

Optimalisasi Metode LPC-16 dan HMM-Forward Pada Sistem Asisten Virtual

Chondro Seto Nur Suryawan, Marisa Premitasari
Institut Teknologi Nasional

Jl. PH.H. Mustofa No.23 Bandung 40124, Bandung
chondroseto@windowslive.com, marisa.premitasari@gmail

Abstrak

Pada umumnya manusia saat ini menggunakan sistem operasi *windows* yang berjalan di perangkat desktop akan memasang banyak aplikasi sesuai kebutuhannya. Semakin banyak aplikasi yang di pasang maka semakin banyak pula *shortcut* yang tampil di bagian *desktop windows*. *Shortcut* sendiri merupakan sebuah objek alternatif yang digunakan untuk mewakili sehingga pengguna dapat dengan mudah membuka aplikasi tanpa harus pengguna membuka tempat dimana aplikasi tersebut terpasang. Banyaknya aplikasi yang terpasang pada sistem operasi *windows* membuat *shortcut* pada bagian *desktop* menjadi banyak dan membuat pengguna kesulitan dalam mencari atau membuka aplikasi yang diinginkan. Oleh karena itu diperlukan aplikasi yang dapat membantu pengguna dalam mencari dan membuka aplikasi dengan mudah tanpa membuat pengguna kesulitan. Aplikasi tersebut adalah virtual asisten yang akan membantu pengguna dalam mencari dan membuka aplikasi yang diinginkan. Cara kerjanya dengan pengguna memasukan suara pengguna lalu di proses ekstraksi ciri menggunakan metode Linear Predictive Coding lalu di klasifikasikan menggunakan metode Hidden Markov Model Forward. Setelah terdeteksi maka aplikasi akan membuka aplikasi sesuai suara yang terdeteksi. Penelitian ini menggunakan 120 data latih yang terdiri dari 6 label yaitu whatsapp, linkedin, Tokopedia, gmail, powerpoint, word. Untuk setiap label memiliki data latih berjumlah 20 data. Data yang diujikan berjumlah 60. Untuk setiap labelnya memiliki 10 data uji.

Kata kunci :

Pengenalan Suara, Linear Predictive Coding, Hidden Markov Model

Abstract

In general, people today using windows operating system that runs on desktop devices will install many applications as needed. The more applications installed, the more shortcuts that appear in the desktop part of windows. Shortcut itself is an alternative object used to represent so that the user can easily open the application. Users can open without having to open the place where the application is installed. The large number of applications installed on the windows operating system makes shortcuts on the desktop become many and makes it difficult for users to find or open cool applications. Therefore, an application is needed that can assist the user in finding and opening the application easily without making it difficult for the user. The app is a virtual assistant that will assist users in finding and opening the desired application. How it works with the user entering the user's voice and then in the process of extracting the characteristics using the Linear Predictive Coding method and then classified using the Hidden Markov Model Forward method. Once detected, the application will open the application according to the detected sound. This study used 120 training data that were divided into 6 labels consisting of Whatsapp, Linkedin, Tokopedia, Gmail, PowerPoint, and Word. For each label has a training data of 20 data. The data tested amounted to 60. Each label has 10 test data.

Keywords :

Speech Recognition, Linear Predictive Coding, Hidden Markov Model

I. PENDAHULUAN

Pada Umumnya manusia saat ini menggunakan sistem operasi windows yang dapat berjalan di perangkat dekstop akan memasang banyak aplikasi karena banyaknya kebutuhan sehingga pada bagian dekstop terdapat banyak aplikasi yang terlihat Umumnya pengguna memasang banyak aplikasi di sistem operasi windowsnya karena digunakan untuk bekerja dan keseharian serta hiburan. pengguna dapat membuka aplikasi dengan shortcut aplikasi pada bagian desktop di sistem operasi windows. Dalam mencari shorcut aplikasi membutuhkan waktu yang tidak sedikit terkadang pengguna keliru dan salah klik aplikasi sehingga membuat pusing pengguna. Dilihat dari proses pencarian aplikasi yang cukup panjang tersebut maka diperlukan suatu solusi yang dapat menyederhanakan suatu proses tersebut untuk membuka aplikasi tanpa harus mencari menggunakan penglihatan pengguna. Solusi yang diperlukan adalah sebuah aplikasi yang dapat membuka hanya dengan menyebutkan nama aplikasinya saja menggunakan suara pengguna tanpa perlu mencari aplikasi dengan penglihatan pengguna

