkan berdasarkan kualitas durasi yang dihasilkan dapat diketahui bahwa penggunaan model trigram lebih banyak menghasilkan durasi yang dapat diterima oleh pendengar. Pada kedua perhitungan yang digunakan, hasil yang didapat didominasi dengan kualitas "sangat buruk". Penyebabnya terjadi hal ini sejalan dengan penyebab besarnya perbedaan hasil pengujian perbandingan durasi fonem di atas.

Pengujian K-Fold Cross Validation Durasi Fonem 5 State

Hasil pengujian dengan penggunaan pembagian data korpus 80%-20% dapat dilihat pada Tabel 11. Sedangkan penggunaan pembagian data korpus 20%-80 dapat dilihat pada Tabel 12.

Tabel 11. Hasil Pengujian K-Fold Cross Validation 80%-20% Durasi Fonem 5 State

Fold	Korpus Uji	Bigram	Trigram
1	1-100	48,44%	80,92%
2	101-200	50,60%	82,46%
3	201-300	47,65%	77,45%
4	301-400	46,30%	72,45%
5	401-500	50,84%	85,25%
Rata-Rata		48,77%	79,71%

Tabel 12. Hasil Pengujian K-Fold Cross Validation 20%-80% Durasi Fonem 5 State

Fold	Korpus Uji	Bigram	Trigram
1	1-400	53,76%	80,72%
2	1-300, 401-500	62,08%	84,21%
3	1-200, 301-500	52,28%	82,60%
4	1-100, 201-500	54,29%	81,68%
5	101-500	50,37%	80,11%
Rata-Rata		54,66%	81,86%

Rata-rata hasil yang telah didapatkan kemudian dianalisis dengan cara dibandingkan terhadap nilai *baseline*. Pada hasil dengan penggunaan korpus latihan 80% diketahui untuk model bigram hasil yang didapatkan tidak jauh berbeda dengan nilai *baseline*, sedangkan sebaliknya dengan trigram terdapat perbedaan yang cukup signifikan. Hal ini dikarenakan dengan model trigram HMM melibatkan lebih banyak kombinasi durasi dan kode suku kata, sehingga perubahan data yang terjadi memberikan dampak yang besar terhadap hasil perhitungan yang didapat. Sedangkan hasil dengan penggunaan korpus latihan 20% diketahui dibandingkan dengan nilai *baseline* baik pada penggunaan kedua

model n-gram terdapat perbedaan yang cukup signifikan. Perbedaan ini terjadi karena berkurangnya jumlah kalimat pada korpus latihan yang digunakan. Masih sama seperti sebelumnya, perbedaan hasil rata-rata perbandingan durasi fonem yang terjadi pada tiap fold terjadi karena perbedaan korpus latihan dan uji yang digunakan.

3) Pengujian Perbandingan Durasi Kalimat

Pengujian ini hampir sama dengan pengujian sebelumnya, hanya saja kali ini yang dibandingkan adalah durasi kalimat.

Penentuan Nilai Baseline Durasi Kalimat 3 State

Nilai *baseline* didapatkan dari rata-rata hasil pengujian. Hasil pengujian dapat dilihat pada Tabel 13. Hasil pengujian yang ada kemudian diolah kembali berdasarkan kualitas durasi pada Tabel 14.

Tabel 13. Hasil Nilai *Baseline* Durasi Kalimat 3 State

Model n-gram	Nilai <i>Baseline</i>
Bigram	20,62%
Trigram	18,96%

Tabel 14. Hasil Nilai *Baseline* Durasi Kalimat 3 State Berdasarkan Kualitas

Kualitas	Jumlah Fonem		Persentase Fonem	
	Bigram	Trigram	Bigram	Trigram
Sangat Baik	80	104	16,00%	20,80%
Baik	333	324	66,60%	64,80%
Buruk	74	58	14,80%	11,60%
Sangat Buruk	13	14	2,60%	2,80%
Total	500	500	100%	100%

