

**OPTIMASI POWER NORMALIZED CEPSTRAL  
COEFFICIENTS DAN GROWING SELF ORGANIZING  
MAPS DALAM PENGENALAN SUARA**



**OLEH :**  
**MEUTIA PUSPITASARI, S.KOM**  
**09042681721007**

**PROGRAM STUDI MAGISTER ILMU KOMPUTER  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS SRIWIJAYA  
TAHUN 2021**

**OPTIMASI POWER NORMALIZED CEPSTRAL  
COEFFICIENTS DAN GROWING SELF ORGANIZING  
MAPS DALAM PENGENALAN SUARA**

**TESIS**

**Diajukan Untuk Melengkapi Salah Satu Syarat**

**Memperoleh Gelar Magister**



**OLEH :**

**MEUTIA PUSPITASARI, S.KOM**

**09042681721007**

**PROGRAM STUDI MAGISTER ILMU KOMPUTER  
FAKULTAS ILMU KOMPUTER  
UNIVERSITAS SRIWIJAYA  
TAHUN 2021**

## LEMBAR PENGESAHAN

### **OPTIMASI POWER NORMALIZED CEPSTRAL COEFFICIENTS DAN GROWING SELF ORGANIZING MAPS DALAM PENGENALAN SUARA**

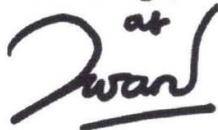
#### TESIS

Diajukan Untuk Melengkapi Salah Satu Syarat  
Memperoleh Gelar Magister

**OLEH :**  
**MEUTIA PUSPITASARI**  
**09042681721007**

Palembang, September 2021

Pembimbing I,



Dr. Iwan Pahendra, M.T.  
NIP. 197403222002121002

Pembimbing II,



Dr. Abdiansah, S.Kom., M.Cs.  
NIP. 198410012009121005

Mengetahui,  
Koordinator Program Studi Magister Ilmu Komputer



Dian Palupi Rini, M.Kom., Ph.D.  
NIP 197802232006042002

## LEMBAR PERSETUJUAN

Pada hari Selasa tanggal 13 Juli 2021 telah dilaksanakan ujian sidang Tesis secara daring oleh Magister Ilmu Komputer Fakultas Ilmu Komputer Universitas Sriwijaya.

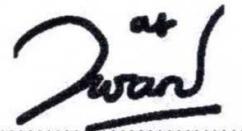
Nama : Meutia Puspitasari

NIM : 09042681721007

Judul : Optimasi *Power Normalized Cepstral Coefficients* dan  
*Growing Self Organizing Maps* dalam Pengenalan Suara

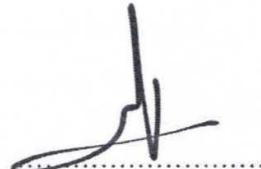
### 1. Pembimbing I

Dr. Iwan Pahendra, M.T.  
NIP. 197403222002121002



### 2. Pembimbing II

Dr. Abdiansah, S.Kom., M.Cs.  
NIP. 198410012009121005



### 3. Penguji I

Dr. Erwin, S.Si., M.Si.  
NIP. 197101291994121001



### 4. Penguji II

Dr. Ermatita, M.Kom.  
NIP.196709132006042001



Mengetahui,

Koordinator Program Studi Magister Ilmu Komputer

Dian Palupi Rini, M.Kom., Ph.D  
NIP 197802232006042002

## LEMBAR PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Meutia Puspitasari  
NIM : 09042681721007  
Program Studi : Magister Ilmu Komputer  
Judul Tesis : Optimasi *Power Normalized Cepstral Coefficients* dan  
*Growing Self Organizing Maps* dalam Pengenalan  
Suara

Hasil Pengecekan Software iThenticate/Turnitin : **13%**

Menyatakan bahwa laporan tesis saya merupakan hasil karya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam laporan tesis ini, maka saya bersedia menerima sanksi akademik dari Universitas Sriwijaya sesuai dengan ketentuan yang berlaku.

Demikian, pernyataan ini saya buat dengan sebenarnya dan tidak ada paksaan oleh siapapun.

Palembang, 23 September 2021



Meutia Puspitasari  
NIM. 09042681721007

## KATA PENGANTAR

Puji dan syukur penulis haturkan kehadiran Allah Swt yang telah melimpahkan rahmat, hidayah, dan kesehatan kepada penulis sehingga dapat menyelesaikan Tesis yang berjudul "**Optimasi Power Normalized Cepstral Coefficients dan Growing Self Organizing Maps dalam Pengenalan Suara**". Pada kesempatan ini, penulis juga hendak mengucapkan banyak terima kasih kepada pihak-pihak yang telah banyak membantu menyelesaikan Tesis ini diantaranya :