Aplikasi ini memerlukan masukan pengguna berupa suara pengguna yang menyebutkan nama aplikasi tersebut lalu di proses dengan mengubahnya menjadi angka digital. Angka digital tersebut lalu di klasifikasikan sesuai dengan data yang sudah di latih sebelumnya. Setelah terdeteksi maka akan tampil teks atau nama aplikasi yang hasil proses pengenalan suara tersebut. Setelah itu sistem akan mencari sesuai hasil pengenalan suara tersebut dan mencocokkannya dengan aplikasi yang sesuai yang sudah tersedia pada aplikasi tersebut. Jika hasil klasifikasi terdapat kecocokan dengan nama aplikasi yang sudah terdaftar maka aplikasi akan dibuka oleh aplikasi dan di jalan kan oleh sistem operasi. Nama aplikasi ini biasanya disebut virtual asisten

Virtual asisten adalah sebuah agen yang cerdas yang memiliki tujuan untuk membantu pengguna dalam melakukan suatu tugas. Umumnya virtual asisten menggunakan masukan suara dari pengguna untuk melakukan suatu tugas. Dalam mengenali pengguna dan melakukan tugas diperlukan pengenalan terlebih dahulu dengan suara pengguna.

Pengenalan suara adalah suatu proses dalam mengenali suatu suara. Dalam mengenali suara, suara harus di ekstraksi terlebih dahulu . dalam penelitian ini penulis menggunakan metode Linear Predictive

Coding untuk mengekstraksi suara. Bukan hanya ekstraksi ciri tapi juga diperlukan metode lain untuk mengenali hasil dari ekstraksi ciri. Dalam penelitian ini juga menggunakan metode Hidden Markov Model.

Linear Predictive Coding merupakan metode ekstraksi dari sinyal suara menjadi bentuk nilai-nilai berupa koefisien Linear Predictive Coding. Biasanya metode ini digunakan untuk *speech recognition*, *speech to text* dan *text to speech*.

Hidden Markov Model merupakan sebuah model statistik dari suatu sistem yang dapat menghitung probabilitas dari suatu kejadian yang tidak dapat observasi berdasarkan kejadian yang dapat observasi. Hidden Markov Model terdiri dari 4 Algoritma yang terdiri dari *Forward*, *Backward*, *Viterbi* dan *Baum Welch*. Algoritma yang digunakan adalah algoritma *forward*. Algoritma ini akan digunakan pada saat proses pencocokan dengan data latih.

Penelitian mengenai pengenalan suara sudah dilakukan oleh (Abdullah, & Ramadhan, 2016) yang berjudul Implementasi Algoritma Hidden Markov Model Sebagai Pengenalan Perintah Suara Pada Aplikasi Winamp yang meneliti tentang pengenalan perintah suara menggunakan metode linear predictive coding dan hidden markov model, untuk parameter linear predictive coding yang digunakan dalam penelitian ini tidak disebutkan dan algoritma yang digunakan pada hidden markov model juga tidak disebutkan.

Oleh karena itu penelitian penulis akan meneliti metode linear predictive coding dengan ordo 16 dan hidden markov model dengan algoritma forward dalam mengenali suara.diharapkan dengan metode linear predictive coding dengan ordo16 akurasi dapat mencapai angka 80%.

II. LANDASAN TEORI

II.1 Virtual Asisten

Virtual Asisten adalah suatu agen cerdas dimana agen ini memiliki tujuan untuk membantu melakukan tugas dengan menyederhanakan suatu kegiatan dengan perintah suara (Massai, Nesi, & Pantaleo, 2019).

Teknologi ini berfungsi untuk memudahkan pengguna dalam melakukan suatu hal yang biasanya di lakukan oleh pengguna, dimana pengguna hanya perlu memanggil apa yang diinginkan lalu virtual asisten akan melakukan apa yang diinginkan

pengguna atau mewakili pengguna dalam melakukan hal tersebut

Virtual asisten ini juga dikenal dengan *virtual personal assistant* (VPA). Contoh virtual asisten yaitu *Google Assistant*, *Microsoft Cortana*, *Apple Siri* dan *Amazon Alexa*.