Analisis terhadap nilai *baseline* hasil pengujian perbandingan durasi kalimat dapat diketahui bahwa pada penggunaan durasi 3 *state*, perhitungan dengan bigram mendapatkan hasil lebih baik, yaitu sebesar 20,62% dan trigram mendapatkan hasil sebesar 18,96%. Namun, perbedaan yang terjadi tidak begitu signifikan. Penyebabnya terjadi hal ini sama seperti pada pengujian perbandingan durasi fonem sedikitnya variasi durasi fonem yang digunakan. Kemudian hasil yang didapat pada pengujian ini jauh lebih baik dibandingkan dengan pengujian perbandingan durasi fonem dikarenakan durasi kalimat diperoleh dengan menjumlahkan durasi fonem, sehingga setiap nilai yang ada saling menutupi dan memperkecil selisih yang terjadi antara durasi hasil prediksi HMM dengan durasi hasil rekaman suara penutur.

Sedangkan berdasarkan kualitas durasi yang dihasilkan dapat diketahui bahwa penggunaan model trigram lebih banyak menghasilkan durasi dengan kualitas yang dapat diterima oleh pendengar. Selain itu, pengujian ini hasilnya didominasi oleh kualitas "baik" yang dikarenakan durasi kalimat berasal dari penjumlahan durasi fonem seperti yang sudah disebutkan sebelumnya.

Pengujian K-Fold Cross Validation Durasi Fonem 3 State

Hasil pengujian dengan penggunaan pembagian data korpus 80%-20% dapat dilihat pada Tabel 15. Sedangkan penggunaan pembagian data korpus 20%-80 dapat dilihat pada Tabel 16.

Tabel 15. Hasil Pengujian *K-Fold Cross Validation* 80%-20% Durasi Kalimat 3 State

Fold	Korpus Uji	Bigram	Trigram
1	1-100	18,08%	14,82%
2	101-200	20,16%	18,90%
3	201-300	20,78%	18,60%
4	301-400	23,22%	21,83%
5	401-500	20,62%	19,38%
Rata-Rata		20,57%	18,71%

Tabel 16. Hasil Pengujian *K-Fold Cross Validation* 20%-80% Durasi Fonem 3 State

Fold	Korpus Uji	Bigram	Trigram
1	1-400	18,97%	15,32%
2	1-300, 401-500	18,83%	15,51%
3	1-200, 301-500	19,64%	17,07%
4	1-100, 201-500	17,89%	16,02%
5	101-500	21,43%	18,26%
Rata-Rata		19,35%	16,44%

Rata-rata hasil yang telah didapatkan kemudian dianalisis dengan cara dibandingkan terhadap nilai *baseline*. Pada hasil dengan penggunaan korpus latihan 80% diketahui bahwa hasil yang didapatkan sangat mendekati dengan nilai *baseline*. Hal ini dikarenakan jumlah kalimat yang digunakan pada kedua korpus latihan tidak berbeda jauh. Sedangkan pada hasil dengan penggunaan korpus latihan 20% diketahui terjadi perbedaan pada kedua model n-gram. Sama seperti sebelumnya, hal ini juga dikarenakan jauhnya perbedaan jumlah kalimat yang digunakan untuk korpus latihan pada kedua pengujian. Sama seperti pada pengujian perbandingan durasi fonem, perbedaan hasil pengujian pada setiap fold terjadi karena memang berbedanya korpus latihan dan uji yang digunakan.

Penentuan Nilai Baseline Durasi Kalimat 5 State

Nilai *baseline* didapatkan dari rata-rata hasil pengujian. Hasil pengujian dapat dilihat pada Tabel 17. Hasil pengujian yang ada kemudian diolah kembali berdasarkan kualitas durasi pada Tabel 18.

Tabel 17. Hasil Nilai *Baseline* Durasi Kalimat 5 State

Model n-gram	Nilai <i>Baseline</i>
Bigram	31,98%
Trigram	16,72%

Tabel 18. Hasil Nilai *Baseline* Durasi Kalimat 5 State Berdasarkan Kualitas

Kualitas	Jumlah Fonem		Persentase Fonem	
	Bigram	Trigram	Bigram	Trigram
Sangat Baik	17	194	3,40%	38,80%
Baik	180	208	36,00%	41,60%
Buruk	191	50	38,20%	10,00%
Sangat Buruk	112	48	22,40%	9,60%
Total	500	500	100%	100%