1. Suami, M Nanda Seftian, SE., M.Si yang telah memberikan dukungan, do'a serta motivasi untuk menyelesaikan tesis.
2. Orang tua, Bapak Erwin Usman dan Ibu Reni Rosarina yang telah memberikan do'a dan dukungan.
3. Mertua, Ibu Nuryani yang telah memberikan dukungan dan semangat serta do'a untuk menyelesaikan tesis.
4. Anak-anak, Alysha Azzahra Medina dan Muhammad Faezya Al Fatih yang selalu menjadi penghibur hati dan memberikan saya semangat dalam menyelesaikan tesis.
5. Dian Palupi Rini, S.Si, M.Kom, Ph.D. selaku Koordinator Program Studi Magister Ilmu Komputer Universitas Sriwijaya dan bimbingan Tesis atas kebijakkan dan dukungannya selama pelaksanaan Tesis.
6. Alm. Reza Firsandaya Malik, S.T. M.T., Ph.D selaku dosen Pembimbing I tesis saya yang memberikan arahan dan masukan terhadap tesis saya.
7. Dr. Iwan Pahendra, M.T selaku dosen Pembimbing I tesis saya yang telah memberikan masukan dan koreksi dalam penulisan tesis.
8. Dr. Abdiansah, S.Kom., M.Cs. selaku dosen Pembimbing II tesis saya yang telah memberikan arahan dan bimbingan dalam penulisan tesis.
9. Dr. Erwin, S.Si., M.Si. dan Dr. Ermatita, M.Kom selaku dosen Pengaji I dan II yang telah memberikan masukan untuk penulisan tesis
10. Semua dosen Program Studi Magister Ilmu Komputer Universitas Sriwijaya yang telah melimpahkan ilmunya kepada penulis selama proses belajar mengajar di Fakultas Ilmu Komputer Universitas Sriwijaya.

11. Ardina Ariani, M.Kom selaku admin Program Studi Magister Ilmu Komputer yang telah banyak memperlancar kegiatan akademik dan siding tesis.
12. Semua pihak yang telah membantu secara langsung maupun secara tidak langsung yang tidak bias penulis jelaskan satu persatu.

Penulis menyadari bahwa masih banyak kekurangan dalam penulisan Tesis ini. Oleh karena itu, kritik dan saran yang membangun sangat diharapkan demi kemajuan karya tulis khususnya yang berkenaan dengan Tesis ini. Penulis berharap semoga Tesis ini dapat bermanfaat bagi semua pihak khususnya Magister Ilmu Komputer Universitas Sriwijaya.

Palembang, 23 September 2021

Penulis

# **OPTIMIZATION POWER NORMALIZED CEPSTRAL COEFFICIENTS AND GROWING SELF ORGANIZING MAPS IN SPEECH RECOGNITION**

**Meutia Puspitasari (09042681721007)**

Dept of Master Computer Science, Computer Science Faculty, Sriwijaya University  
Email : meutiaps@gmail.com

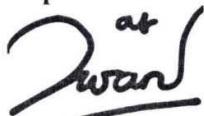
## **Abstraction**

Speech recognition system is a technology that allows devices to recognize words by digitizing words and matching digital signals with certain patterns stored in computer devices. The purpose of this study is to develop a speech recognition technique by implementing the Power Normalized Cepstral Coefficients for feature extraction method and classifying it using the Self Organizing Maps and Growing Self Organizing Maps methods. Model analysis technique is using the Confusion Matrix. The dataset is used in the form of voice data taken manually with a total of 200 voice data from 20 respondents. The results of this analysis obtained several conclusions, the first Power Normalized Cepstral Coefficients method is able to make voice data that has noise become a good feature for training and testing. The second, methods of Growing Self Organizing Maps is able to produce the same accuracy as the method of Self Organizing Maps with a smaller number of iterations and final weight nodes than the Self Organizing Maps method requires. The third, the spread factor value has an influence to form the right model by determining the growth of nodes. The fourth, Self Organizing Maps and Growing Self Organizing Maps are both capable of producing good models in terms of model performance measurement with the support of the Power Normalized Cepstral Coefficients feature achieving an accuracy of 95%.

**Keywords:** Speech Recognition, Power Normalized Cepstral Coefficients, Self Organizing Maps, Growing Self Organizing Maps, Confusion Matrix

## **Acknowledged,**

### **Supervisor I**



Dr. Iwan Pahendra, M.T.  
NIP. 197403222002121002

**Palembang, September 2021**

### **Supervisor II**



Dr. Abdiansah, S.Kom., M.CS.  
NIP. 198410012009121005

## **Head of Department Master Computer Science**



Dian Palupi Rini, M.Kom., Ph.D  
NIP. 197802232006042002

**OPTIMASI POWER NORMALIZED CEPSTRAL COEFFICIENTS DAN GROWING  
SELF ORGANIZING MAPS DALAM PENGENALAN SUARA**

**Meutia Puspitasari (09042681721007)**

Jurusan Magister Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Sriwijaya Email :  
meutiaps@gmail.com

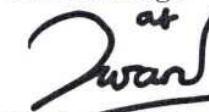
**Abstrak**

Sistem pengenalan suara adalah teknologi yang memungkinkan perangkat untuk mengenali kata-kata dengan cara digitalisasi kata dan mencocokkan sinyal digital dengan pola tertentu yang tersimpan dalam perangkat komputer. Tujuan dari penelitian ini adalah mengembangkan teknik pengenalan suara dengan mengimplementasikan metode extraksi fitur *Power Normalized Cepstral Coefficients* dan mengklasifikasikannya dengan metode *Self Organizing Maps* dan *Growing Self Organizing Maps*. Teknik analisis model menggunakan *Confusion Matrix*. Dataset yang digunakan berupa data suara yang diambil secara manual dengan jumlah 200 data suara dari 20 orang responden. Hasil analisis dari penelitian ini memperoleh beberapa kesimpulan yaitu pertama metode *Power Normalized Cepstral Coefficients* mampu menjadikan data suara yang mempunyai kebisingan menjadi fitur yang baik untuk pelatihan dan pengujian. Kedua metode *Growing Self Organizing Maps* mampu menghasilkan akurasi yang sama dengan metode *Self Organizing Maps* dengan jumlah iterasi dan node bobot akhir yang sedikit dari pada metode *Self Organizing Maps* perlukan. Ketiga, nilai *spread factor* punya pengaruh untuk membentuk model yang tepat dengan menentukan pertumbuhan node. Keempat, *Self Organizing Maps* dan *Growing Self Organizing Maps* keduanya mampu menghasilkan model yang baik dilihat dari pengukuran kinerja model dengan dukungan fitur *Power Normalized Cepstral Coefficients* mencapai akurasi sebesar 95%.