II.2 Pengenalan Suara

Pengenalan Suara (*Speech Recognition*) adalah suatu proses dalam mengenali kata yang diucapkan oleh komputer (Abdullah, & Ramadhan, 2016). Input yang diperlukan dari pengenalan suara ialah berupa suara atau ucapan yang diubah menjadi sinyal digital dengan cara mengubah gelombang suara menjadi sekumpulan angka lalu disesuaikan dengan kode dan pola yang tersimpan di dalam suatu perangkat. Setiap ucapan memiliki pola ciri yang berbeda

Terdapat dua tipe pengenalan suara (*speech recognition*) berdasarkan ketergantungan pembicara yaitu *independent speech recognition* dan *dependent speech recognition*.

Independent speech recognition adalah sistem pengenalan ucapan yang tidak dipengaruhi oleh pembicara, tetapi mempunyai keterbatasan dalam jumlah kata. Contohnya asisten virtual *Microsoft Cortana*, *Apple Siri* dan *Google Assistant*

Dependent speech recognition adalah sistem pengenalan ucapan yang dipengaruhi oleh pembicara. Sistem ini memerlukan pelatihan khusus dari pembicara dimana Hasil pelatihan dari setiap pembicara akan di simpan pada kelas. kelas inilah yang akan digunakan pada sistem untuk mengenali pembicara untuk berinteraksi dengan sistem. Contohnya pada sistem untuk membuka kunci dengan suara

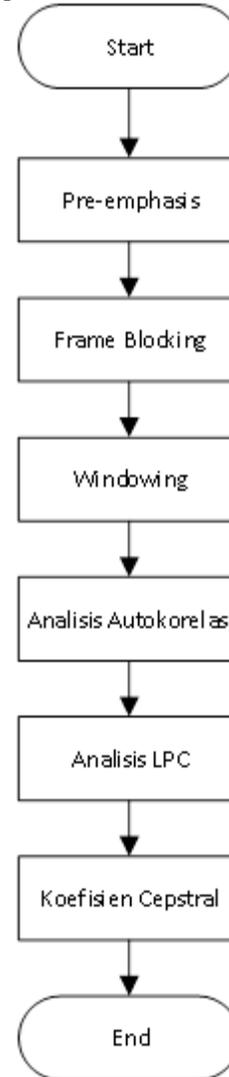
II.3 Linear Predictive Coding

LPC (*Linear Predictive Coding*) adalah suatu metode ekstraksi ciri dari sinyal suara yang dapat mengubah ciri dari suatu sinyal menjadi ke dalam bentuk nilai berupa koefisien *Linear Predictive Coding* atau LPC (Abdullah, & Ramadhan, 2016). Metode *Linear Predictive Coding* lebih sering digunakan pada aplikasi yang *speech recognition*, *speech to text* dan *text to speech*. Proses *linear predictive coding* membutuhkan beberapa parameter yaitu :

N = jumlah sample per frame

M = jarak antara frame yang berurutan
P = ordo LPC

Dalam proses ekstraksi ciri menggunakan metode *Linear Predictive Coding* terdapat 6 tahapan yang ditunjukkan pada Gambar 2.1 berikut ini :



Gambar 1 Flowchart Metode LPC

II.4 Hidden Markov Model

HMM (*Hidden Markov Model*) merupakan sebuah model statistik dari suatu sistem yang dapat melakukan perhitungan probabilitas dari suatu kejadian yang tidak dapat di observasi berdasarkan kejadian yang dapat di observasi (Widhiyanti, 2012).

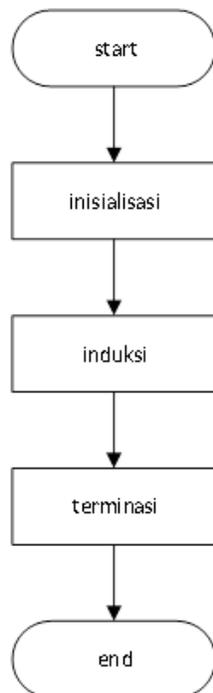
Hidden Markov Model memiliki 2 macam bagian yaitu observed state dan hidden state. Observed state adalah bagian yang dapat diamati sedangkan hidden state adalah bagian yang tidak dapat diamati (Widhiyantil, 2012)

Metode hidden markov model dapat digunakan untuk aplikasi di bidang pattern recognition contohnya pengenalan suara, pengenalan tulisan, pengenalan gestur, bioinformatika, kompresi kalimat, computer vision, dan lain-lain

Hidden Markov Model memiliki 4 algoritma yaitu Forward, Backward, Viterbi dan Baum Welch. Algoritma yang digunakan pada penelitian ini adalah algoritma forward yang akan digunakan saat proses pencocokan dengan data hasil training.

