



ITS
Institut
Teknologi
Sepuluh Nopember

TUGAS AKHIR - TJ 141502

**KLASIFIKASI *OBSTRUCTIVE SLEEP APNEA (OSA)*
BERDASARKAN FITUR STATISTIK DARI RR INTERVAL
PADA SINYAL ECG**

Desanti Nurma Risalah
NRP 07211440000030

Dosen Pembimbing
Prof. Dr. Ir. Yoyon K. Suprpto, M.Sc.
Dr. Diah Puspito W, ST., M.Sc.

DEPARTEMEN TEKNIK KOMPUTER
Fakultas Teknologi Elektro
Institut Teknologi Sepuluh Nopember
Surabaya 2018



ITS
Institut
Teknologi
Sepuluh Nopember

TUGAS AKHIR - TJ 141502

**KLASIFIKASI *OBSTRUCTIVE SLEEP APNEA (OSA)*
BERDASARKAN FITUR STATISTIK DARI RR INTERVAL
PADA SINYAL ECG**

Desanti Nurma Risalah
NRP 0721144000030

Dosen Pembimbing
Prof. Dr. Ir. Yoyon K. Suprpto, M.Sc.
Dr. Diah Puspito W, ST., M.Sc.

DEPARTEMEN TEKNIK KOMPUTER
Fakultas Teknologi Elektro
Institut Teknologi Sepuluh Nopember
Surabaya 2018



FINAL PROJECT - TJ 141502

CLASSIFICATION OF OBSTRUCTIVE SLEEP APNEA (OSA) BASED ON STATISTICAL FEATURES OF RR INTERVAL ON ECG SIGNALS

Desanti Nurma Risalah
NRP 0721144000030

Advisors

Prof. Dr. Ir. Yoyon K. Suprpto, M.Sc.
Dr. Diah Puspito W, ST., M.Sc.

DEPARTEMENT OF COMPUTER ENGINEERING
Faculty of Electrical Engineering
Sepuluh Nopember Institute of Technology
Surabaya 2018

PERNYATAAN KEASLIAN

TUGAS AKHIR

Dengan ini saya menyatakan bahwa isi sebagian maupun keseluruhan Tugas Akhir saya dengan judul "Klasifikasi *Obstructive Sleep Apnea* (OSA) berdasarkan Fitur Statistik dari RR Interval pada Sinyal ECG" adalah benar-benar hasil karya intelektual mandiri, diselesaikan tanpa menggunakan bahan-bahan yang tidak diijinkan dan bukan karya pihak lain yang saya akui sebagai karya sendiri.

Semua referensi yang dikutip maupun dirujuk telah ditulis secara lengkap pada daftar pustaka.

Apabila ternyata pernyataan ini tidak benar, saya bersedia menerima sanksi sesuai peraturan yang berlaku.

Surabaya, Juli 2018



Desanti Nurma Risalah
NRP. 0721144000030

LEMBAR PENGESAHAN

Klasifikasi *Obstructive Sleep Apnea (OSA)* berdasarkan Fitur Statistik dari RR Interval pada Sinyal ECG

Tugas Akhir ini disusun untuk memenuhi salah satu syarat memperoleh gelar Sarjana Teknik di Institut Teknologi Sepuluh Nopember Surabaya

Oleh : Desanti Nurma Risalah (NRP: 07211440000030)

Tanggal Ujian : 28 Juni 2018

Periode Wisuda : September 2018

Disetujui oleh:

Prof. Dr. Ir. Yoyon K. Suprpto, M.Sc.
NIP: 195409251978031001

(Pembimbing I)

Dr. Diah Puspito Wulandari, ST., M.Sc.
NIP: 198012192005012001

(Pembimbing II)

Dr. Adhi Dharma Wibawa, ST., MT.
NIP: 197605052008121003

(Penguji I)

Dr. I Ketut Eddy Purnama, ST., MT.
NIP: 196907301995121001

(Penguji II)

Eko Pramunanto, ST., MT.
NIP: 197409072002121001

(Penguji III)

Mengetahui
Kepala Departemen Teknik Komputer

Dr. I Ketut Eddy Purnama, ST., MT.
NIP: 196907301995121001

DEPARTEMEN
TEKNIK KOMPUTER

ABSTRAK

Nama Mahasiswa : Desanti Nurma Risalah
Judul Tugas Akhir : Klasifikasi *Obstructive Sleep Apnea* (OSA) berdasarkan Fitur Statistik dari RR Interval pada Sinyal ECG
Pembimbing : 1. Prof. Dr. Ir. Yoyon K. Suprpto, M.Sc.
2. Dr. Diah Puspito W, ST., M.Sc.

Obstructive sleep apnea (OSA) merupakan bentuk gangguan umum yang menyebabkan penderitanya mengalami henti napas saat tidur yang dapat memicu munculnya penyakit kardiovaskular jika tidak tertangani dengan benar. Secara umum, pengujian apnea biasanya menggunakan *polysomnography* (PSG), yang merupakan prosedur standar untuk diagnosis semua gangguan tidur. Namun sebagian besar kasus *sleep apnea* saat ini tidak terdiagnosis karena masalah biaya yang mahal dan kurang tersedia karena PSG dilakukan di laboratorium khusus, dimana para staf yang bertugas harus bekerja semalaman. Penelitian ini mengusulkan sebuah metode untuk mendeteksi OSA berdasarkan sinyal ECG. Sinyal ECG dapat dimanfaatkan untuk membuat sistem deteksi yang lebih sederhana dan penanganan yang lebih cepat daripada PSG. Data berupa sinyal ECG didapatkan dari Physionet Database dengan menggunakan 12 subjek rekaman berbeda. Data direkam selama 2 malam berturut-turut menggunakan perekaman *single-channel* ECG. Sinyal tersebut kemudian dilakukan ekstraksi fitur-fitur statistiknya berdasarkan jarak antar puncak R atau RR interval sinyal ECG yang diproses dalam durasi periode yang singkat. Hasil ekstraksi fitur tersebut digunakan sebagai input klasifikasi menggunakan metode *Support Vector Machine* untuk dilatih dan diujikan pada rekaman apnea dan non-apnea dari subjek positif dan negatif OSA. Dilakukan pula pengujian untuk mengetahui kinerjanya dengan mendapatkan nilai *Accuracy*, *Sensitivity*, dan *Specificity* dari sistem. Hasil menunjukkan bahwa sistem yang telah dibuat dapat mendeteksi periode gangguan OSA dengan akurasi yang cukup tinggi, yakni sekitar 83.67%. Kedepannya diharapkan sistem yang telah dibuat dapat terus dikembangkan dan dapat digunakan sebagai dasar pengembangan alat deteksi OSA di kemudian hari.

Kata Kunci : ECG; OSA; RR interval; *Support Vector Machine*

Halaman ini sengaja dikosongkan

ABSTRACT

Name : Desanti Nurma Risalah
Title : Classification of Obstructive Sleep Apnea (OSA)
based on Statistical Features of RR Interval on ECG
Signals
Advisors : 1. Prof. Dr. Ir. Yoyon K. Suprpto, M.Sc.
2. Dr. Diah Puspito W, ST., M.Sc.

Obstructive sleep apnea (OSA) is a common disorder that person stop breathing during sleep that can cause cardiovascular disease if it is not handled well. Most of sleep apnea cases were tested with polysomnography (PSG), which is the standard procedure for diagnose all sleep disorders. But most of sleep apnea cases are currently undiagnosed because of expenses and inconvenient, because PSG is done at sleep labs, where an expert human observer is needed to work overnight. This research proposes a method for detecting OSA based on ECG signal. ECG signals can be utilized to make detection system simpler and handling faster than PSG. The ECG signals that used are available from Physionet Database with 12 different records. Data is recorded for 2 consecutive nights using ECG single-channel recording. Statistical features are extracted based on RR intervals which processes short duration epochs of ECGs. Features extraction results are used as classification input using Support Vector Machine method to be trained and tested on sleep apnea records from subjects with and without OSA. System could determine its performance by obtaining Accuracy, Sensitivity, and Specificity score. The results show that the system that has been created can recognize epochs of OSA with a high degree of accuracy, approximately 83.67%. In the future, it is expected that the created system can be developed and can be used as basis for future development of a tool for OSA screening.

Keywords : ECG; OSA; RR interval; Support Vector Machine

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Puji syukur kehadiran Allah SWT atas segala limpahan berkah, rahmat, serta hidayah-Nya, penulis dapat menyelesaikan tugas akhir dengan judul Klasifikasi *Obstructive Sleep Apnea* (OSA) berdasarkan RR Interval Sinyal ECG sesuai dengan batas waktu yang telah ditentukan guna memenuhi salah satu syarat kelulusan Program Sarjana Teknik Komputer, ITS Dalam pengerjaannya, tugas akhir ini dapat terselesaikan karena bantuan dari berbagai pihak. Oleh sebab itu penulis mengucapkan terimakasih kepada:

1. Keluarga, Ibu, dan Ayah yang telah memberikan dorongan spiritual dan material dalam penyelesaian tugas akhir ini. Tanpa doa kalian penulis tidak akan bisa seperti ini.
2. Bapak Kepala Departemen Teknik Komputer Dr. I Ketut Eddy Purnama, ST., MT.
3. Bapak Prof. Dr. Ir. Yoyon Kusnendar Suprpto M.Sc. dan Ibu Dr. Diah Puspito Wulandari, ST., M.Sc. selaku dosen pembimbing, atas dukungan dan bimbingan selama mengerjakan tugas akhir.
4. Bapak-ibu dosen pengajar serta staff Departemen Teknik Komputer atas pengajaran, bimbingan, serta perhatian yang diberikan kepada penulis selama ini.
5. Teman-teman seperjuangan Teknik Komputer 2014 dan seluruh e54 yang pasti saling mendoakan yang terbaik untuk kita semua.
6. Mbak Vika S2, Mbak Ifut, Mbak Enny, Mas Tegar dan lain-lain, yang telah memberikan sedikit sumbangsuhnya namun sangat berarti dan menjadi bantuan besar untuk penulis. Serta seluruh pihak yang telah membantu kelancaran tugas akhir ini yang tidak bisa penulis sebutkan satu persatu.

Kesempurnaan hanya milik Allah SWT, untuk itu penulis memohon segenap kritik dan saran yang membangun. Semoga penelitian ini dapat memberikan manfaat bagi kita semua.