**Kata Kunci :** Sistem Pengenalan Suara, *Power Normalized Cepstral Coefficients*, *Self Organizing Maps*, *Growing Self Organizing Maps*, *Confusion Matrix*

Mengetahui,

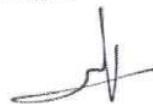
Pembimbing I



Dr. Iwan Pahendra, M.T.  
NIP. 197403222002121002

Palembang, September 2021

Pembimbing II



Dr. Abdiansah, S.Kom., M.CS.  
NIP. 198410012009121005

Koordinator Program Studi Magister Ilmu Komputer



Dian Palupi Rini, M.Kom., Ph.D  
NIP 197802232006042002

## DAFTAR ISI

<b>HALAMAN JUDUL</b>	<b>ii</b>
<b>HALAMAN PENGESAHAN</b>	<b>iii</b>
<b>HALAMAN PERSETUJUAN</b>	<b>iii</b>
<b>HALAMAN PERNYATAAN</b>	<b>iii</b>
<b>KATA PENGANTAR</b>	<b>vi</b>
<b>ABSTRACTION</b>	<b>viii</b>
<b>ABSTRAK</b>	<b>ix</b>
<b>DAFTAR ISI</b>	<b>x</b>
<b>DAFTAR GAMBAR</b>	<b>xiii</b>
<b>DAFTAR TABEL</b>	<b>xv</b>
<b>BAB I. PENDAHULUAN</b>	<b>1</b>
1.1    Latar Belakang Macalah	1
1.2    Perumusan Macalah	3
1.3    Tujuan Penelitian	3
1.4    Batasan Macalah	3
1.5    Manfaat Penelitian	4
1.6    Sistematika Penulisan	4
<b>BAB II. TINJAUAN PUSTAKA</b>	<b>6</b>
2.1    Ucapan	6
2.2    Pengenalan Ucapan	6
2.3    State Of The Art	7
2.3.1    Ekstraksi Fitur Power Normalized Cepstral Coefficients (PNCC)	9
2.3.1.1    Pre-Emphasis	10
2.3.1.2    Magnitude Square Of STFT	10
2.3.1.3    Medium Time Power Calculation	10
2.3.1.4    Asymmetric Noise Suppression with Temporal Masking	11
2.3.1.5    Weight Smoothing	11
2.3.1.6    Time-Frequency Normalization	11
2.3.1.7    Mean Power Normalization	11

2.3.9	Power Function Non Linyer	12
2.3.10	Discrete Transform (DCT)	12
2.4	Klasifikasi Self Organizing Maps (SOM)	12
2.4.1	Normalisasi Input	14
2.4.2	Menghitung Keluaran dari Tiap Output Neuron Untuk Menentukan Neuron Pemenang	14
2.4.3	Learning Rate	15
2.4.4	Penyesuaian Bobot	15
2.4.5	Menghitung Error	16
2.5	Klasifikasi dengan Growing Self Organizing Maps (GSOM)	16
2.6	Pengujian dan Evaluasi Klasifikasi	17
2.6.1	Pengujian dengan K-Fold Cross Validation	17
2.6.2	Evaluasi Hasil dengan Confusion Matrix	19
BAB III.	METODOLOGI PENELITIAN	20
3.1	Pengumpulan Data	20
3.2	Kerangka Penelitian	20
3.3	Skenario Penelitian	22
3.4	Tahapan Penelitian	23
3.5	Ekstraksi Fitur Power Normal Cepstral Coefficient (PNCC)	25
3.6	Implementasi Growing Self Organizing Maps (GSOM)	26
3.6.1	Pelatihan Growing Self Organizing Maps (GSOM)	26
3.6.2	Pengujian Growing Self Organizing Maps (GSOM)	28
3.7	Perangkat yang digunakan	28
3.8	Pengujian pada hasil klasifikasi dengan Growing Self Organizing Maps (GSOM)	29
3.9	Analisis Hasil	29
3.10	Kesimpulan dan Saran	29
3.12	Jadwal Penelitian	30
BAB IV.	HASIL DAN ANALISA	31
4.1	Audio Database	30
4.1.1	Plot Sinyal Suara	31
4.1.2	Ekstraksi Fitur dan Plot Fitur PNCC	35
4.2	Hasil Pelatihan dan Pengujian Model SOM	36

4.2.1	Model 1	36
4.2.2	Model 2	39
4.2.3	Model 3	41
4.2.4	Model 4	44
4.2.5	Model 5	47
4.3	Hasil Pelatihan dan Pengujian Model GSOM	49
4.3.1	Model 1	50
4.3.2	Model 2	53
4.3.3	Model 3	55
4.3.4	Model 4	58
4.3.5	Model 5	61
4.4	Analisis Perbandingan Model	64
4.4.1	Analisis Perbandingan Model SOM	72
4.4.2	Analisis Perbandingan Model GSOM	74
4.5	Analisis Perbandingan SOM dengan GSOM	75
BAB V.	KESIMPULAN DAN SARAN	77
5.1	Kesimpulan	77
5.2	Saran	77
DAFTAR PUSTAKA		78