Hidden Markov Model menggabungkan dua atau lebih rantai markov dimana satu rantai yang terdiri dari state yang dapat diobservasi dan rantai lainnya membentuk state yang tidak dapat diobservasi, yang akan mempengaruhi hasil state yang dapat diobservasi.

Pada penelitian ini digunakan algoritma yang digunakan adalah algoritma forward yang terdiri dari 3 tahapan yang ditunjukkan pada Gambar 2 berikut.

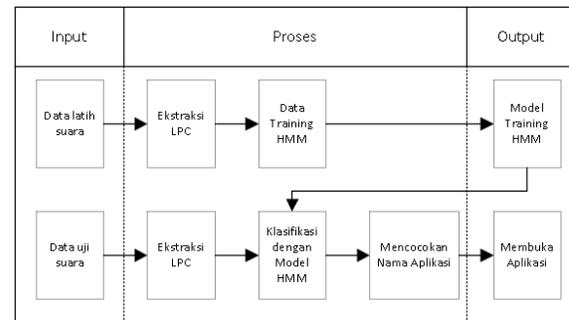


Gambar 2 Flowchart Metode HMM

III. ANALISIS DAN PERANCANGAN

III.1 Blok Diagram

Dibawah ini adalah blok diagram sistem ditunjukkan pada Gambar 3 berikut ini.



Gambar 3 Blok Diagram Sistem

Berdasarkan gambar blok diagram diatas ini, terdapat dua proses yaitu proses data latih suara dan data suara uji. Untuk proses data latih suara, data latih suara akan di ekstraksi dengan metode Linear Predictive Coding lalu di latih menjadi data training. Sedangkan untuk data suara uji akan di ekstraksi Linear Predictive Coding lalu hasil data training akan di masukan ke klasifikasi Hidden Markov Model setelah itu akan memulai proses mencocokkan nama aplikasi. Hasil outputnya akan tampil teks hasil input suara dan membuka aplikasi yang dituju.

III.2 Pengambilan Data

Data latih suara dan data uji suara menggunakan suara penulis. Format file data suara yang digunakan untuk data latih suara dan data uji suara adalah format data WAV(Waveform Audio File). Sample rate yang digunakan untuk data latih suara dan data uji suara adalah 8000 sample rate. data latih suara dan data uji suara berdurasi dua detik lebih dan tidak lebih dari tiga detik.

Data latih suara diperoleh dari suara penulis dengan jumlah total 120 data suara yang terdiri dari enam label. Masing-masing label berisi 20 data suara. Data uji suara diperoleh dari suara penulis dengan jumlah total 60 data suara yang terdiri dari enam label. Masing-masing label berisi 10 data suara. Label terdiri dari Gmail, Whatsapp, Tokopedia, LinkedIn, Word dan Powerpoint. Setiap label memiliki data latih suara dan data uji suara.

III.3 Hasil Pengujian

Dibawah ini hasil pengujian data yang ditunjukkan pada Tabel 4.7 berikut ini

Tabel 1 Hasil Pengujian Data

File Suara	Label Sebenarnya	Label Hasil Prediksi
Whatsappuji1.wav	Whatsapp	Whatsapp
Whatsappuji2.wav	Whatsapp	Whatsapp
Whatsappuji3.wav	Whatsapp	Whatsapp
Whatsappuji4.wav	Whatsapp	Whatsapp
Whatsappuji5.wav	Whatsapp	Whatsapp
Whatsappuji6.wav	Whatsapp	Tokopedia
Whatsappuji7.wav	Whatsapp	Whatsapp
Whatsappuji8.wav	Whatsapp	Whatsapp
Whatsappuji9.wav	Whatsapp	Whatsapp
Whatsappuji10.wav	Whatsapp	Tokopedia
Linkedinuji1.wav	Linkedin	Linkedin
Linkedinuji2.wav	Linkedin	Linkedin
Linkedinuji3wav	Linkedin	Linkedin
Linkedinuji4.wav	Linkedin	Linkedin
Linkedinuji5.wav	Linkedin	Linkedin
Linkedinuji6.wav	Linkedin	Linkedin
Linkedinuji7.wav	Linkedin	Linkedin
Linkedinuji8.wav	Linkedin	Linkedin
Linkedinuji9.wav	Linkedin	Tokopedia