Suarabaya, Juli 2018

Penulis

Halaman ini sengaja dikosongkan

DAFTAR ISI

ABSTRAK.....	i
ABSTRACT.....	iii
KATA PENGANTAR	v
DAFTAR ISI.....	vii
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xiii
DAFTAR NOMENKLATUR.....	xv
1. BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Permasalahan	2
1.3 Tujuan	2
1.4 Batasan Masalah	2
1.5 Sistematika Penulisan.....	3
2. BAB 2 TINJAUAN PUSTAKA.....	5
2.1 <i>Sleep Apnea Syndrome</i>	5
2.1.1 <i>Obstructive Sleep Apnea</i>	5
2.2 Hubungan OSA dengan Perubahan ECG	6
2.3 <i>Electrocardiogram (ECG)</i>	7
2.3.1 Pengertian <i>Electrocardiogram</i>	7
2.3.2 RR Interval	8
2.4 <i>Heart Rate Variability</i>	10
2.4.1 Pengukuran HRV	10
2.5 <i>Support Vector Machine (SVM)</i>	13
2.5.1 Klasifikasi Linear pada SVM.....	14

2.5.2	Klasifikasi Non-linear pada SVM.....	15
2.5.3	Karakteristik SVM.....	17
2.5.4	Kelebihan dan Kekurangan SVM.....	17
2.5.5	SVM pada Python.....	19
2.6	Pengukuran Unjuk Kerja.....	20
3.	BAB 3 DESAIN DAN IMPLEMENTASI SISTEM.....	23
3.1	Desain Sistem.....	23
3.2	Implementasi Sistem.....	24
3.2.1	Pengumpulan Data.....	24
3.2.2	<i>Pre-processing</i>	26
3.2.3	Ekstraksi Fitur.....	30
3.2.4	Klasifikasi.....	32
3.2.5	Evaluasi Kinerja.....	35
4.	BAB 4 PENGUJIAN DAN ANALISIS.....	39
4.1	Hasil dan Pembahasan.....	39
4.1.1	Pengumpulan Data.....	39
4.1.2	<i>Pre-processing</i>	40
4.1.3	Ekstraksi Fitur.....	42
4.2	Klasifikasi dan Evaluasi Kinerja.....	47
4.2.1	Klasifikasi.....	47
4.2.2	Evaluasi Kinerja.....	49
4.2.3	Pengujian terhadap Komposisi Data.....	52
4.2.4	Pengujian terhadap Durasi Segmentasi Data.....	55
4.2.5.	Pengujian Terhadap Kernel SVM yang Digunakan.....	58
4.2.6.	Pengujian Pengurangan Fitur.....	60
5.	BAB 5 PENUTUP.....	65

5.1	Kesimpulan	65
5.2	Saran	66
DAFTAR PUSTAKA		67
LAMPIRAN.....		69
Lampiran I.		69
Lampiran II.....		92
Lampiran III.		93
Lampiran IV.		95
Lampiran V.		97
Lampiran VI.		100
BIODATA PENULIS		107

Halaman ini sengaja dikosongkan

DAFTAR GAMBAR

Gambar 2.1 Proses pengujian polisomnografi [18]	6
Gambar 2.2 Proses <i>electrocardiography</i> [3].	7
Gambar 2.3 Siklus normal sinyal ECG [9].	9
Gambar 2.4 Konsep Kerja SVM dalam memisahkan 2 <i>class</i> [14].	14
Gambar 3.1 Blok diagram penelitian	23
Gambar 3.2 Contoh rekaman sinyal dengan anotasinya pada menit ke 12:00 dan 13.00.....	25
Gambar 3.3 Visualisasi penentuan anotasi pada rekaman sinyal ECG .	26
Gambar 3.4 Listing program <i>training</i> data A.....	28
Gambar 3.5 Listing program <i>training</i> data B.....	29
Gambar 3.6 Hasil deteksi puncak R pada beberapa model sinyal.....	30
Gambar 3.7 Listing program <i>training</i> data C.....	31
Gambar 3.8 Listing program <i>training</i> data D.....	33
Gambar 3.9 Listing program <i>Testing</i> Data.....	34
Gambar 3.10 Flowchart program Training untuk klasifikasi	37
Gambar 3.11 Flowchart program Testing untuk klasifikasi	38
Gambar 4.1 Contoh hasil plot sinyal dengan durasi 15 detik.....	40
Gambar 4.2 Hasil plot sinyal saat deteksi R pada keadaan apnea.....	41
Gambar 4.3 Hasil plot sinyal saat deteksi R pada keadaan non-apnea..	42
Gambar 4.4 Visualisasi teknik pengujian yang dilakukan	48
Gambar 4.5 Grafik persentase nilai <i>Accuracy</i> pada setiap percobaan...	50
Gambar 4.6 Grafik persentase nilai <i>Sensitivity</i> pada setiap percobaan....	50
Gambar 4.7 Grafik persentase nilai <i>Specificity</i> pada setiap percobaan	51
Gambar 4.8 Visualisasi variasi komposisi data apnea dan non-apnea dalam pengujian	53
Gambar 4.9 Hasil akurasi terhadap variasi komposisi data apnea dan non-apnea sebagai data <i>training</i>	54
Gambar 4.10 Hasil persentase rata-rata <i>Accuracy</i> pada pengujian terhadap durasi segmentasi data.....	57
Gambar 4.11 Hasil persentase rata-rata <i>Accuracy</i> pada pengujian terhadap kernel SVM yang digunakan.	59
Gambar 4.12 Grafik tren hasil <i>accuracy</i> rata-rata pada pengujian pertama.....	62

Gambar 4.13 Grafik tren hasil accuracy rata-rata pada pengujian kedua	63
1 Bentuk sinyal rekaman a03 periode apnea durasi 15 detik	70
2 Bentuk sinyal rekaman a03 periode non-apnea durasi 15 detik.....	71
3 Bentuk sinyal rekaman a05 periode apnea durasi 15 detik	72
4 Bentuk sinyal rekaman a05 periode non- apnea durasi 15 detik.....	73
5 Bentuk sinyal rekaman a07 periode apnea durasi 15 detik	74
6 Bentuk sinyal rekaman a07 periode non- apnea durasi 15 detik.....	75
7 Bentuk sinyal rekaman a08 periode apnea durasi 15 detik	76
8 Bentuk sinyal rekaman a08 periode non-apnea durasi 15 detik.....	77
9 Bentuk sinyal rekaman a11 periode apnea durasi 15 detik.	78
10 Bentuk sinyal rekaman a11 periode non-apnea durasi 15 detik.....	79
11 Bentuk sinyal rekaman a13 periode apnea durasi 15 detik	80
12 Bentuk sinyal rekaman a13 periode non-apnea durasi 15 detik.....	81
13 Bentuk sinyal rekaman a15 periode apnea durasi 15 detik	82
14 Bentuk sinyal rekaman a15 periode non-apnea durasi 15 detik.....	83
15 Bentuk sinyal rekaman a19 periode apnea durasi 15 detik	84
16 Bentuk sinyal rekaman a19 periode non-apnea durasi 15 detik.....	85
17 Bentuk sinyal rekaman b01 periode apnea durasi 15 detik.....	86
18 Bentuk sinyal rekaman b01 periode non-apnea durasi 15 detik.....	87
19 Bentuk sinyal rekaman b03 periode apnea durasi 15 detik.....	88
20 Bentuk sinyal rekaman b03 periode non-apnea durasi 15 detik.....	89
21 Bentuk sinyal rekaman c06 periode non-apnea durasi 15 detik.....	90
22 Bentuk sinyal rekaman c07 periode non-apnea durasi 15 detik.....	91

DAFTAR TABEL

Tabel 2.1 Fitur ECG dan durasi normalnya [9]	9
Tabel 2.2 Beberapa Macam Fungsi Kernel SVM [11][16]	16
Tabel 3.1 Sruktur pada file CSV hasil ekstraksi fitur.....	32
Tabel 3.2 <i>Confusion Matrix</i>	35
Tabel 4.1 Informasi subjek perekaman	39
Tabel 4.2 Contoh data yang diperoleh dari Physionet Database	40
Tabel 4.3 Contoh hasil ekstraksi fitur	44
Tabel 4.4. Tabel <i>Equality of Group Means</i> ANOVA.....	45
Tabel 4.5 Arti nama VAR untuk masing-masing nama fitur	46
Tabel 4.6 Contoh tabel data <i>testing</i> hasil klasifikasi	49
Tabel 4.7 Skenario pengujian terhadap durasi segmentasi data	55
Tabel 4.8 Skenario pembagian data <i>training</i> dan <i>testing</i>	56
Tabel 4.9 Skenario pengujian terhadap kernel SVM yang digunakan ..	58
Tabel 4.10 <i>Structure Matrix</i>	60
Tabel 4.11 Hasil <i>accuracy</i> rata-rata pengujian pertama	62
Tabel 4.12 Hasil <i>accuracy</i> rata-rata pengujian kedua	63
1 Tingkat keberhasilan klasifikasi pada komposisi data Apnea:Non 40:60	92
2 Tingkat keberhasilan klasifikasi pada komposisi data Apnea:Non 50:50	92
3 Tingkat keberhasilan klasifikasi pada komposisi data Apnea:Non 60:40	92
4 Tingkat keberhasilan klasifikasi durasi segmentasi 10 detik	93
5 Tingkat keberhasilan klasifikasi durasi segmentasi 15 detik	93
6 Tingkat keberhasilan klasifikasi durasi segmentasi 20 detik	93
7 Tingkat keberhasilan klasifikasi durasi segmentasi 25 detik	94
8 Tingkat keberhasilan klasifikasi durasi segmentasi 30 detik	94
9 Tingkat keberhasilan klasifikasi menggunakan kernel Linear	95
10 Tingkat keberhasilan klasifikasi menggunakan kernel RBF	95
11 Tingkat keberhasilan klasifikasi menggunakan kernel Polinomial derajat 3.....	95
12 Tingkat keberhasilan klasifikasi menggunakan kernel Sigmoid	96

13	<i>Analysis Case Processing Summary</i>	97
14	<i>Tests of Equality of Group Means</i>	98
15	<i>Wilks' Lambda</i>	99
16	<i>Functions at Group Centroids</i>	99
17	Hasil akurasi tiap percobaan menggunakan 10 fitur	100
18	Hasil akurasi tiap percobaan menggunakan 9 fitur	100
19	Hasil akurasi tiap percobaan menggunakan 8 fitur	101
20	Hasil akurasi tiap percobaan menggunakan 7 fitur	101
21	Hasil akurasi tiap percobaan menggunakan 6 fitur	101
22	Hasil akurasi tiap percobaan menggunakan 5 fitur	102
23	Hasil akurasi tiap percobaan menggunakan 4 fitur	102
24	Hasil akurasi tiap percobaan menggunakan 3 fitur	102
25	Hasil akurasi tiap percobaan menggunakan 2 fitur	103
26	Hasil akurasi tiap percobaan menggunakan 1 fitur	103
27	Hasil akurasi tiap percobaan menggunakan 10 fitur	103
28	Hasil akurasi tiap percobaan menggunakan 9 fitur	104
29	Hasil akurasi tiap percobaan menggunakan 8 fitur	104
30	Hasil akurasi tiap percobaan menggunakan 7 fitur	104
31	Hasil akurasi tiap percobaan menggunakan 6 fitur	105
32	Hasil akurasi tiap percobaan menggunakan 5 fitur	105
33	Hasil akurasi tiap percobaan menggunakan 4 fitur	105
34	Hasil akurasi tiap percobaan menggunakan 3 fitur	106
35	Hasil akurasi tiap percobaan menggunakan 2 fitur	106
36	Hasil akurasi tiap percobaan menggunakan 1 fitur	106

DAFTAR NOMENKLATUR

rr_i	=	RR interval pada iterasi ke-i
i	=	Indeks iterasi
\bar{x}	=	<i>Mean</i>
n	=	Jumlah data
S	=	Standar Deviasi RR interval
R	=	Selisih RR interval yang berdekatan
RMSSD	=	<i>Root mean square of successive differences</i>
SDSD	=	<i>Standard deviation of successive differences</i>
$K(x_i, x_j)$	=	Fungsi kernel linear
x_i	=	Data pada iterasi ke-i
w	=	Normal bidang
b	=	Posisi bidang relatif terhadap pusat koordinat.
y_i	=	Label kelas dari data x_i
$f(x_d)$	=	Fungsi keputusan dari data yang akan diklasifikasikan
s_i	=	<i>Support vector</i>
ns	=	Jumlah <i>support vector</i>
x_d	=	Data yang akan diklasifikasikan
α	=	Nilai bobot setiap titik data
$K(x_i, y_i)$	=	Fungsi Kernel
c	=	Nilai konstanta
d	=	Nilai derajat polinomial
exp	=	Fungsi eksponensial, basis logaritma natural (2,71828 18284 59045 23536 02874 71352)
σ^2	=	Varians
$tanh$	=	Formula tangen hiperbola

Halaman ini sengaja dikosongkan

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Rata-rata manusia harus tidur sekitar 7 jam per hari. Itu artinya manusia tidur sekitar 1/3 tahun dari hidup mereka. Merasa kurang enak badan setelah semalaman tanpa tidur adalah cara tubuh mengingatkan kita bahwa tidur adalah kebutuhan seperti halnya makan, minum dan bernafas. Saat kita tidur, tubuh kita melakukan perbaikan diri. Tanpa tidur, kita tidak dapat menjalankan aktivitas semaksimal mungkin karena kondisi tubuh yang buruk.