## DAFTAR GAMBAR

<b>Gambar 2.1 Proses <i>Power Normalized Cepstral Coefficients (PNCC)</i>.....</b>	<b>16</b>
<b>Gambar 2.2 Struktur Jaringan Kohonen .....</b>	<b>20</b>
<b>Gambar 2.3 Node Awal <i>Growing Self Organizing Maps (GSOM)</i> dan Pertumbuhan Node.....</b>	<b>24</b>
<b>Gambar 2.4 Ilustrasi <i>4-Fold Cross Validation</i>.....</b>	<b>25</b>
<b>Gambar 3.1 Kerangka Penelitian.....</b>	<b>28</b>
<b>Gambar 3.2Skenario Penelitian .....</b>	<b>29</b>
<b>Gambar 3.3 Diagram Alir Tahapan Penelitian .....</b>	<b>31</b>
<b>Gambar 3.4 Proses <i>Power Normalized Cepstral Coefficients (PNCC)</i>.....</b>	<b>32</b>
<b>Gambar 3.5Proses Pelatihan <i>Growing Self Organizing Maps (GSOM)</i> .....</b>	<b>33</b>
<b>Gambar 4.1Plot Fitur PNCC.....</b>	<b>36</b>
<b>Gambar 4.2Confusion Matrix Model 1 SOM K-Fold=5 .....</b>	<b>37</b>
<b>Gambar 4.3Confusion Matrix Model 1 SOM K-Fold=10 .....</b>	<b>39</b>
<b>Gambar 4.4Confusion Matrix Model 2 SOM K-Fold=5 .....</b>	<b>40</b>
<b>Gambar 4.5Confusion Matrix Model 2 SOM K-Fold=10 .....</b>	<b>41</b>
<b>Gambar 4.6Confusion Matrix Model 3 SOM K-Fold=5 .....</b>	<b>43</b>
<b>Gambar 4.7Confusion Matrix Model 3 SOM K-Fold=10 .....</b>	<b>44</b>
<b>Gambar 4.8Confusion Matrix Model 4 SOM K-Fold=5 .....</b>	<b>45</b>
<b>Gambar 4.9Confusion Matrix Model 4 SOM K-Fold=10 .....</b>	<b>47</b>
<b>Gambar 4.10Confusion Matrix Model 5 SOM K-Fold=5 .....</b>	<b>48</b>
<b>Gambar 4.11Confusion Matrix Model 5 SOM K-Fold=10 .....</b>	<b>50</b>
<b>Gambar 4.10Confusion Matrix Model 5 SOM K-Fold=5 .....</b>	<b>48</b>
<b>Gambar 4.11Confusion Matrix Model 5 SOM K-Fold=10 .....</b>	<b>50</b>
<b>Gambar 4.12Confusion Matrix Model 1 GSOM K-Fold=5 .....</b>	<b>52</b>
<b>Gambar 4.13Confusion Matrix Model 1 GSOM K-Fold=10 .....</b>	<b>53</b>
<b>Gambar 4.14Confusion Matrix Model 2 GSOM K-Fold=5 .....</b>	<b>54</b>
<b>Gambar 4.15Confusion Matrix Model 2 GSOM K-Fold=10 .....</b>	<b>56</b>
<b>Gambar 4.16Confusion Matrix Model 3 GSOM K-Fold=5 .....</b>	<b>57</b>
<b>Gambar 4.17Confusion Matrix Model 3 GSOM K-Fold=10 .....</b>	<b>58</b>
<b>Gambar 4.18Confusion Matrix Model 4 GSOM K-Fold=5 .....</b>	<b>60</b>

<b>Gambar 4.19Confusion Matrix Model 4 GSOM K-Fold=10 .....</b>	<b>61</b>
<b>Gambar 4.20Confusion Matrix Model 5 GSOM K-Fold=5 .....</b>	<b>62</b>
<b>Gambar 4.21Confusion Matrix Model 5 GSOM K-Fold=10 .....</b>	<b>64</b>

## DAFTAR TABEL

<b>Tabel 2.1</b> <i>State of The Art</i> .....	13
<b>Tabel 2.2</b> Ilustrasi Evaluasi <i>Confusion Matrix</i> .....	26
<b>Tabel 3.1</b> Jadwal Penelitian.....	37
<b>Tabel 4.1</b> Plot Sinyal Suara Satu Orang yang sama dari Datasets .....	32
<b>Tabel 4.2</b> Plot Sinyal Suara dari 3 Orang yang Berbeda .....	33
<b>Tabel 4.3</b> <i>Hyperparameter Model 1 SOM</i> .....	37
<b>Tabel 4.4</b> Akurasi Model 1 SOM K-Fold=5 .....	37
<b>Tabel 4.5</b> Hasil Perhitungan Confusion Matrix Model 1 SOM K-Fold=5.....	38
<b>Tabel 4.6</b> Akurasi Model 1 SOM K-Fold=10 .....	38
<b>Tabel 4.7</b> Hasil Perhitungan Confusion Matrix Model 1 SOM K-Fold=10.....	39
<b>Tabel 4.8</b> <i>Hyperparameter Model 2 SOM</i> .....	40
<b>Tabel 4.9</b> Akurasi Model 2 SOM K-Fold=5 .....	40
<b>Tabel 4.10</b> Hasil Perhitungan Confusion Matrix Model 2 SOM K-Fold=5.....	40
<b>Tabel 4.11</b> Akurasi Model 2 SOM K-Fold=10.....	41
<b>Tabel 4.12</b> Hasil Perhitungan Confusion Matrix Model 2 SOM K-Fold=10.....	42
<b>Tabel 4.13</b> <i>Hyperparameter Model 3 SOM</i> .....	42
<b>Tabel 4.14</b> Akurasi Model 3 SOM K-Fold=5.....	42
<b>Tabel 4.15</b> Hasil Perhitungan Confusion Matrix Model 3 SOM K-Fold=5.....	43
<b>Tabel 4.16</b> Akurasi Model 3 SOM K-Fold=10.....	44
<b>Tabel 4.17</b> Hasil Perhitungan Confusion Matrix Model 3SOM K-Fold=10.....	44
<b>Tabel 4.18</b> <i>Hyperparameter Model 4 SOM</i> .....	45
<b>Tabel 4.19</b> Akurasi Model 4 SOM K-Fold=5.....	45
<b>Tabel 4.20</b> Hasil Perhitungan Confusion Matrix Model 4 SOM K-Fold=5.....	46
<b>Tabel 4.21</b> Akurasi Model 4 SOM K-Fold=10.....	46
<b>Tabel 4.22</b> Hasil Perhitungan Confusion Matrix Model 4SOM K-Fold=10.....	47
<b>Tabel 4.23</b> <i>Hyperparameter Model 5 SOM</i> .....	48
<b>Tabel 4.24</b> Akurasi Model 5 SOM K-Fold=5.....	48
<b>Tabel 4.25</b> Hasil Perhitungan Confusion Matrix Model 5 SOM K-Fold=5.....	49
<b>Tabel 4.26</b> Akurasi Model 5 SOM K-Fold=10.....	49