Linkedinuji10.wav	Linkedin	Linkedin
Tokopediauji1.wav	Tokopedia	Tokopedia
Tokopediauji2.wav	Tokopedia	Tokopedia
Tokopediauji3.wav	Tokopedia	Tokopedia
Tokopediauji4.wav	Tokopedia	Tokopedia
Tokopediauji5.wav	Tokopedia	Tokopedia
Tokopediauji6.wav	Tokopedia	Tokopedia
Tokopediauji7.wav	Tokopedia	Tokopedia
Tokopediauji8.wav	Tokopedia	Whatsapp
Tokopediauji9.wav	Tokopedia	Tokopedia
Tokopediauji10.wav	Tokopedia	Whatsapp
Gmailuji1.wav	Gmail	Tokopedia
Gmailuji2.wav	Gmail	Gmail
Gmailuji3.wav	Gmail	Gmail
Gmailuji4.wav	Gmail	Tokopedia
Gmailuji5.wav	Gmail	Gmail
Gmailuji6.wav	Gmail	Gmail
Gmailuji7.wav	Gmail	Gmail
Gmailuji8.wav	Gmail	Gmail
Gmailuji9.wav	Gmail	Gmail
Gmailuji10.wav	Gmail	Gmail
Powerpointuji1.wav	Powerpoint	Powerpoint
Powerpointuji2.wav	Powerpoint	Powerpoint
Powerpointuji3.wav	Powerpoint	Powerpoint

Powerpointuji4.wav	Powerpoint	Powerpoint
Powerpointuji5.wav	Powerpoint	Powerpoint
Powerpointuji6.wav	Powerpoint	Powerpoint
Powerpointuji7.wav	Powerpoint	Powerpoint
Powerpointuji8.wav	Powerpoint	Powerpoint
Powerpointuji9.wav	Powerpoint	Word
Powerpointuji10.wav	Powerpoint	Powerpoint
Worduji1.wav	Word	Word
Worduji2.wav	Word	Word
Worduji3.wav	Word	Word
Worduji4.wav	Word	Word
Worduji5.wav	Word	Word
Worduji6.wav	Word	Word
Worduji7.wav	Word	Word
Worduji8.wav	Word	Tokopedia
Worduji9.wav	Word	Word
Worduji10.wav	Word	Word

Dibawah ini adalah tabel jumlah yang terdeteksi berdasarkan Tabel 4.7 yang ditunjukkan pada Tabel 4.8 berikut ini

Tabel 4.8 Jumlah yang terdeteksi

No.	Nama Aplikasi	Terdeteksi	Tidak Terdeteksi	Persentase
1	Whatsapp	8	2	80%
2	Linkedin	9	1	90%

3	Tokopedia	8	2	80%
4	Gmail	8	2	80%
5	Powerpoint	9	1	90%
6	Word	9	1	90%

Berdasarkan tabel di atas whatsapp memiliki persentase 80% yang terdeteksi dari 10 suara uji. Linkedin memiliki persentase 90% yang terdeteksi dari 10 suara uji. Tokopedia memiliki persentase 80% yang terdeteksi dari 10 suara uji. Gmail memiliki persentase 80% yang terdeteksi dari 10 suara uji. Powerpoint memiliki persentase 90% yang terdeteksi dari 10 suara uji. Word memiliki persentase 90% yang terdeteksi dari 10 suara uji.

Tabel 4.9 Hasil Pengujian Berdasarkan Ordo LPC

No	Nama Aplikasi	Ordo LPC				
		16	15	14	13	12
1	Whatsapp	80%	80%	60%	50%	20%
2	Linkedin	90%	90%	80%	80%	50%
3	Tokopedia	80%	80%	60%	50%	70%
4	Gmail	80%	60%	50%	70%	50%
5	Powerpoint	90%	90%	90%	100%	100%
6	Word	90%	80%	80%	70%	50%
Rata-Rata		85%	80%	70%	70%	57%

Berdasarkan data tabel hasil pengujian diatas didapatkan hasil berupa nilai rata-rata persentase pengujian deteksi. pengujian dengan ordo 16 memiliki rata-rata persentase 85%. pengujian dengan ordo 15 memiliki rata-rata persentase 80%. pengujian dengan ordo 14 memiliki rata-rata persentase 70%. pengujian dengan ordo 13 memiliki rata-rata persentase 70%. pengujian dengan ordo 12 memiliki rata-rata

persentase 57%. Pengujian dengan ordo 16 memiliki persentase terbesar daripada pengujian dengan ordo yang diujikan.