Sampai saat ini, ada lebih dari 70 jenis gangguan tidur yang telah ditemukan, termasuk yang paling umum seperti insomnia, *sleep apnea*, narkolepsi dan sindrom kaki gelisah [1]. Hal paling lumrah yang terjadi di masyarakat terkait gangguan tidur adalah mendengkur. Banyak orang menyepelkan tidur dengan mendengkur. Padahal, mendengkur bisa menjadi gejala serius penyakit gangguan pernafasan saat kita tidur, seperti yang terjadi pada beberapa kasus Apnea.

Tidak hanya terjadi di malam hari, gejala apnea juga terjadi pada siang hari, seperti pada gangguan tidur pada umumnya [3]. Apnea membuat penderitanya merasa lelah terus menerus. Pada siang hari penderita kantuk yang berlebihan, gangguan konsentrasi, depresi, bahkan kehilangan ingatan. Hal-hal tersebut umum terjadi karena penderita gangguan tidur tidak memiliki tidur yang berkualitas di malam harinya [3].

Menurut *American Thoracic Society*, beberapa kondisi kardiovaskular bisa terjadi karena OSA yang tidak diobati dengan benar. Penderita OSA akan cenderung memiliki tekanan darah tinggi (hipertensi). Hipertensi itu kemudian memicu penderita terkena serangan jantung, stroke, dan aritmia [3]. Selain itu, karena terhentinya napas secara berulang menyebabkan jantung mengalami stres atau tertekan sehingga dapat menyebabkan pembengkakan [8].

Secara umum, pengujian apnea biasanya menggunakan *polysomnography* (PSG), yang merupakan prosedur standar untuk diagnosis semua gangguan tidur. PSG mencatat seluruh aliran udara napas, gerakan pernafasan, saturasi oksigen, posisi tubuh, *electroencephalogram* (EEG), *electrooculogram* (EOG), *elektromiogram* (EMG), dan *Electrocardiogram* (ECG) [5]. Sebagian besar kasus apnea

tidak terdiagnosis karena ketidaknyamanan pengujian, biaya pengujian yang cukup mahal, serta pengujian yang tidak selalu tersedia [1]. Pengujian PSG memerlukan biaya yang mahal karena hanya dilakukan pada laboratorium khusus di rumah sakit dan mengharuskan mesin, dokter, teknisi dan staf untuk bekerja semalaman. Pengujian menjadi kurang nyaman karena pasien harus bermalam di ruang pengujian saat pemeriksaan dilakukan. Terlebih, ketersediaan pengujian PSG terbatas karena ruang pengujian yang ada selalu penuh bahkan pasien yang mendaftar harus masuk dalam *waiting list* dan menunggu selama berbulan-bulan. Untuk itu, perlu adanya sistem deteksi yang lebih sederhana untuk penanganan yang lebih cepat daripada PSG. Saat ini semakin banyak teknik-teknik terkait dengan biometri yang diteliti dan dikembangkan untuk diagnosa kesehatan, salah satunya ECG.

1.2 Permasalahan

Belum adanya sistem yang dapat mengklasifikasikan sinyal ECG sehingga dapat diketahui terjadi OSA atau tidak pada pasien tersebut.

1.3 Tujuan

Tujuan dari penelitian ini adalah untuk melakukan klasifikasi sinyal ECG yang terjadi apnea dan non-apnea menggunakan fitur-fitur yang diekstrak dari RR Interval sinyal ECG pasien sehingga dapat diketahui apakah apnea terjadi atau tidak. Output yang diharapkan sistem yang dapat mengklasifikasikan sinyal ECG ke dalam 2 kelompok yakni apnea dan non-apnea. Manfaat untuk ke depannya, diharapkan dapat menjadi referensi untuk penelitian berikutnya dan dapat dikembangkan lebih lanjut untuk dapat memantau gangguan tidur pasien dari sinyal ECG secara *real-time*.

1.4 Batasan Masalah

Batasan masalah pada penelitian ini adalah:

1. Data sinyal ECG *single-channel* diperoleh dari *PhysioNet Database* [2].
2. Pada tahap *pre-processing* dilakukan penghitungan RR Interval.
3. Fitur yang digunakan berupa fitur statistik yang didapat dengan menghitung RR Interval sinyal dengan total sebanyak 11 fitur.
4. Metode yang digunakan untuk klasifikasi adalah *Support Vector Machine* dengan menggunakan kernel linear.
5. Klasifikasi yang dilakukan adalah apnea dan non-apnea.

1.5 Sistematika Penulisan

Laporan penelitian Tugas akhir ini tersusun dalam sistematika dan terstruktur sehingga mudah dipahami dan dipelajari oleh pembaca maupun seseorang yang ingin melanjutkan penelitian ini. Alur sistematika penulisan laporan penelitian ini yaitu :

1. BAB I Pendahuluan

Bab ini berisi uraian tentang latar belakang permasalahan, penegasan dan alasan pemilihan judul, sistematika laporan, tujuan dan metodologi penelitian.

2. BAB II Dasar Teori

Pada bab ini berisi tentang uraian secara sistematis teori-teori yang berhubungan dengan permasalahan yang dibahas pada penelitian ini. Teori-teori ini digunakan sebagai dasar dalam penelitian, yaitu informasi terkait sinyal ECG, penyakit apnea, dan teori-teori penunjang lainnya.

3. BAB III Perancangan Sistem dan Implementasi

Bab ini berisi tentang penjelasan-penjelasan terkait eksperimen yang akan dilakukan dan langkah-langkah data diolah hingga menghasilkan data klasifikasi yang diinginkan. Guna mendukung itu digunakan blok diagram atau *work flow* agar sistem yang akan dibuat dapat terlihat dan mudah dibaca untuk implementasi pada pelaksanaan tugas akhir.

4. BAB IV Pengujian dan Analisa

Bab ini menjelaskan tentang pengujian eksperimen yang dilakukan terhadap data dan analisisnya. Beberapa ekstraksi fitur sinyal akan ditunjukkan hasilnya pada bab ini dan dilakukan analisa terhadap hasil *training* data yang telah didapat dari hasil ekstraksi fitur.

5. BAB V Penutup

Bab ini merupakan penutup yang berisi kesimpulan yang diambil dari penelitian dan pengujian yang telah dilakukan. Saran dan kritik yang membangun untuk pengembangan lebih lanjut juga dituliskan pada bab ini.

Halaman ini sengaja dikosongkan

BAB 2

TINJAUAN PUSTAKA

2.1 Sleep Apnea Syndrome

Sleep apnea syndrome adalah suatu sindrom dengan ditemukannya kejadian apnea (penghentian aliran udara selama 10 detik sehingga menyebabkan 2-4% penurunan saturasi oksigen) atau hipopnea (penurunan aliran udara paling sedikit 30-50% sehingga menyebabkan penurunan saturasi oksigen) pada saat tidur [5]. Apnea dapat disebabkan kelainan sentral, obstruktif jalan nafas, atau campuran. Obstruktif apnea adalah berhentinya aliran udara pada hidung dan mulut walaupun dengan usaha nafas, Sentral apnea adalah penghentian pernafasan yang tidak disertai dengan usaha bernafas akibat tidak adanya rangsangan nafas, sedangkan apnea campuran adalah campuran antara keduanya.

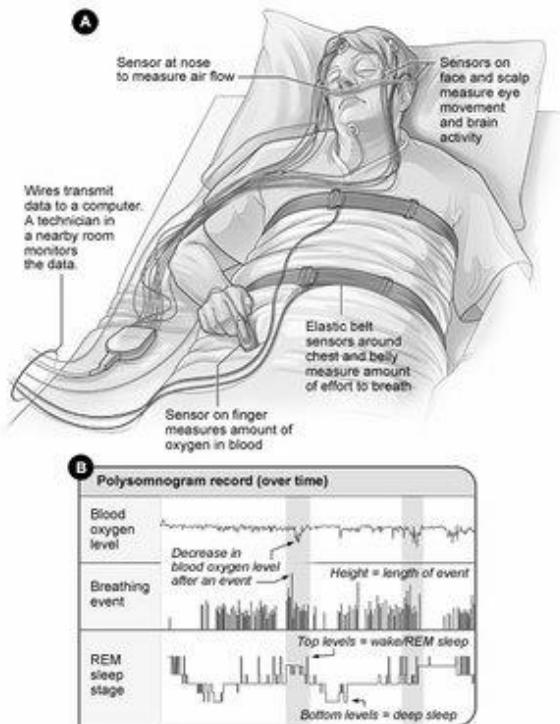
Identifikasi apnea didasarkan pada nilai AHI (Apnea-Hypopnea Indeks). Angka AH1 kurang dari 5 dianggap sebagai normal; AH1 5-15 kali per jam sebagai *obstructive sleep apnea-hypopnea syndrome* (OSAHS) ringan; AH1 15-30 kali per jam sebagai OSAHS sedang; dan AH1 diatas 30 per jam dianggap sebagai OSAHS berat.

2.1.1 Obstructive Sleep Apnea

Obstructive sleep apnea atau selanjutnya disebut OSA, merupakan bentuk umum *sleep-disordered breathing* (SDB) atau gangguan napas ketika tidur. OSA terjadi karena sumbatan total atau sebagian jalan napas yang terjadi secara berulang pada saat tidur sehingga menyebabkan aliran udara ke paru menjadi terhambat. Kejadian apnea terjadi selama 10-60 detik dan OSA yang ekstrim dapat terjadi berulang setiap 30 detik [4][5]. Berhentinya aliran udara ini berakibat meningkatnya tekanan pada jantung dan tekanan darah, sehingga penderita OSA sulit memasuki fase *rapid eye movement* (REM). Kelainan anatomi saluran nafas merupakan faktor utama terjadinya OSA. Selain itu, terdapat faktor lain yang mempengaruhi seperti umur, jenis kelamin, obesitas, dan gaya hidup [5].