<b>Tabel 4.27Hasil Perhitungan Confusion Matrix Model 5SOM K-Fold=10 .....</b>	50
<b>Tabel 4.28<i>Hyperparameter</i> Model 1 GSOM .....</b>	51
<b>Tabel 4.29Akurasi Model 1 GSOM K-Fold=5 .....</b>	51
<b>Tabel 4.30Hasil Perhitungan Confusion Matrix Model 1 GSOM K-Fold=5 .....</b>	52
<b>Tabel 4.31Akurasi Model 1 GSOM K-Fold=10 .....</b>	52
<b>Tabel 4.32Hasil Perhitungan Confusion Matrix Model 1 GSOM K-Fold=10 .....</b>	53
<b>Tabel 4.33<i>Hyperparameter</i> Model 2GSOM .....</b>	54
<b>Tabel 4.34Akurasi Model 2GSOM K-Fold=5 .....</b>	54
<b>Tabel 4.35Hasil Perhitungan Confusion Matrix Model 2GSOM K-Fold=5 .....</b>	55
<b>Tabel 4.36Akurasi Model 2GSOM K-Fold=10 .....</b>	56
<b>Tabel 4.37Hasil Perhitungan Confusion Matrix Model 2 GSOM K-Fold=10 .....</b>	56
<b>Tabel 4.38<i>Hyperparameter</i> Model 3GSOM .....</b>	57
<b>Tabel 4.39Akurasi Model 3GSOM K-Fold=5 .....</b>	57
<b>Tabel 4.40Hasil Perhitungan Confusion Matrix Model 3GSOM K-Fold=5 .....</b>	57
<b>Tabel 4.41Akurasi Model 3GSOM K-Fold=10 .....</b>	58
<b>Tabel 4.42Hasil Perhitungan Confusion Matrix Model 3 GSOM K-Fold=10 .....</b>	59
<b>Tabel 4.43<i>Hyperparameter</i> Model 4GSOM .....</b>	59
<b>Tabel 4.44Akurasi Model 4GSOM K-Fold=5 .....</b>	59
<b>Tabel 4.45Hasil Perhitungan Confusion Matrix Model 4GSOM K-Fold=5 .....</b>	60
<b>Tabel 4.46Akurasi Model 4GSOM K-Fold=10 .....</b>	61
<b>Tabel 4.47Hasil Perhitungan Confusion Matrix Model 4 GSOM K-Fold=10 .....</b>	61
<b>Tabel 4.48<i>Hyperparameter</i> Model 5GSOM .....</b>	62
<b>Tabel 4.49Akurasi Model 5GSOM K-Fold=5 .....</b>	62
<b>Tabel 4.50Hasil Perhitungan Confusion Matrix Model 5GSOM K-Fold=5 .....</b>	63
<b>Tabel 4.51Akurasi Model 5GSOM K-Fold=10 .....</b>	63
<b>Tabel 4.52Hasil Perhitungan Confusion Matrix Model 5 GSOM K-Fold=10 .....</b>	64
<b>Tabel 4.53Data Sample .....</b>	65
<b>Tabel 4.54Hasil <i>Pre Emphasis</i> .....</b>	66
<b>Tabel 4.55Hasil Perhitungan <i>Short Time Fourier Transform (STFM)</i> .....</b>	67
<b>Tabel 4.56Hasil Perhitungan <i>Gammatone Frequency Integration</i> .....</b>	67
<b>Tabel 4.57Hasil Perhitungan <i>Medium Teim Power Calculation</i> .....</b>	67