IV. KESIMPULAN DAN SARAN

Berdasarkan hasil pengujian data suara dari 6 nama aplikasi yang terdiri dari Whatsapp, LinkedIn, Tokopedia, Gmail, Powerpoint, Word yang menggunakan metode Linear Predictive Coding dengan ordo 16 dan menggunakan metode Hidden Markov Model Forward menghasilkan persentase rata-rata dalam mendeteksi data suara adalah 85% serta memiliki nilai rata-rata tertinggi diantara pengujian ordo Linear Predictive Coding 15, 14,13, dan 12. Sehingga ordo 16 adalah ordo Linear Predictive Coding yang paling optimal berdasarkan rata-rata deteksi tertinggi.

Keakurasian deteksi ini juga di pengaruhi oleh banyak faktor yang bisa menyebabkan keberagaman hasil akurasi diantaranya data latih, data uji dan nilai ordo yang digunakan pada metode Linear predictive coding serta Algoritma Hidden Markov Model

REFERENSI

- Champion, C., & Houghton, S. M. (2015). Application of continuous state Hidden Markov Models to a classical problem in speech recognition. *Computer Speech & Language*, 1-18. <https://doi.org/10.1016/j.csl.2015.05.001>
- DEWI, I. A., ZULKARNAIN, A., & LESTARI, A. A. (2018). Identifikasi Suara Tangisan Bayi menggunakan Metode LPC dan Euclidean Distance. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, 6(1), 153. <https://doi.org/10.26760/elkomika.v6i1.153>
- Khairunizam, Danuri, & Jaroji. (2017). *Aplikasi Pemutar Musik Menggunakan Speech Recognition*. 2(2).
- Khandade, S., & Khot, S. (2016). *Speaker Recognition Based Home Automation Using Matlab*. (9), 1627-1631.
- Massai, L., Nesi, P., & Pantaleo, G. (2019). PAVAL: A location-aware virtual personal assistant for retrieving geolocated points of interest and location-based services. *Engineering Applications of Artificial Intelligence*, 77(September 2018), 70-85. <https://doi.org/10.1016/j.engappai.2018.09.013>
- Mouaza, B., Abderrahim, B. H., & Abdelmajid, E. (2019). ScienceDirect ScienceDirect Speech Recognition of Dialect Using Hidden Markov 29 Moroccan Models Speech Recognition of Moroccan Dialect Using Hidden Markov. *Procedia Computer Science*, 151(2018), 985-991. <https://doi.org/10.1016/j.procs.2019.04.138>
- Nada, Q., Ridhuandi, C., Santoso, P., & Apriyanto, D. (2019). *Speech Recognition dengan Hidden Markov Model untuk*. 5(1), 19-26.
- Orosanu, L., & Jouviet, D. (2018). Detection of sentence modality on French automatic speech-to-text transcriptions. *Procedia Computer Science*, 128, 38-46. <https://doi.org/10.1016/j.procs.2018.03.006>
- Pseudocode, J., Abdullah, D., & Ramadhan, R. (2016). *IMPLEMENTASI ALGORITMA HIDDEN MARKOV MODEL SEBAGAI PENGENALAN*. III, 15-25.
- Shadiev, R., & Huang, Y. M. (2016). Facilitating cross-cultural understanding with learning activities supported by speech-to-text recognition and computer-aided translation. *Computers and Education*, 98, 130-141. <https://doi.org/10.1016/j.compedu.2016.03.013>
- Sharma, S., Kumar, V., & Rana, K. P. S. (2019). Automatic oscillations detection and quantification in process control loops using linear predictive coding. *Engineering Science and Technology, an International Journal*, (xxxx). <https://doi.org/10.1016/j.jestch.2019.04.006>
- Widhiyantil, K. (2012). *Dengan HMM dan Rule Based*. 8(2).