Analisis terhadap nilai *baseline* hasil pengujian perbandingan durasi kalimat pada durasi 5 *state*, diketahui bahwa trigram mendapatkan hasil lebih baik, yaitu sebesar 16,72% dan bigram mendapatkan hasil 31,98%. Perbedaan yang terjadi di antara

kedua model perhitungan sangat signifikan. Hal ini dikarenakan perhitungan dengan trigram mendorong HMM untuk memilih durasi berdasarkan tiga suku kata, sehingga durasi yang terpilih berdurasi panjang. Akibatnya durasi fonem tersebut ketika dijumlahkan sebagai durasi kalimat lebih mampu menutupi selisih yang terjadi dengan durasi kalimat hasil rekaman suara penutur. Dari hasil yang didapat pada kedua model perhitungan juga diketahui bahwa penggunaan model trigram masih lebih banyak menghasilkan durasi dengan kualitas tergolong dapat diterima oleh pendengar.

Pengujian K-Fold Cross Validation Durasi Fonem 3 State

Hasil pengujian dengan penggunaan pembagian data korpus 80%-20% dapat dilihat pada Tabel 19. Sedangkan penggunaan pembagian data korpus 20%-80 dapat dilihat pada Tabel 20.

Tabel 19. Hasil Pengujian K-Fold Cross Validation 80%-20% Durasi Kalimat 5 State

Fold	Korpus Uji	Bigram	Trigram
1	1-100	29,02%	19,69%
2	101-200	30,60%	14,57%
3	201-300	30,86%	18,17%
4	301-400	35,41%	18,52%
5	401-500	32,18%	19,86%
Rata-Rata		31,61%	18,16%

Tabel 20. Hasil Pengujian K-Fold Cross Validation 20%-80% Durasi Fonem 5 State

Fold	Korpus Uji	Bigram	Trigram
1	1-400	23,27%	13,85%
2	1-300, 401-500	21,31%	16,09%
3	1-200, 301-500	30,24%	15,26%
4	1-100, 201-500	26,02%	15,15%
5	101-500	32,24%	14,72%
Rata-Rata		26,62%	15,01%

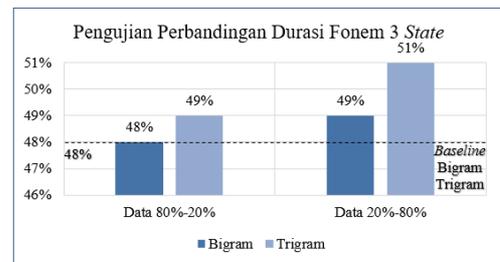
Rata-rata hasil yang telah didapatkan kemudian dianalisis dengan cara dibandingkan terhadap nilai *baseline*. Pada hasil dengan penggunaan korpus latihan 80% diketahui bahwa hasil yang didapatkan mendekati nilai *baseline*, karena perubahan probabilitas yang terjadi antara penggunaan kedua korpus latihan tidak terdapat perbedaan yang signifikan. Sedangkan pada hasil dengan penggunaan korpus latihan 20% diketahui bahwa dengan kedua model n-gram terjadi perbedaan yang signifikan. Hal ini dikarenakan jauhnya perbedaan jumlah kalimat yang digunakan untuk korpus latihan pada kedua pengujian, sehingga probabilitas yang terbentuk juga berbeda. Kemudian, sama seperti hasil-hasil pengujian sebelumnya, variasi hasil yang didapatkan pada setiap *fold* terjadi karena perbedaan pada korpus latihan dan uji yang digunakan.

4) Pembahasan Hasil Pengujian

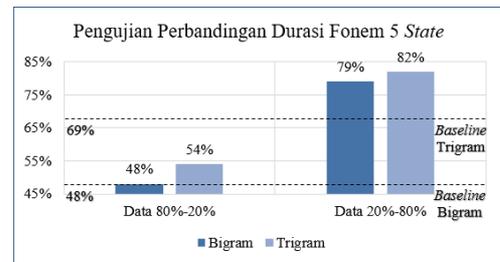
Berdasarkan nilai *baseline* yang telah diperoleh, penggunaan durasi 3 state mendapatkan hasil yang lebih baik. Hal ini dikarenakan meskipun pada penggunaan 5 state variasi durasi yang digunakan meningkat, namun ini juga meningkatkan kemungkinan terpilihnya durasi fonem dengan probabilitas kemunculan rendah. Akibatnya durasi yang terpilih oleh HMM menjauhi durasi hasil rekaman suara penutur. Selain itu, jika dilihat berdasarkan model n-gram yang digunakan, pada prediksi dengan bigram mendapatkan hasil yang lebih baik. Namun jika dilihat lebih detil terhadap kualitas durasi yang dihasilkan, trigram mendapatkan hasil yang lebih baik. Hal ini dikarenakan