<b>Tabel 4.58Hasil Perhitungan <i>Asymetric Noise Sippression Temporal Masking</i></b> .....	67
<b>Tabel 4.59Hasil Perhitungan <i>Weight Smoothing</i></b> .....	67
<b>Tabel 4.60Hasil Perhitungan <i>Time Frequency Normalization</i></b> .....	68
<b>Tabel 4.61Perhitungan <i>Mean Power Normalizaion</i></b> .....	68
<b>Tabel 4.62Hasil Perhitungan <i>Power Normalization Non Linearty</i></b> .....	68
<b>Tabel 4.63Hasil Perhitungan <i>Discrete Transfrom (DCT)</i></b> .....	68
<b>Tabel 4.64Parameter <i>Self Organizing Maps (SOM)</i></b> .....	69
<b>Tabel 4.65Nilai <i>PowerNormalized Cepstral Coefficients (PNCC)</i></b> .....	69
<b>Tabel 4.66Bobot <i>Self Organizing Maps (SOM)2x2</i></b> .....	69
<b>Tabel 4.67<i>Self Organizing Maps (SOM)</i> Iterasi 1</b> .....	69
<b>Tabel 4.68Hasil Perhitungan <i>Self Organizing Maps (SOM)</i></b> .....	70
<b>Tabel 4.69Parameter <i>Growing Self Organizing Maps (GSOM)</i></b> .....	70
<b>Tabel 4.70Nilai <i>PowerNormalized Cepstral Coefficients (PNCC)</i></b> .....	70
<b>Tabel 4.71Bobot Awal <i>Growing Self Organizing Maps (GSOM)</i></b> .....	71
<b>Tabel 4.72Bobot <i>Growing Self Organizing Maps (GSOM)</i></b> .....	71
<b>Tabel 4.73Bobot Akhir <i>Growing Self Organizing Maps (GSOM)</i></b> .....	71
<b>Tabel 4.74Ringkasan Pelatihan dan Pengujian Model dengan Metode SOM</b> ...	73
<b>Tabel 4.75Ringkasan Pelatihan &amp; Pengujian Model dengan Metode GSOM</b> ....	74

## BAB I. PENDAHULUAN

### 1.1 Latar Belakang Masalah

Pengenalan suara adalah teknologi yang memungkinkan perangkat untuk mengenali kata-kata dengan cara digitalisasi kata dan mencocokkan sinyal digital dengan pola tertentu yang tersimpan dalam perangkat komputer. Pada sistem pengenalan suara ini, agar kata-kata yang diucapkan bisa dikenali dengan baik, maka kondisi lingkungan yang ada di sekitar mesin atau alat harus dalam keadaan tenang dan bebas dari gangguan suara lain atau *noise* yang dapat menghambat atau mengganggu kinerja sistem. Teknik untuk sistem pengenalan suara ini dibagi menjadi dua proses yaitu *feature extraction* dan *voice feature matching* atau proses klasifikasi (Saini dkk., 2017).

Sistem pengenalan suara sebagian besar keunggulannya ditentukan oleh metode ekstraksi fitur yang digunakan. Ada banyak metode ekstraksi fitur yang sering digunakan untuk sistem pengenalan suara ini seperti *Power Normalized Cepstral Coefficients (PNCC)*, *Mel Frequency Cepstrum Coefficients (MFCC)*, *Modified Group Delay Function (ModGDF)*, *Perceptual Linear Prediction*, *Vector Quantization*, dan *K-Means Algorithm* (Saini dkk., 2017; Alasadi dkk., 2020; Lee dkk., 2020). Pada penelitian yang dilakukan oleh Alasadi dkk (2020) tentang penggunaan metode *Power Normalized Cepstral Coefficients (PNCC)* sebagai ekstraksi fitur untuk pengenalan suara dalam bahasa Arab, metode tersebut mencapai akurasi tertinggi yaitu 93% dibandingkan dengan metode *Mel Frequency Cepstrum Coefficients (MFCC)* dengan akurasi 88% dan *Modified Group Delay Function (ModGDF)* yang mendapatkan akurasi 90% untuk suara yang mempunyai tingkat kebisingan yang tinggi. Sedangkan pada penelitian yang dilakukan oleh Kurnian dkk (2020) tentang penggunaan metode *Power Normalized Cepstral Coefficients (PNCC)* sebagai ekstraksi fitur untuk *Forensic Automatic Speaker Recognition (FASR)* dengan menggunakan database Malayan menunjukkan bahwa *Power Normalized Cepstral Coefficients (PNCC)* lebih tangguh pada suara yang mempunyai kebisingan dengan persentase *Equal Error*

*Rate (ERR)* 29,99% lebih kecil dibandingkan *Mel Frequency Cepstrum Coefficients (MFCC)* dengan *Equal Error Rate (ERR)* 41,97%.

*Power Normalized Cepstral Coefficients (PNCC)* tidak bergantung pada kualitas ucapan sedangkan *Mel Frequency Cepstrum Coefficients (MFCC)* sangat bergantung pada kualitas ucapan, semakin tinggi tingkat kebisingan maka akurasi yang didapatkan akan semakin menurun. *Power Normalized Cepstral Coefficients (PNCC)* memberikan hasil yang lebih baik dalam sisi akurasi dibandingkan metode lain (Al-Kaltakchi dkk., 2019).

Setelah proses ekstraksi fitur, hasil dari proses tersebut dilanjutkan ke tahap proses klasifikasi. *Self Organizing Maps (SOM)* pertama kali diperkenalkan oleh Profesor yang berasal dari Finlandia, yaitu Teuvo Kohonen. *Self Organizing Maps (SOM)* adalah model jaringan saraf tiruan dengan metode pembelajaran tidak terawasi. Salah satu keunggulan dari algoritma *Self Organizing Maps (SOM)* adalah mampu memetakan data dalam dimensi tinggi ke bentuk dimensi rendah (Nawaratne dkk., 2019). Menurut penelitian yang dilakukan oleh Devi K.J. dkk (2020), penggunaan *Self Organizing Maps (SOM)* sebagai teknik klasifikasi dalam pengenalan suara mendapatkan akurasi 93%.