Pada umumnya deteksi penyakit OSA dapat dilakukan dengan menggunakan test polisomnografi (PSG). PSG merupakan alat yang dapat merekam fungsi-fungsi tubuh selama tidur, seperti gerakan bola mata, gelombang otak, regangan otot, aliran udara nafas, gerakan nafas, kerja jantung, dan kadar oksigen. Karena banyaknya sensor yang dipasang untuk dapat mendeteksi OSA, maka dirasa pasien tidak nyaman untuk

melakukan deteksi OSA menggunakan polisomnografi [1]. Ilustrasi pengujian polisomnografi dapat dilihat pada Gambar 2.1



Gambar 2.1 Proses pengujian polisomnografi [www.alphasleeplab.com]

2.2 Hubungan OSA dengan Perubahan ECG

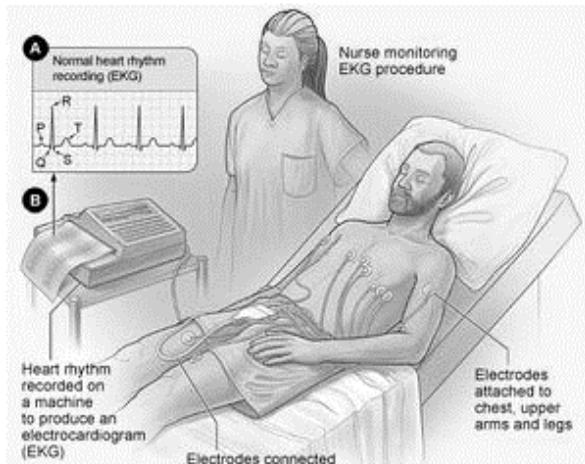
Beberapa penelitian menunjukkan bahwa terdapat hubungan yang kuat antara OSA dengan penyakit kardiovaskuler. Beberapa kondisi kardiovaskular bisa terjadi karena OSA yang tidak diobati dengan benar. Penderita OSA akan cenderung memiliki tekanan darah tinggi (hipertensi) dan sulit untuk mengendalikan tekanan darahnya. Dari orang dengan hipertensi, sekitar 30% memiliki OSA. Orang yang menderita OSA, ada kemungkinan 50% juga memiliki hipertensi [8].

Selain itu, masalah dengan ritme jantung juga bisa terjadi seperti fibrilasi atrium (tidak beraturan denyut jantung), atau bradikardia (denyut jantung lambat). Pasien dengan OSA berat empat kali lebih mungkin mengalami fibrilasi atrium dibandingkan mereka yang tidak memiliki OSA [8]. Gangguan fibrilasi atrium dan bradikardia itulah yang kemudian menjadikan pasien penderita OSA memiliki denyut jantung yang tidak teratur, terutama pada saat OSA terjadi. Saat OSA terjadi, denyut jantung penderita dapat melambat dan kemudian langsung berubah menjadi lebih cepat dari denyut normal dalam periode yang sangat singkat [8]. Denyut jantung yang terjadi pada pasien selanjutnya dapat direpresentasikan dengan gambaran aktivitas listrik yang biasa disebut rekaman *electrocardiogram* (ECG).

2.3 *Electrocardiogram* (ECG)

2.3.1 Pengertian *Electrocardiogram*

Electrocardiogram atau selanjutnya disebut ECG, adalah representasi atau gambaran aktivitas listrik jantung yang direkam melalui elektrode yang diletakkan pada permukaan tubuh [7]. Proses memperoleh sinyal ECG dapat dilakukan dengan cara meletakkan elektrode langsung pada permukaan kulit di area jantung yang disebut *electrocardiography*, dengan visualisasi seperti pada Gambar 2.2.



Gambar 2.2 Proses *electrocardiography* [3].

Jantung menghasilkan impuls listrik kecil yang menyebar melalui otot jantung. Impuls ini bisa dideteksi oleh mesin ECG. Mesin ECG mencatat aktivitas listrik jantung dan menampilkan data tersebut sebagai jejak di atas kertas atau yang kita kenal sebagai electrocardiogram. Data ini kemudian ditafsirkan oleh para praktisi medis. ECG membantu untuk menemukan penyebab gejala atau nyeri dada dan juga membantu mendeteksi irama jantung abnormal atau kelainan jantung.

Sinyal ECG memiliki frekuensi antara 0,05 Hz sampai dengan 100Hz. Amplitudo sinyal ECG berkisar antara 0,02V sampai 5V [6]. ECG dari jantung yang sehat dan normal memiliki bentuk yang khas. Ketidakteraturan dalam irama jantung atau kerusakan pada otot jantung bisa mengubah aktivitas listrik jantung, sehingga bentuk ECG akan berubah. Gangguan jantung yang bisa dideteksi dengan menggunakan ECG, meliputi irama jantung abnormal, serangan jantung, dan pembesaran jantung. Namun seiring dengan perkembangan teknologi dan penelitian, sinyal jantung saat ini mampu menjadi data untuk banyak gangguan kesehatan pada tubuh manusia, termasuk gangguan nafas saat tidur karena nafas yang tidak teratur juga akan menyebabkan irama jantung menjadi tidak normal. Pada pasien OSA, electrocardiogram pun akan berubah karena pasien mengalami henti napas saat tidur [12].

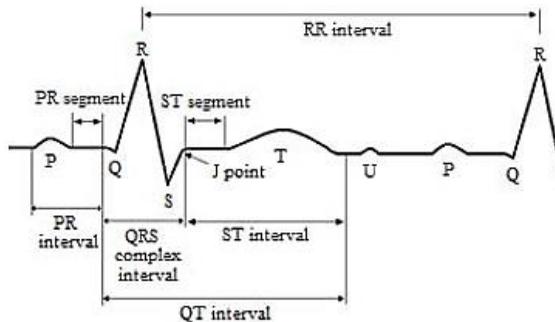
Klasifikasi sinyal ECG menggunakan teknik pembelajaran mesin dapat memberikan masukan yang substansial kepada dokter untuk memastikan diagnosis [9]. Klasifikasi dan deteksi dapat membantu dalam mengidentifikasi kelainan yang ada pada sinyal ECG pasien. Setelah mengidentifikasi kelainan, penyakit dapat dideteksi dan penanganan pasien yang lebih baik dapat dilakukan. Klasifikasi ECG yang akurat memberikan informasi yang cukup untuk mendeteksi penyakit dan membantu dokter dalam menemukan terapi pengobatan terbaik untuk pasien.

2.3.2 RR Interval

Satu sinyal ECG utuh terdiri dari beberapa irama/detakan. Setiap detakan berisi gelombang P, kompleks QRS, dan gelombang T. Setiap puncak (P, Q, R, S, T, dan U), interval (PR, RR, QRS, ST, dan QT) dan segmen (PR dan ST) sinyal ECG memiliki nilai amplitudo atau durasi normal. Puncak, interval, dan segmen inilah yang disebut fitur ECG. Gambar 2.3 menunjukkan fitur-fitur untuk satu siklus jantung ECG, yang kemudian dijelaskan pada Tabel 2.1. Tabel 2.1 menyajikan fitur ECG beserta deskripsi dan durasinya.

Tabel 2.1 Fitur ECG dan durasi normalnya [9]

Fitur	Deskripsi	Durasi
RR	Jarak antara gelombang R dan gelombang R selanjutnya	600-1200 ms
P	Bukit kecil pertama pada sinyal ECG	80ms
PR	Ukuran dari permulaan gelombang P hingga permulaan kompleks QRS	120-200ms
QRS	Umumnya dimulai dengan titik lembah Q, titik bukit besar R dan diakhiri dengan titik lembah S.	80-120ms
PR Segment	Menghubungkan gelombang P dan kompleks QRS	50-120ms
J-Point	Titik dimana kompleks QRS selesai dan ST segment dimulai	-
ST Segment	Menghubungkan kompleks QRS dengan gelombang T	80-120ms
T	Umumnya berbentuk gelombang bukit yang rendah	160ms
ST	Ukuran dari J-point hingga akhir gelombang T	320ms
QT	Ukuran dari permulaan kompleks QRS hingga akhir gelombang T	420ms
U	Biasanya memiliki amplitudo rendah dan sering kali tidak ada sama sekali	Tidak disebutkan



Gambar 2.3 Siklus normal sinyal ECG [9].

RR interval juga dikenal dengan istilah *normal-to-normal* (NN) interval. Dari Gambar 2.3, terlihat bahwa RR interval merupakan jarak antar puncak gelombang R ke puncak R berikutnya pada kompleks QRS. Dalam sinyal ECG, keteraturan dari interval tersebut sangat diperhatikan. RR Interval bermanfaat dalam mengetahui irama jantung atau *heart rate* (HR) yang selanjutnya digunakan untuk mengukur *Heart Rate Variability* (HVR) [13]. Waktu RR Interval pada detak sinyal ECG yang dihasilkan mengikuti Persamaan 2.1.

$$rr_i = r_{(i+1)} - r_i, \quad i = 1, 2, \dots, n - 1 \quad (2.1)$$

2.4 Heart Rate Variability

Pada kondisi normal tanpa ada pengaruh dari luar, besarnya *heart rate* atau detakan jantung dalam keadaan istirahat adalah sekitar 90 bpm (*beat per minute*). Artinya, jantung berdetak sebanyak 90 kali setiap menitnya. Nilai *heart rate* ini dapat mengalami perubahan, baik perubahan yang bersifat fisiologis maupun patofisiologis. Perubahan nilai *heart rate* yang bersifat fisiologis misalnya adalah karena adanya pengaruh sistem syaraf otonomik, volume darah yang kembali ke jantung (*venous return*), respirasi, dan lain sebagainya. Perubahan nilai *heart rate* yang bersifat patofisiologis misalnya adalah yang disebabkan oleh adanya penyakit aritmia [10].

Perubahan secara terus menerus pada nilai *heart rate* tersebut menunjukkan adanya *heart rate variability* (HRV). HRV ini menggambarkan variasi antar *inter-beat-interval* (IBI) berturut-turut dari detak jantung.

2.4.1 Pengukuran HRV

HRV dapat dievaluasi atau dianalisis dengan sejumlah metode. Secara umum, metode tersebut dapat dikelompokkan menjadi 2 kelompok besar yakni dengan domain waktu (*time domain*) atau domain frekuensi (*frequency domain*).