penggunaan model trigram mendorong HMM untuk melakukan perhitungan yang lebih presisi. Tetapi karena sedikitnya informasi yang terkandung pada kode suku kata, presisi perhitungan ini hanya terjadi pada kata yang dengan tiga suku kata. Akibatnya secara keseluruhan nilai *baseline* yang didapat menjadi lebih buruk. Sehingga dapat disimpulkan berdasarkan model n-gram, trigram dapat memprediksi durasi fonem lebih baik.

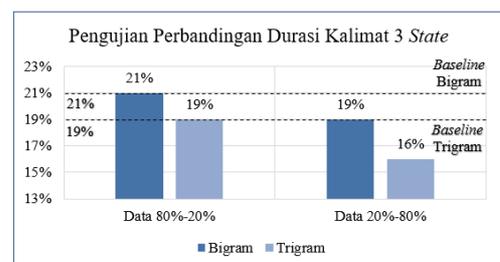
Berdasarkan pengujian dengan *k-fold cross validation*, untuk pembagian data 80%-20% secara keseluruhan hasil yang didapat tidak jauh berbeda dengan nilai *baseline*. Ini dikarenakan penggunaan 500 kalimat pada korpus latihan untuk pengujian nilai Baseline dengan 400 kalimat pada korpus latihan untuk pengujian ini, menghasilkan data probabilitas yang tidak jauh berbeda. Sedangkan untuk pembagian data 20%-80% secara keseluruhan hasil yang didapat cukup berbeda dengan nilai *baseline*. Ini dikarenakan perbedaan yang jauh dalam jumlah kalimat yang digunakan pada korpus latihan. Seperti yang telah disebutkan sebelumnya, korpus latihan pada pembagian data 20%-80% berisi 100 kalimat saja. Akibatnya data probabilitas yang terbentuk jauh berbeda. Sedangkan perbedaan yang terjadi pada tiap *fold* untuk kedua pembagian data, dikarenakan perbedaan kalimat untuk setiap *fold*. Perbandingan hasil setiap pengujian *k-fold cross validation* terhadap nilai *baseline* dapat terlihat pada sejumlah grafik di bawah ini.



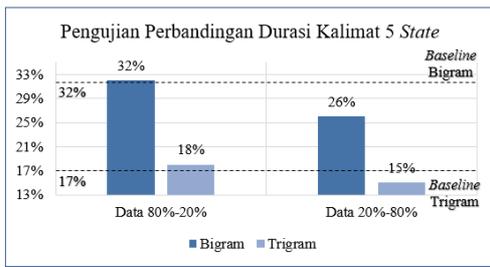
Gambar 7. Grafik hasil pengujian perbandingan durasi fonem 3 state



Gambar 8. Grafik hasil pengujian perbandingan durasi fonem 5 state



Gambar 9. Grafik hasil pengujian perbandingan durasi kalimat 3 state



Gambar 10. Grafik hasil pengujian perbandingan durasi kalimat 5 state

KESIMPULAN

Berdasarkan hasil analisis dan pengujian terhadap prediksi durasi fonem pada bahasa Melayu Pontianak berdasarkan suku kata menggunakan Hidden Markov Model (HMM), maka dapat ditarik kesimpulan sebagai berikut:

1. Berdasarkan pengujian perbandingan durasi fonem untuk penentuan nilai *baseline*, hasil durasi fonem yang didapatkan mayoritas berkualitas "sangat buruk", dengan prediksi pada durasi 3 *state* dengan model bigram mendapatkan hasil terbaik.
2. Berdasarkan pengujian perbandingan durasi kalimat untuk penentuan nilai *baseline*, hasil durasi yang didapatkan mayoritas berkualitas "baik", dengan prediksi pada durasi 5 *state* dengan model trigram mendapatkan hasil terbaik.
3. Berdasarkan state durasi yang digunakan, penggunaan 3 *state* mendapatkan hasil yang lebih baik, sedangkan berdasarkan model *n*-gram, penggunaan trigram mendapatkan hasil yang lebih baik.
4. Berdasarkan keseluruhan nilai *baseline* yang didapat, prediksi durasi fonem pada bahasa Melayu Pontianak berdasarkan suku kata dengan penggunaan durasi hasil pengelompokkan 3 dan 5 *state* masih belum mampu menghasilkan durasi fonem yang tepat.
5. Berdasarkan pengujian *k-fold cross validation* yang didapat, pemodelan HMM dalam memprediksi durasi fonem berdasarkan suku kata sudah cukup mampu dalam mencapai nilai *baseline* yang diharapkan.

UCAPAN TERIMA KASIH

Terima kasih kepada Allah SWT dan junjungan Nabi Muhammad SAW. Terima kasih juga kepada keluarga dan teman-teman yang selalu mendukung saya. Terakhir terima kasih kepada berbagai pihak lainnya yang tidak dapat saya sebutkan satu persatu.

DAFTAR PUSTAKA

- [1] Krishna, N. S., Talukdar, P. P., Bali, K., dan Ramakrishnan, A. G. 2004. Duration Modeling for Hindi Text-To-Speech Synthesis System. Bangalore: Indian Institute of Science.
- [2] Na'im, Akhsan dan Hendry S. 2010. Kewarganegaraan, Suku Bangsa, Agama, dan Bahasa Sehari-hari Penduduk Indonesia, Hasil Sensus Penduduk 2010. Jakarta: Badan Pusat Statistik.
- [3] Novianti, E. 2011. Menilik Nasib Bahasa Melayu Pontianak. Semarang: Universitas Diponegoro.
- [4] Benesty, J., M. Sondhi, M., dan Yiteng, H. 2008. Springer Handbook of Speech Processing. Berlin: Springer-Verlag.

- [5] Nugraha, T. A. 2014. Prediksi Jeda dalam Ucapan Kalimat Bahasa Indonesia dengan Hidden Markov Model. Pontianak: Program Studi Informatika, Fakultas Teknik, Universitas Tanjungpura.
- [6] Prasetyo Budi, M. E. 2011. Teori Dasar Hidden Markov Model. Bandung: Program Studi Sistem dan Teknologi Informasi, Institut Teknologi Bandung.
- [7] Alwi, H., Dardjowidjojo, S., Lapoliwa, H., dan M. Moeliono, A. 2000. Tata bahasa baku bahasa Indonesia (Edisi Ketiga). Jakarta: Balai Pustaka
- [8] Erick, P. D. P. dan Hartati, U. 2016. Perbedaan Bahasa Melayu Pontianak Kalimantan Barat dengan Bahasa Indonesia Standar. Yogyakarta: FKIP Universitas Sarjanwiyata Tamansiswa Yogyakarta.
- [9] Ratnawati, N. K. M. dan Anak, A. N. B. J. D. 2015. Fonem Suprasegmental / Ciri-ciri Prosodi. Singaraja: Jurusan Pendidikan Bahasa dan Sanstra Indonesia, Fakultas Bahasa dan Senis, Universitas Pendidikan Ganesha.
- [10] Herho, S. H. S. 2017. Tutorial Pemrograman Python 2 Untuk Pemula. Bandung: WPCL Press.
- [11] Muniyati, E. F. 2019. Prediksi Jeda dalam Ucapan Kalimat Bahasa Melayu Pontianak Menggunakan Hidden Markov Model Berbasis Part of Speech. Pontianak: Jurusan Informatika, Fakultas Teknik, Universitas Tanjungpura.
- [12] Wicaksono, A. F. dan Ayu, P. 2010. HMM Based Part-Of-Speech Tagger for Bahasa Indonesia. Bandung: Teknik Informatika, Institut Teknologi Bandung.
- [13] Numrotul, H. 2019. Analisis Implementasi Model Durasi Klatt pada Bahasa Melayu Pontianak Menggunakan Diphone Bahasa Indonesia. Pontianak: Jurusan Informatika, Fakultas Teknik, Universitas Tanjungpura.