Ada pula metode *Growing Self Organizing Maps (GSOM)* yang merupakan varian atau versi perluasan dari *Self Organizing Maps (SOM)*. *Growing Self Organizing Maps (GSOM)* dikembangkan untuk mengatasi masalah ukuran yang cocok untuk peta *Self Organizing Maps (SOM)*. *Growing Self Organizing Maps (GSOM)* memiliki parameter *spread factor* dan efek dari *spread factor* dapat digunakan untuk mengukur dan mengontrol penyebaran *Growing Self Organizing Maps (GSOM)*. *Spread factor* tidak bergantung pada dimensi data dan dengan demikian dapat digunakan sebagai ukuran pengontrol untuk menghasilkan peta dengan dimensi yang berbeda, yang kemudian dapat dibandingkan dan dianalisis dengan lebih akurat (Alahakoon dkk., 2000; Nawaratne dkk., 2019).

Berdasarkan penjelasan tersebut, maka tesis ini akan menerapkan proses ekstraksi fitur menggunakan *Power Normalized Cepstral Coefficients (PNCC)* dan *Growing Self Organizing Maps (GSOM)* untuk proses klasifikasi pengenalan suara yang berfokus pada angka. Hasil penelitian ini akan mengkonversi sinyal suara manusia ke dalam bentuk teks.

## **1.2 Perumusan Masalah**

Dari latar belakang diatas terdapat beberapa isu yang akan dibahas dalam penelitian ini :

1. Bagaimana proses ekstraksi fitur pada data suara menggunakan metode *Power Normalized Cepstral Coefficients (PNCC)* mempengaruhi hasil klasifikasi.
2. Bagaimana klasifikasi suara dengan menggunakan metode *Growing Self Organizing Maps (GSOM)* dapat dilakukan.
3. Bagaimana tingkat akurasi yang dihasilkan pada klasifikasi suara dengan *Growing Self Organizing Maps (GSOM)* dibandingkan dengan *Self Organizing Maps (SOM)* konvensional.
4. Bagaimana parameter *spread factor* pada *Growing Self Organizing Maps (GSOM)* dapat mempengaruhi hasil klasifikasi hingga mencapai titik optimal.

## **1.3 Tujuan Penelitian**

Secara umum tujuan dari penelitian ini adalah mengembangkan teknik pengenalan suara dengan mengimplementasikan metode extraksi fitur *Power Normalized Cepstral Coefficients (PNCC)* dan mengklasifikasikannya dengan metode *Growing Self Organizing Maps (GSOM)*.

Adapun tujuan yang ingin dicapai dari penelitian ini adalah :

1. Mengetahui performa klasifikasi *Growing Self Organizing Maps (GSOM)*
2. Melakukan percobaan trial and error parameter dasar atau *spread factor* pada *Growing Self Organizing Maps (GSOM)* untuk mendapatkan *spread factor* yang cocok
3. Mengukur tingkat keakuratan metode *Growing Self Organizing Maps (GSOM)* dalam pengenalan suara.
4. Menganalisis perbandingan akurasi pengenalan suara antara metode *Growing Self Organizing Maps (GSOM)* dan *Self Organizing Maps (SOM)*.

#### **1.4 Batasan Masalah**

Dalam pengenalan pola sinyal suara dengan *Power Normalized Cepstral Coefficients (PNCC)* dan *Growing Self Organizing Maps (GSOM)* terdapat pembatasan masalah sebagai berikut:

1. Data suara memiliki 1 kanal (mono) dan pada lingkungan *noise*.
2. Output dalam penelitian ini adalah teks yang dikonversikan dari suara berdasarkan hasil klasifikasi.
3. Data suara diambil dan direkam dalam bentuk berkas (*file*) bukan dalam keadaan *real time*.
4. Data suara yang digunakan berjumlah 20 data suara dari 20 orang responden. Pada proses *training* digunakan 15 data suara, sedangkan pada proses pengenalan digunakan 5 data suara.
5. *Dataset public* juga akan digunakan yaitu dataset *Speech Accent Archive*.

#### **1.5 Manfaat Penelitian**

Secara umum tujuan dari penelitian ini adalah mengembangkan teknik pengenalan suara dengan mengimplementasikan metode extraksi fitur *Power Normalized Cepstral Coefficients (PNCC)* dan mengklasifikasikannya dengan metode *Growing Self Organizing Maps (GSOM)*.

Adapun manfaat yang ingin dicapai dari penelitian ini adalah :

1. Dapat menjadi pilihan alternatif untuk penelitian selanjutnya.
2. Setelah penelitian ini diharapkan ke depannya, dapat diwujudkan ke dalam aplikasi yang dapat diimplementasikan dalam kehidupan sehari-hari.

#### **1.6 Sistematika Penulisan**

Agar memperoleh gambaran jelas mengenai penelitian ini, maka dibuatlah suatu sistematika penulisan yang berisi gambaran dalam tiap bab suatu penelitian, yaitu:

##### **1. BAB I : Pendahuluan**

Bab ini menjelaskan tentang latar belakang, perumusan masalah, batasan masalah, tujuan dan manfaat serta sistematika penulisan dari topic yang dipilih berupa pengoptimasi *Power Normalized Cepstral Coefficients*

(PNCC) dan *Growing Self Organizing Maps* (GSOM) dalam Pengenalan Suara.

**2. BAB II : Tinjauan Pustaka**

Bab ini menjelaskan mengenai literature *review* yang berhubungan dengan pengoptimasian *Power Normalized Cepstral Coefficients* (PNCC) dan *Growing Self Organizing Maps* (GSOM) dalam Pengenalan Suara.