1. Metode domain waktu

Metode ini adalah metode yang paling sederhana dalam pengukuran HRV [13]. Metode ini didasarkan pada perhitungan interval *beat-to-beat* atau NN. Menurut [13], dalam metode ini dapat dibedakan menjadi 2 pengukuran, yakni metode statistik dan metode geometris. Pada metode statistik dapat dibagi menjadi dua kelas, yaitu yang berasal dari pengukuran langsung NN interval atau RR Interval dan yang berasal dari perbedaan

antara NN interval yang berdekatan. Pengukuran-pengukuran tersebut dilakukan untuk mendapat berbagai variabel seperti :

a. Mean

Perhitungan nilai rata-rata dari RR interval. Diperoleh dari hasil pembagian dari penjumlahan seluruh interval dengan banyaknya RR interval, yang sebelumnya telah didapat dari Persamaan 2.1. Secara matematis, Mean dapat dirumuskan seperti pada Persamaan 2.2.

$$\bar{x} = \frac{rr_1+rr_2+rr_3+\dots+rr_i}{n} \quad (2.2)$$

Dimana, \bar{x} = Mean

i = indeks iterasi

rr_i = RR interval ke- i

n = Banyak data

b. Median

Nilai tengah RR interval.

c. SDNN

Standar deviasi RR interval, dirumuskan seperti pada Persamaan 2.3.

$$S = \sqrt{\frac{\sum_{i=1}^n (rr_i - \bar{x})^2}{n}} \quad (2.3)$$

Dimana, S = Standar deviasi RR interval (SDNN)

d. *Root mean square of successive differences* (RMSSD)

Didefinisikan sebagai akar kuadrat dari rata-rata selisih antara RR interval yang berdekatan. Secara matematis dapat dirumuskan seperti pada Persamaan 2.4.

$$R = \sqrt{rr_{i+1} - rr_i} \quad (2.4)$$

Dimana, R = selisih antara RR interval yang berdekatan.

Dengan mengacu pada Persamaan 2.4, maka didapatkan rumus RMSSD seperti pada persamaan 2.5.

$$RMSSD = \sqrt{\frac{\sum_{i=1}^n (R_i)^2}{n}} \quad (2.5)$$

Dimana, *RMSSD* = *Root mean square of successive differences*

- e. *Standard deviation of successive differences (SDSD)*, seperti pada Persamaan 2.6.

$$SDSD = \sqrt{\frac{\sum_{i=1}^n (R_i - \bar{x})^2}{n}} \quad (2.6)$$

Dimana, *SDSD* = *Standard deviation of successive differences*

- f. *NNxx*
Jumlah pasangan RR interval yang berdekatan dengan selisih lebih dari xx ms.
- g. *pNNxx*
(Proporsi) *NNxx* dibagi dengan jumlah RR Interval. *pNNxx* biasanya dinyatakan dalam persen.
- h. *Kisaran Interkuartil*
Perbedaan antara persentil ke-75 dan ke-25 dari distribusi RR interval
- i. *Mean Absolute Deviation (MAD)*
Didefinisikan sebagai rata-rata dari nilai mutlak yang diperoleh dari pengurangan rata-rata RR interval dari semua nilai RR interval. *MAD* dapat ditulis secara matematis seperti pada Persamaan 2.7.

$$MAD = \frac{\sum_{i=1}^n |x_i - \bar{x}|}{n} \quad (2.7)$$

Dimana, *MAD* = *Mean Absolute Deviation*

Rangkaian RR interval juga dapat diubah menjadi pola geometrik dengan :

- a. Pola distribusi kepadatan sampel dari durasi RR interval.
- b. Distribusi kepadatan sampel dari selisih RR interval yang berdekatan.
- c. *Scatterplot* setiap RR interval dengan RR interval sebelumnya, atau disebut plot Lorenz.

Rumus sederhana kemudian digunakan untuk menilai variabilitas dari sifat geometrik dan/atau grafis dari pola yang dihasilkan [13].

2. Metode domain frekuensi

Metode domain frekuensi menetapkan band-band frekuensi dan kemudian menghitung jumlah RR Interval yang sesuai dengan masing-masing band. Band-band tersebut biasanya meliputi :

- a. *High frequency* (HF) dari 0,15-0,4 Hz.
- b. *Low frequency* (LF) dari 0,04 sampai 0,15 Hz.
- c. *Very low frequency* (VLF) dari 0,0033 sampai 0,04 Hz.

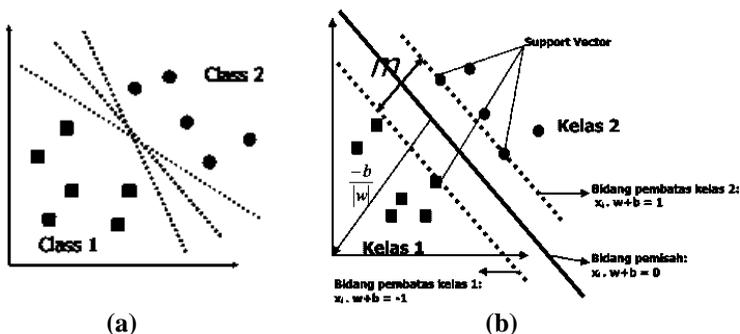
2.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) dikembangkan oleh Boser, Guyon, Vapnik, dan pertama kali dipresentasikan pada tahun 1992 di *Annual Workshop on Computational Learning Theory*. Konsep dasar SVM sebenarnya merupakan kombinasi harmonis dari teori-teori komputasi yang telah ada puluhan tahun sebelumnya, seperti *margin hyperplane*, kernel diperkenalkan oleh Aronszajn tahun 1950, dan demikian juga dengan konsep-konsep pendukung yang lain.

Konsep SVM dapat dijelaskan secara sederhana sebagai usaha mencari *hyperplane* terbaik yang berfungsi sebagai pemisah dua buah *class* pada *input space*. Masalah klasifikasi dapat diterjemahkan dengan usaha menemukan garis (*hyperplane*) yang memisahkan antara kedua kelompok. Usaha untuk mencari lokasi *hyperplane* ini merupakan inti dari proses pembelajaran pada SVM [11].

2.5.1 Klasifikasi Linear pada SVM

Klasifikasi Linear merupakan data yang dapat dipisahkan secara linier. Misalkan $\{x_1, \dots, x_n\}$ adalah dataset untuk $i = 1, 2, \dots, n$, yang mana n adalah banyaknya data, dan $y_i \in \{+1, -1\}$ adalah label kelas dari data x_i sebagai *class 1* dan *class 2*, seperti yang ditunjukkan pada Gambar 2.4.



Gambar 2.4 Konsep Kerja SVM dalam memisahkan 2 *class* [14].

Pada Gambar 2.4(a) memperlihatkan beberapa data yang merupakan anggota dari dua buah *class* : +1 dan -1. Data yang tergabung pada *class* -1 disimbolkan dengan kotak, sedangkan data pada *class* +1, disimbolkan dengan lingkaran.

Hyperplane pemisah terbaik antara kedua *class* dapat ditemukan dengan mengukur *margin hyperplane* tersebut, dan mencari titik maksimalnya. Margin adalah jarak antara *hyperplane* tersebut dengan data terdekat dari masing-masing *class* yang dilambangkan dengan m . Adapun data yang berada pada bidang pembatas ini disebut *support vector*. Garis solid pada Gambar 2.4(b) menunjukkan *hyperplane* yang terbaik, yaitu yang terletak tepat pada tengah-tengah kedua *class*.

Bidang pembatas pertama membatasi kelas pertama sedangkan bidang pembatas kedua membatasi kelas kedua, sehingga diperoleh Persamaan 2.8.

$$x_i \cdot w + b \geq -1 \text{ untuk } y_i = +1 \quad (2.8)$$

$$x_i \cdot w + b \leq -1 \text{ untuk } y_i = -1$$

Dimana, w = normal bidang

b = posisi bidang relatif terhadap pusat koordinat.

y_i = label kelas dari data x_i

Margin terbesar dapat ditemukan dengan memaksimalkan nilai jarak antara *hyperplane* dan titik terdekatnya, yaitu $1/|w|$.

Pencarian bidang pemisah terbaik dengan nilai margin terbesar dapat dirumuskan menjadi masalah optimasi konstrain dengan *Quadratic Programming (QP) Problem*. Permasalahan QP ini dapat dipecahkan dengan berbagai teknik komputasi, diantaranya dengan *Lagrange Multiplier*. Setelah solusi permasalahan *quadratic programming* ditemukan, maka kelas dari data pengujian x dapat ditentukan berdasarkan nilai dari fungsi keputusan pada Persamaan 2.9.

$$f(x_d) = a_0 + \sum_{i=1}^{ns} \alpha_i y_i s_i \cdot x_d + b, \quad (2.9)$$

Dimana, s_i = *support vector*

ns = jumlah *support vector*

x_d = data yang akan diklasifikasikan

α = nilai bobot setiap titik data

2.5.2 Klasifikasi Non-linear pada SVM

Pada mulanya teknik *machine learning* dikembangkan dengan asumsi kelinearan. Sehingga algoritma yang dihasilkan terbatas untuk kasus-kasus yang linear saja. Namun, pada umumnya masalah dalam dunia nyata jarang yang bersifat *linear separable*. Kebanyakan bersifat non linear. Untuk menyelesaikan problem non linear, SVM dimodifikasi dengan memasukkan fungsi Kernel [11]. Beberapa fungsi berikut disajikan secara singkat pada Tabel 2.2.

Dalam non linear SVM, pertama-tama data x dipetakan oleh fungsi $\Phi(x)$ ke ruang vektor yang berdimensi lebih tinggi. Pada ruang vektor yang baru ini, *hyperplane* yang memisahkan kedua class tersebut dapat dikonstruksikan [11].

Tabel 2.2 Beberapa Macam Fungsi Kernel SVM [11][16]

No	Nama Kernel	Definisi Fungsi
1	<i>Linier</i>	$K(x_i, y_i) = x_i \cdot y_i$
2	<i>Polinomial of degree d</i>	$K(x_i, y_i) = (x_i \cdot y_i)^d$
3	<i>Polinomial of degree up to d</i>	$K(x_i, y_i) = (x_i \cdot y_i + c)^d$
4	<i>Gaussian Radial Basis Function (RBF)</i>	$K(x_i, y_i) = \exp\left(-\frac{ x_i - y_i ^2}{2\sigma^2}\right)$
5	<i>Sigmoid</i>	$K(x_i, y_i) = \tanh(\alpha x_i y_i + c)$

Dimana, $K(x_i, y_i)$ = Fungsi Kernel
 c = Nilai konstanta
 d = Nilai derajat polinomial
 \exp = Fungsi eksponensial, basis logaritma natural
 (2,71828 18284 59045 23536 02874 71352)
 σ^2 = Varians
 \tanh = Formula tangen hiperbola

Kernel Linier digunakan ketika data yang akan diklasifikasi dapat terpisah dengan sebuah garis/hyperplane. Sedangkan Kernel Non-Linier digunakan ketika data hanya dapat dipisahkan dengan garis lengkung atau sebuah bidang kurva pada ruang dimensi yang lebih tinggi [16]. Pilihan kernel tergantung pada masalah yang akan dimodelkan. Kernel polinomial cocok untuk digunakan pada data pelatihan yang dinormalisasi, dan memungkinkan untuk memodelkan fitur dengan bentuk polinomial. *Radial Basis Function* dapat digunakan ketika untuk memilih batas berupa lingkaran [1].

Fungsi kernel $K(x_i, y_i)$ yang mendefinisikan secara implisit transformasi Φ karena umumnya transformasi Φ ini tidak diketahui, dan sangat sulit untuk difahami secara mudah. Hal ini disebut sebagai *Kernel Trick*, yang dirumuskan seperti pada Persamaan 2.10.

$$K(x_i, y_i) = \Phi(x_i) \cdot \Phi(y_i) \quad (2.10)$$

Kernel trick memberikan berbagai kemudahan, karena dalam proses pembelajaran SVM, untuk menentukan *support vector*, maka cukup dengan mengetahui fungsi kernel yang dipakai, dan tidak perlu mengetahui wujud dari fungsi non-linear [11].

2.5.3 Karakteristik SVM

SVM memiliki beberapa karakteristik sebagai berikut :

1. Secara prinsip SVM adalah *linear classifier*
2. *Pattern recognition* dilakukan dengan mentransformasikan data pada *input space* ke ruang yang berdimensi lebih tinggi, dan optimisasi dilakukan pada ruang vector yang baru tersebut. Hal ini membedakan SVM dari solusi *pattern recognition* pada umumnya, yang melakukan optimisasi parameter pada ruang hasil transformasi yang berdimensi lebih rendah daripada dimensi *input space*.
3. Menerapkan strategi *Structural Risk Minimization* (SRM)
4. Prinsip kerja SVM pada dasarnya hanya mampu menangani klasifikasi dua *class*

2.5.4 Kelebihan dan Kekurangan SVM

Dalam menentukan metode dan solusi untuk menyelesaikan suatu permasalahan, kelebihan dan kelemahan masing-masing metode adalah hal yang harus diperhatikan. Hal ini bertujuan agar metode yang dipilih adalah metode yang tepat dengan memperhatikan karakteristik data yang diolah. Dalam hal SVM, walaupun berbagai studi telah menunjukkan kelebihan metode SVM dibandingkan metode konvensional lainnya, SVM juga memiliki berbagai kelemahan [11]. Kelebihan dan kelemahan SVM antara lain adalah sebagai berikut menurut [11] :

1. Kelebihan SVM
 - a. Generalisasi
Generalisasi didefinisikan sebagai kemampuan suatu metode *machine learning* untuk mengklasifikasikan suatu *pattern* atau pola tertentu, yang tidak termasuk data yang dipakai dalam fase pembelajaran metode itu. Strategi pembelajaran pada *neural network* dan umumnya metode *machine learning* difokuskan pada usaha untuk meminimalkan *error* pada *training-set* atau data yang digunakan dalam fase pembelajaran. Strategi ini disebut *Empirical Risk Minimization* (ERM). Selain meminimalkan *error* pada *training-set*, metode SVM juga

dapat meminimalkan faktor kedua. Strategi ini disebut *Structural Risk Minimization* (SRM), dan dalam SVM diwujudkan dengan memilih *hyperplane* dengan margin terbesar. Berbagai studi empiris menunjukkan bahwa pendekatan SRM pada SVM memberikan *error* generalisasi yang lebih kecil daripada yang diperoleh dari strategi ERM pada *neural network* maupun metode yang lain [11].

- b. *Curse of dimensionality*
Curse of dimensionality didefinisikan sebagai masalah yang dihadapi suatu metode *pattern recognition* dalam mengestimasi parameter dikarenakan jumlah sampel data yang relatif lebih sedikit dibandingkan dengan dimensional ruang vektor data tersebut. Semakin tinggi dimensi dari ruang vektor informasi yang diolah, semakin besar pula jumlah data yang dibutuhkan dalam proses pembelajaran. Pada kenyataannya, seringkali terjadi data yang diolah berjumlah terbatas dan untuk mengumpulkan data yang lebih banyak sangat sulit dilakukan karena adanya berbagai kendala baik besarnya biaya maupun kesulitan teknis. Vapnik V.N, dalam penelitiannya tentang teori pembelajaran statistik membuktikan bahwa tingkat generalisasi yang diperoleh SVM tidak dipengaruhi oleh dimensi dari input vector [11]. Hal ini merupakan alasan utama mengapa SVM dijadikan salah satu metode yang tepat digunakan untuk memecahkan masalah berdimensi tinggi namun dalam keterbatasan sampel data yang ada.
- c. Landasan Teori
Sebagai metode yang berbasis statistik, SVM dinilai memiliki landasan teori yang dapat dianalisa dengan jelas [11].
- d. *Feasibility*
SVM dapat diimplementasikan dengan relatif mudah, karena proses penentuan *support vector* dapat dirumuskan dalam QP *problem* [11].

2. Kelemahan SVM

- a. Sulit dipakai dalam masalah berskala besar. Skala besar dalam hal ini merujuk pada jumlah sampel yang diolah berjumlah sangat banyak [11].
- b. SVM secara teoritik dikembangkan untuk melakukan klasifikasi masalah menjadi dua kelas. Namun saat ini SVM semakin dikembangkan dan telah dimodifikasi agar dapat menyelesaikan masalah dengan lebih dari dua kelas, meskipun pengembangan tersebut masih memiliki banyak kekurangan [11].

2.5.5 SVM pada Python

SVM yang merupakan salah satu pembelajaran mesin dapat dilakukan pada banyak bahasa pemrograman, termasuk Python. Dalam Python, pembelajaran mesin dapat dilakukan dengan memanggil library *machine learning* yang disebut Scikit-learn. Scikit-learn adalah library *open source* pembelajaran mesin gratis pada Python yang dibuat oleh David Cournapeau pada Juni 2007. David Cournapeau merupakan seorang *data scientist* yang bekerja di perusahaan *software* di Jepang. Scikit-learn pada awalnya ia ciptakan saat mengadakan project Google Summer of Code. Nama Scikit itu sendiri merupakan akronim dari SciPy Toolkit, pengembang terpisah yang mendistribusikan SciPy, library umum Python untuk memudahkan menggunakan dalam bidang sains (*scientific*). Seiring berjalannya waktu, library tersebut terus dikembangkan oleh ilmuwan-ilmuwan lainnya. Scikit-learn memiliki berbagai fitur seperti klasifikasi, regresi, dan algoritma *clustering* termasuk SVM, random forests, gradient boosting, k-means, dan DBSCAN. Scikit-learn didesain untuk dapat digunakan beriringan dengan library numerik dan *scientific* seperti NumPy dan SciPy..

Fitur, metode, formula matematika yang ada pada Scikit-learn sangat beragam sesuai dengan kebutuhan penggunaannya. Dalam SVM, klasifikasi dapat dilakukan dengan metode SVC, NuSVC, dan LinearSVC. Setiap metode tersebut berbeda formula matematika yang digunakan. SVC dan NuSVC adalah metode yang hampir sama, namun sedikit berbeda dalam penetapan parameter dan formula matematikanya. Sedangkan LinearSVC adalah implementasi klasifikasi SVM yang digunakan untuk kasus klasifikasi linear saja [www.scikit-learn.org].

2.6 Pengukuran Unjuk Kerja

Ilmuwan, spesialis, bahkan praktisi menjumpai berbagai kasus yang berbeda setiap harinya. Mereka harus menggunakan alat diagnosa tambahan seperti tes laboratorium dan studi pencitraan untuk dapat menangani kasus-kasus tersebut [15]. Seiring dengan berkembangnya ilmu pengetahuan, saat ini hasil akurasi dari tes-tes tersebut dapat diketahui. Tes dengan kesalahan terkecil dan akurasi paling tinggi lah yang lebih diinginkan oleh masyarakat. Oleh karena itu, tes dengan akurasi 100% harus yang menjadi pilihan utama.

Namun, hal ini tidak terjadi dalam kehidupan nyata karena keakuratan suatu tes bervariasi untuk kasus yang berbeda dan dalam situasi yang berbeda. Dalam dunia kedokteran, dapat terjadi akurasi yang tinggi pada pasien risiko rendah dan akurasi rendah pada penyakit dengan risiko tinggi. Karakteristik dari tes yang mencerminkan kemampuan tersebut dikenal dengan akurasi (*Accuracy*), sensitivitas (*sensitivity*), spesifisitas (*specificity*), nilai prediksi positif dan negatif dan rasio kemungkinan positif dan negatif [15]. Berikut adalah istilah-istilah yang digunakan dalam menghitung keakuratan suatu tes beserta pengertian dan perhitungan matematis masing-masing istilah :

- a. *Positive* :
Umumnya, positive merujuk pada pasien pengidap penyakit dari kasus yang ada.
- b. *Negative* :
Merujuk pada orang yang sehat (bukan pengidap penyakit).
- c. *True Positive (TN)*:
Jumlah kasus dengan benar diidentifikasi sebagai pasien.
- d. *False positive (FP)* :
Jumlah kasus salah diidentifikasi sebagai pasien.
- e. *True negative (TN)* :
jumlah kasus yang diidentifikasi dengan benar sebagai sehat.
- f. *False negative (FN)* :
Jumlah kasus yang salah diidentifikasi sebagai sehat.
- g. *Accuracy* :
Kemampuannya untuk membedakan pasien dan kasus yang sehat dengan benar. Untuk mendapatkan *Accuracy*, kita harus menghitung proporsi *true positive* dan *true negative* dalam semua

kasus yang dievaluasi. Secara matematis, ini dapat dinyatakan sebagai Persamaan 2.11.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.11)$$

h. *Sensitivity* :

Kemampuannya untuk menentukan kasus pasien dengan benar. Untuk menghitungnya, kita harus menghitung proporsi positif yang benar dalam kasus pasien. Secara matematis, ini dapat dinyatakan sebagai Persamaan 2.12.

$$Sensitivity = \frac{TP}{TP + FN} \quad (2.12)$$

i. *Specificity* :

Kemampuannya untuk menentukan kasus yang sehat dengan benar, dengan menghitung proporsi negatif yang sebenarnya dalam kasus yang sehat. Secara matematis, ini dapat dinyatakan sebagai Persamaan 2.13.

$$Specificity = \frac{TN}{TN + FP} \quad (2.13)$$

Sensitivity dan *Specificity* dapat dikatakan sebagai dua nilai yang saling berlawanan. Ini artinya apabila nilai *Sensitivity* meningkat, maka nilai *Specificity* akan berkurang. Begitu juga sebaliknya, apabila nilai *Sensitivity* berkurang maka nilai *Specificity* akan bertambah.

Halaman ini sengaja dikosongkan

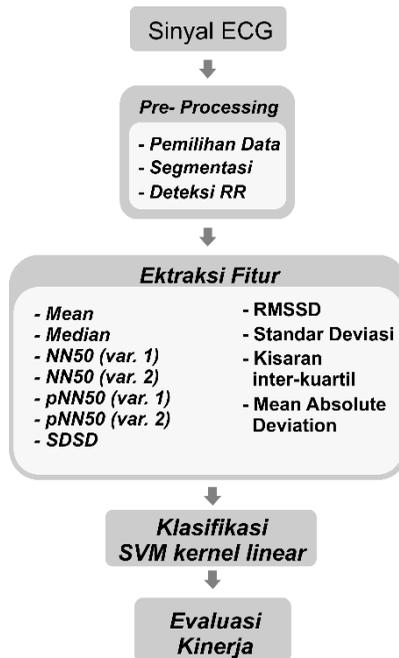
BAB 3

DESAIN DAN IMPLEMENTASI SISTEM

Penelitian ini dilaksanakan sesuai dengan desain sistem berikut dengan implementasinya. Desain sistem merupakan konsep dari pembuatan dan perancangan infrastruktur dan kemudian diwujudkan dalam bentuk blok-blok alur yang harus dikerjakan. Pada bagian implementasi merupakan pelaksanaan teknis untuk setiap blok pada desain sistem.