**3. BAB III : Metodologi Penelitian**

Bab ini menjelaskan tahapan atau metode secara terperinci tentang langkah-langkah yang digunakan untuk mencari, mengumpulkan dan menganalisa terkait pengoptimasian *Power Normalized Cepstral Coefficients* (PNCC) dan *Growing Self Organizing Maps* (GSOM) serta model yang digunakan sehingga tujuan dari penulisan dapat tercapai.

**4. BAB IV : Hasil dan Analisa**

Bab ini menjelaskan hasil dan analisa terkait pengoptimasian *Power Normalized Cepstral Coefficients* (PNCC) dan *Growing Self Organizing Maps* (GSOM).

**5. BAB V : Kesimpulan dan Saran**

Bab ini menjelaskan kesimpulan dan saran dari penelitian yang dilakukan terkait pengoptimasian *Power Normalized Cepstral Coefficients* (PNCC) dan *Growing Self Organizing Maps* (GSOM).

## DAFTAR PUSTAKA

- Al-Kaltakchi, M,T,S, Taha, H, A, A, Shehab, M, A, and Abdullah, M, A, M. (2020). Comparison of feature extraction and normalization methods for speaker recognition using grid-audiovisual database. Indonesian Journal of Electrical Engineering and Computer Science Vol. 18, No. 2, May 2020, pp. 782~789.
- Alahakoon,D, Halgamuge, S, and Bala, S. (2000). Dynamic Self-Organizing Maps With Controlled Growth For Knowledge Discovery. IEEE Transactions on Neural Networks 11(3):601 – 614
- Alasadi, A.A, Adhyani, T,H,H. (2020). Efficient Feature Extraction Algorithms to Develop an Arabic Speech Recognition System. Engineering, Technology & Applied Science Research Vol. 10, No. 2, 2020, 5547-5553.
- Benkerzaz, S, Y. Elmir, and A. Dennai. 2019. “A Study on Automatic Speech Recognition”. 10(3). 77-85. [10.6025/jitr/2019/10/3/77-85](https://doi.org/10.6025/jitr/2019/10/3/77-85)
- Du, X. P. and He, P. L. (2006). The Clustering Solution of Speech Recognition Models with SOM. *In Proceedings of the Third international conference on Advances in Neural Networks*. Heidelberg (Berlin), 150-157.
- Dutta, A, and Rao K. 2015. Robust Language Identification using Power Normalized Cepstral Coefficients. 2015 Eighth International Conference on Contemporary Computing (IC3)
- Goronescu, Florin. (2011). Data Minig: Concepts and Tecniques. Verlag Berlin Heidelberg Springer
- Han, J., and Kamber M., Pei (2012). Data Mining:Concept and Techniques. New

York:Morgan Kaufmann Publisher

Henny Leidiyana. (2013). Penerapan Algoritma K-Nearest Neighbour untuk penentuan resiko kredit kepemilikan bermotor.

Khan U, Ahmad M, Sarim M, and Shafiq F. 2019. Feature Extraction and Modeling Tecniques in Speech Recognition: A Review

Kim C and Stern R, 2016. Power-Normalized Cepstral Coefficients (PNCC) for Robust Speech Recognition. IEEE/ACM Transactions on Audio, Speech, and Language Processing (Volume: 24 , Issue: 7 , July 2016)

Kurian, B, Sreehari, V, R, and Mary, L. (2020). PNCC for Forensic Automatic Speaker Recognition. Proceedings of the International Conference on Microelectronics, Signals and Systems 2019 AIP Conf. Proc. 2222, 030004-1–030004-6; <https://doi.org/10.1063/5.0003967> Published by AIP Publishing. 978-0-7354-1981-0/\$30.00

Kurian, B, Shanavaz K T, and Nikhil G Kurup. 2017. “PNCC Based Speech Enhancement and Its Performance Evaluation using SNR loss.” International Conference on Networks & Advances in Computational Technologies (NetACT). 20-22 July 2017.

Nawaratne, R, Adikari, A, Alahakoon, D, Silva, D,D, and Chilamkurti, N. (2020). Recurrent Self-Structuring Machine Learning for Video Processing using Multi-Stream Hierarchical Growing Self-Organizing Maps. Research Centre for in Data Analytics and Cognition, La Trobe University, Melbourne, Australia.

Ortega, C. A. D. L., Gonzales, M. M., Medina M. A. A., Romo, J. C. M., and Villar, V. E. G. D. (2009). Analysis of Kohonen’s Neural Network with

Application to Speech Recognition. *In proceeding of: MICAI 2009, Workshop Computer Vision and Pattern Recognition – WCVPR.* Juny.

Saini. N, Singh, D. 2017. A Hybrid Approach for Voice Recognition Using MFCC, VQ and KMeans Algorithm. International Journal of Advanced Research in Computer Science.

Nawaratne, R, Alahakoon, D, Silva, D,D and Yu, X. (2019). HT-GSOM: Dynamic Self-organizing Map with Transience for Human Activity Recognition. *IEEE 2019*.

Nilashi, M, Ahmadi, H, Sheikhtaheri, A, Naemi, R, Alotaibi, R and Zhao, J. (2020). Remote tracking of Parkinson's Disease progression using ensembles of Deep Belief Network and Self-Organizing Map. *Expert Systems with Applications*. Volume 159, 30 November 2020, 113562

Zhang Y, Xu K, and Wan J.2018.Rubost Feature for Underwater Targets Recognition UsingPower-Normalized Cepstral Coefficients.” IEEE International Conference on Signal Processing.