3.1 Desain Sistem

Penelitian ini bertujuan untuk melakukan deteksi OSA menggunakan fitur-fitur yang diekstrak dari RR Interval sinyal ECG. Adapun tahapan proses yang dikerjakan dapat dilihat pada Gambar 3.1.



Gambar 3.1 Blok diagram penelitian

Sinyal ECG yang didapat dari aktivitas kerja jantung didapatkan dari *apnea ECG database* pada *Physionet Database*. Sinyal-sinyal tersebut kemudian diseleksi untuk dipisahkan bagian reguler dan apnea dari sinyal utuh. Potongan sinyal yang telah terseleksi sebagai reguler dan apnea kemudian masing-masing dipartisi sehingga menjadi rentang waktu yang lebih kecil. Sinyal kemudian dideteksi interval antar puncak R sinyal (RR-interval) pada rentang waktu tertentu. RR interval yang telah didapat kemudian digunakan untuk memperoleh karakteristik sinyal melalui tahap ekstraksi fitur. Adapun fitur-fitur yang diekstrak pada proses ekstraksi fitur berjumlah 11 fitur statistik yang seluruhnya didapat dari perhitungan nilai RR-interval. Nilai-nilai fitur yang sudah didapat, kemudian dikelompokkan dan diklasifikasi menggunakan teknik *Support Vector Machine* (SVM) menjadi apnea dan non-apnea.

3.2 Implementasi Sistem

3.2.1 Pengumpulan Data

Pada proses pengumpulan data, ada beberapa hal yang perlu diketahui yang berhubungan dengan proses pengumpulan data, seperti sumber data, panjang data, frekuensi sampling, dan lain sebagainya.

Seperti yang telah dijelaskan pada bab sebelumnya, data pada penelitian ini diambil dari Apnea-ECG Database pada Physionet Database. Website PhysioNet adalah *public service* dari *PhysioNet Resource* yang dikelola oleh *National Institutes of Health's NIBIB* dan *NIGMS*. Pada *Apnea Database* ini terdapat 70 rekaman sinyal ECG yang direkam dari subjek yang berbeda dengan variasi umur 27-53 tahun, baik pria maupun wanita. Data direkam selama 2 malam berturut-turut menggunakan perekaman *single-channel* ECG dengan masing-masing 1 hingga 4 kali perekaman untuk setiap subjek, tergantung dari kualitas sinyal yang dihasilkan. Durasi data bervariasi mulai dari 7 hingga 10 jam dengan frekuensi *sampling* 100 Hz.

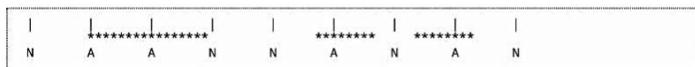
Pada masing-masing data sudah diberi label A untuk apnea dan N untuk non-apnea dalam tiap 1 menit, seperti yang ditunjukkan pada Gambar 3.2. Durasi waktu pernapasan non-apnea dalam setiap data bervariasi mulai 11 menit hingga 533 menit. Sedangkan durasi waktu pernapasan dengan apnea terjadi antara 0-534 menit. Data kemudian dikelompokkan menjadi 2 kelompok yakni durasi apnea dan non-apnea.

Gambar 3.2 Contoh rekaman sinyal dengan anotasi pada menit ke 12:00 dan 13:00

Time	Sample #	Type	Sub Chan	Num
0:00,000	0	N	0	0
1:00,000	6000	N	0	0
2:00,000	12000	N	0	0
3:00,000	18000	N	0	0
4:00,000	24000	N	0	0
5:00,000	30000	N	0	0
6:00,000	36000	N	0	0
7:00,000	42000	N	0	0
8:00,000	48000	N	0	0
9:00,000	54000	N	0	0
10:00,000	60000	N	0	0
11:00,000	66000	N	0	0
12:00,000	72000	N	0	0
13:00,000	78000	A	0	0
14:00,000	84000	A	0	0
15:00,000	90000	A	0	0
16:00,000	96000	A	0	0
17:00,000	102000	A	0	0
18:00,000	108000	A	0	0
19:00,000	114000	A	0	0
20:00,000	120000	A	0	0



Pada Gambar 3.2 ditunjukkan contoh bentuk rekaman sinyal ECG beserta anotasinya yang disajikan pada *Physionet Database*. Gambar yang ada di sisi kiri menunjukkan bentuk sinyal ECG selama durasi yang ditentukan (maksimal 60 detik) sesuai dengan kebutuhan. Dibawah sinyal tersebut terdapat anotasi yang berulang setiap 1 menit. Di masing-masing rekaman, anotasi pertama ditempatkan pada detik ke-0 dan berkaitan dengan interval 1 menit berikutnya (misal waktu 0-59.99 detik dari awal rekaman). Anotasi kedua ditempatkan pada detik ke-60 dan berkaitan dengan interval 1 menit berikutnya (60-119,99 detik), dan seterusnya. Setiap anotasi "A" menunjukkan bahwa apnea sedang berlangsung di awal menit tersebut, sedangkan setiap anotasi "N" menunjukkan bahwa apnea tidak sedang berlangsung di awal menit yang tersebut [2]. Anotasi "N" seringkali disederhanakan dengan simbol titik atau bulatan kecil. Gambar 3.3 menyajikan visualisasi yang memudahkan penjelasan anotasi diatas.



Gambar 3.3 Visualisasi penentuan anotasi pada rekaman sinyal ECG

Dalam Gambar 3.3, waktu berjalan mulai dari tepi kiri. Tanda '*' adalah dimana periode apnea terjadi. Tanda '|' adalah waktu dimana terdapat anotasi (detik ke 0, 60, 120,...). Simbol "N" adalah anotasi non-apnea dan "A" adalah anotasi untuk apnea.

Sedangkan Gambar 3.2 pada sisi kanan berisi ringkasan anotasi pada rekaman sinyal tersebut, sehingga dari tampilan tersebut dapat dilihat pada menit berapa dan hingga menit berapa apnea sedang berlangsung tanpa harus melihat keseluruhan bentuk rekaman sinyal. Dari tahap pengumpulan data ini diketahui dimenit keberapa saja apnea dan non-apnea terjadi dalam setiap perekaman untuk selanjutnya digunakan pada proses *pre-processing*.

3.2.2 *Pre-processing*

Setelah data sinyal dipilih di waktu keberapa daerah apnea dan non-apnea yang digunakan dalam penelitian ini, maka langkah selanjutnya adalah dilakukan *pre-processing*. Pada tahap ini, terdiri dari

tiga tahap yaitu pemilihan data, segmentasi dan deteksi RR Interval data apnea dan non-apnea.

1. Pemilihan Data

Data rekaman yang diambil dari database berupa format *Comma Separated Value* (CSV) yang berisi informasi waktu sampel pada kolom 1 dan besar amplitudo dengan satuan mV untuk setiap 10 ms pada kolom 2. Karena dalam pengambilan data dengan format .csv tidak bisa dilakukan untuk satu rekaman utuh, maka pengambilan data dilakukan berdasarkan waktu data terpilih dengan durasi 10-30 menit pada durasi yang telah ditentukan pada tahap pengumpulan data. Struktur data yang diambil dari database ditunjukkan

Setiap sinyal ECG dipilih antara 10 hingga 30 menit pada masing-masing periode apnea dan non-apnea. Bagian data yang dipilih dengan periode tertentu tersebut ditentukan dengan memperhatikan beberapa aspek yakni sinyal pada periode tersebut bersih (bebas *noise*), jelas fitur gelombangnya, dan tidak bercampur (terdapat apnea dan non-apnea dalam 1 pilihan periode). Jadi, setiap subjek perekaman sekurang-kurangnya memiliki 2 jenis .csv, yaitu untuk periode apnea dan non-apnea yang masing-masing berdurasi 10-30 menit.

2. Segmentasi

Bagian sinyal apnea dan non-apnea yang telah dipilih sebelumnya disegmentasi setiap 15 detik untuk dilakukan proses selanjutnya yakni deteksi R. Hal ini dikarenakan apnea didefinisikan sebagai henti napas yang terjadi ≥ 10 detik. Karena sinyal ECG memiliki frekuensi sampling 100 Hz, maka dalam 15 detik terdiri dari 1500 sampel waktu dan amplitudonya. Setiap baris dalam *file* mewakili setiap 10 ms.

3. Deteksi RR Interval

Data berdurasi 15 detik yang sudah didapat selanjutnya akan dideteksi setiap gelombang R yang ada untuk dinyatakan sebagai puncak (*peak*). Gelombang R terlebih dahulu harus dibedakan dari gelombang sinyal ECG lainnya. Oleh karena itu, dikembangkan dua kondisi berikut, agar puncak R dapat terdeteksi dengan tepat:

- a. Harus merupakan lokal maksima, yang dideteksi menggunakan fungsi *max* dalam *window* 30ms
- b. Puncak lokal maksima harus sekurang-kurangnya 3,5 standar deviasi diatas *mean*.

Segmentasi dan deteksi R dilakukan secara otomatis pada pemrograman Python, dengan tahapan seperti *pseudocode* pada Gambar 3.4 dan Gambar 3.5.

```
#Code Program Training Data
```

```
IMPORT LIBRARY
```

```
from sklearn import svm
```

```
Deklarasi
```

```
Inisialisasi list kosong A, class, data, interval, No., X, Y
```

```
INT sampel = 0 , angka = 1
```

```
INPUT nama file .CSV, 1500, label "1" / "0"
```

```
append label ke list class
```

```
Membuka file .CSV
```

```
read = membaca file .CSV per baris dengan delimiter ";"
```

```
FOR var IN read , DO
```

```
    append var[1] ke list data
```

```
    sampel = sampel + 1
```

```
    IF sampel = 1500, DO
```

```
        angka = angka +1
```

```
        FOR row IN data , DO
```

```
            #Membuat kondisi syarat 1
```

```
                IF dalam window 30 ms , DO
```

```
                    Cari lokal maksima
```

```
                    append lokal maksima ke list A
```

```
...
```

Gambar 3.4 Listing program *training* data A

```

...
#Membuat kondisi syarat 2
mencari nilai rata-rata data
mencari nilai standar deviasi
mencari nilai (standar deviasi x 3,5 ) + rata-rata

#Melakukan Deteksi R
IF nilai data dalam A > (standar deviasi x 3,5 ) +
rata-rata :
    titik tersebut adalah R
    menghitung nilai interval antar R
    append nilai interval antar R ke list interval

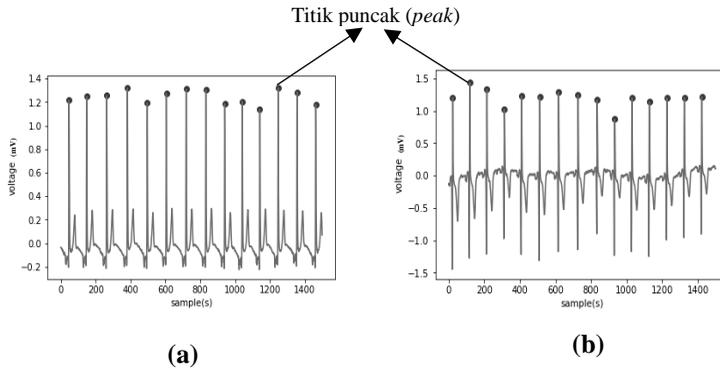
```

Gambar 3.5 Listing program *training* data B

Library *sklearn* merupakan library *machine learning* yang digunakan untuk proses klasifikasi sebagai yang telah dijelaskan pada subbab 2.5.5. Angka 1500 pada INPUT mengacu pada jumlah sampel yang didapat sesuai dengan waktu segmentasi. Seperti yang telah dijelaskan sebelumnya, segmentasi 15 detik akan mendapatkan 1500 sampel karena frekuensi sampling sinyal adalah 100 Hz. Sedangkan input label berupa angka 1 dan 0 untuk merepresentasikan Apnea (1) dan non-apnea (0).

Indeks ke-1 yang dimasukkan kedalam list *data* dipilih karena indeks ke-0 dari file tersebut merupakan informasi waktu dari sinyanya (kolom Time). Karena kolom tersebut tidak digunakan dalam pemrosesan data, maka yang digunakan hanya kolom mV yang merupakan indeks 1.

Ketika dua kondisi syarat diatas terpenuhi, maka titik tersebut merupakan puncak R sinyal. Apabila puncak R telah terdeteksi, maka selanjutnya didapat nilai RR interval. Contoh hasil deteksi puncak R dapat dilihat pada Gambar 3.6.



Gambar 3.6 Hasil deteksi puncak R pada beberapa model sinyal

Pada Gambar 3.6, disajikan contoh 2 sinyal yang berbeda dengan dikenai perlakuan deteksi R. Puncak R disajikan dalam lingkaran pada setiap titik puncak yang terdeteksi.

3.2.3 Ekstraksi Fitur

Menurut Almazaydeh [1], fitur ECG berikut adalah fitur yang paling efektif untuk mendeteksi apnea dengan perhitungan :

1. Mean periode RR Interval.
2. Standar deviasi dari periode RR-Interval.
3. Median RR-interval.
4. Ukuran NN50 (varian 1), yang didefinisikan sebagai jumlah pasangan RR interval yang berdekatan, dimana RR-interval pertama melebihi RR-interval kedua dengan perbedaan lebih dari 50 ms.
5. Ukuran NN50 (varian 2), yang didefinisikan sebagai jumlah pasangan RR interval yang berdekatan, dimana RR-interval kedua melebihi RR-interval pertama dengan perbedaan lebih dari 50 ms.
6. Dua ukuran pNN50 (untuk masing-masing NN50), yang didefinisikan sebagai setiap pengukuran NN50 dibagi dengan jumlah total RR-interval.
7. SDSD, yang didefinisikan sebagai standar deviasi perbedaan antara RR-interval yang berdekatan.
8. RMSSD, Akar kuadrat dari rata-rata selisih yang berurutan antara RR interval yang berdekatan.

9. Kisaran inter kuartil (IQR), yaitu perbedaan antara persentil ke-75 dan ke-25 dari distribusi nilai RR-interval.
10. *Mean Absolute Deviation (MAD)*, yaitu mean dari nilai absolut yang diperoleh dari pengurangan nilai rerata RR interval dari semua nilai RR interval dalam periode.

Ekstraksi fitur dilakukan seperti *pseudocode* pada Gambar 3.7.

```

...
#Ekstraksi nilai-nilai fitur
    Menghitung nilai Mean STD, Median, NN50a,
    NN50b, pNN50a, pNN50b, SDSD, RMSSD, IQR,
    MAD dari list interval

    append setiap nilai angka ke list No
    sampel = 0
    clear list data

#Membuat Database hasil ekstraksi fitur

    Simpan/tambahkan nilai
    (No, Mean, STD, Median, NN50a, NN50b,
    pNN50a, pNN50b, SDSD, RMSSD, IQR, MAD,
    class ) pada file fitur.csv

...

```

Gambar 3.7 Listing program *training* data C

Data-data nilai yang diperoleh dari ekstraksi fitur selanjutnya dijadikan satu dalam sebuah *file*. *File* ini nantinya akan dijadikan *database* atau kumpulan data yang kemudian diolah untuk proses klasifikasi. Pada proses ini, seluruh nilai disimpan ke dalam file dalam format CSV dengan menyertakan label kelasnya dan kemudian akan dipanggil kembali saat proses klasifikasi. Struktur file .CSV sesuai hasil output pada perhitungan ekstraksi fitur dapat dilihat pada Tabel 3.1.

Tabel 3.1 Struktur pada file CSV hasil ekstraksi fitur

No	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	Class
----	----	----	----	----	----	----	----	----	----	-----	-----	-------

Pada Tabel 3.1, kolom F1 menunjukkan fitur Mean, F2 adalah fitur Standar Deviasi, dan seterusnya berurutan sesuai dengan yang telah dijelaskan sebelumnya hingga F11 yang menunjukkan fitur MAD. Nilai fitur beserta labelnya tersebut berfungsi sebagai bahan acuan pembelajaran (*data training*) untuk dapat memprediksi data baru berdasarkan pola nilai fiturnya.

3.2.4 Klasifikasi

Pada tahap ini, file CSV hasil ekstraksi fitur yang telah didapat sebelumnya digunakan dengan dipanggil ke dalam program Phyton. Klasifikasi dilakukan dengan menggunakan metode *Support Vector Machine* mengacu pada [1] yang juga melakukan penelitian serupa. Karena metode SVM merupakan algoritma *supervised learning* atau pembelajaran terarah, maka data-data tersebut perlu untuk dibagi menjadi dua kelompok data yang berbeda yakni data *training* dan data *testing*.

Data *training* merupakan data yang akan diproses sebagai parameter untuk penentuan kelas atau proses klasifikasi, sedangkan data *testing* merupakan data yang akan dicari kelasnya atau data yang diklasifikasikan. Pada tahap ini, data *training* yang digunakan sejumlah 80% dan data *testing* sejumlah 20% dari total data yang ada untuk masing-masing label kelompok apnea dan non-apnea. Kernel yang digunakan adalah fungsi kernel seperti yang digunakan pada [1].

Sebagaimana dijelaskan dalam subbab 2.5.2, kernel linear SVM memerlukan 2 buah masukan (x_i, y_i). Dalam hal ini, x_i merupakan nilai 11 fitur yang didapatkan dari proses ekstraksi fitur. Sedangkan y_i adalah target dari x_i yang merupakan label kelas dari nilai fitur tersebut.

Proses klasifikasi sebenarnya sama dengan langkah membuat database fitur seperti yang telah dijelaskan, namun pada proses klasifikasi file database tersebut dipanggil kembali untuk digunakan dalam proses klasifikasi. Proses klasifikasi dapat dilihat pada Gambar 3.8.

```

...
#Proses Klasifikasi

Deklarasi
Inisiasi list kosong X, Y

Membuka file fitur.CSV
readcsv = membaca file fitur.CSV per baris dengan delimeter ";"
FOR a IN readcsv , DO
    append setiap index terakhir ke list Y
    FOR x IN a
        DO append x[1 :11] ke list X

#membuat objek klasifikasi
clf = SVM.SVC(kernel = "linear" , C=1 )

#memasukkan parameter klasifikasi X, Y
clf.fit(X, Y)

END

```

Gambar 3.8 Listing program *training* data D

Indeks ke 1 sampai 11 pada setiap baris merujuk pada 11 fitur yang terdapat pada file CSV (indeks 0 adalah isi kolom No). Sedangkan label terdapat pada indeks terakhir karena terletak di kolom terakhir pada struktur file CSV seperti pada Tabel 3.1.

Selanjutnya, SVM akan melakukan perhitungan untuk mencari *hyperplane* yang memisahkan kedua kelas. Parameter C yang mengatur seberapa jauh jarak *hyperplane* terhadap kedua data diberikan nilai 1 karena C=1 adalah nilai yang umumnya digunakan.

Data *training* dan data *testing* dipisahkan sebelum dilakukan klasifikasi. Data *testing* yang digunakan adalah data yang tidak digunakan sebagai *training* untuk menghindari *overfitting*. *Overfitting* adalah salah satu kasus yang sering terjadi pada *machine learning*, yaitu kejadian dimana *classifier* kehilangan kemampuan generalisasi pelatihannya sehingga menghasilkan prediksi kelas yang tidak akurat [17]. Maka dari

itu, data yang termasuk dalam data *training* tidak digunakan kembali dalam *testing* klasifikasi.

Selanjutnya dilakukan prediksi dengan menggunakan data *testing*. Seperti sebelumnya, nilai-nilai fitur pada data *testing* digunakan sebagai bahan untuk mendapatkan prediksi label kelasnya. Proses prediksi untuk mengklasifikasikan data *testing* dilakukan seperti *pseudocode* pada Gambar 3.9.

```
#Code Program Testing Data
```

```
IMPORT LIBRARY
```

```
from sklearn import svm
```

```
Deklarasi
```

```
Inisialisasi list kosong A, truth, interval
```

```
INPUT nama file .CSV, label "1" / "0"
```

```
Append label ke list truth
```

```
Membuka file .CSV
```

```
read = membaca file .CSV per baris dengan delimiter ";"
```

```
FOR row IN data , DO
```

```
    A ← setiap index 1
```

```
    #Membuat kondisi syarat 1 (seperti pada program Training Data)
```

```
    #Membuat kondisi syarat 2 (seperti pada program Training Data)
```

```
    #Melakukan Deteksi R (seperti pada program Training Data)
```

```
    #Ekstraksi nilai-nilai fitur (seperti pada program Training Data)
```

```
#Membuat objek prediksi data
```

```
    hasil = clf.predict (Mean, STD, Median, NN50a, NN50b,  
                          pNN50a, pNN50b, SDSD, RMSSD, IQR, MAD)
```

```
    Print hasil prediksi
```

```
    Simpan/tambahkan (nama file .csv, truth, hasil)  
    pada file testing.csv
```

```
END
```

Gambar 3.9 Listing program *Testing Data*

Hasil pengujian dituliskan pada *file testing.csv*. Dari *file* tersebut kemudian dapat dilihat sebaik apa sistem bekerja melalui klasifikasi yang dihasilkan. *Flowchart* keseluruhan dari program *Training Data* dan *Testing Data* dapat dilihat pada Gambar 3.10 dan Gambar 3.11 yang terlampir pada akhir bab 3.

3.2.5 Evaluasi Kinerja

Dari sistem klasifikasi yang telah dilakukan belum tentu didapatkan hasil klasifikasi yang sempurna. Untuk itu, perlu dilakukan pengujian yang bertujuan untuk mengukur kinerja terhadap sistem klasifikasi yang telah dilakukan sebelumnya. Pengukuran dilakukan dengan melihat kembali hasil prediksi yang dihasilkan oleh sistem pada *file testing.csv*.

Untuk mengevaluasi kinerja sistem klasifikasi digunakan 2 indikator statistik tambahan, yakni *Sensitivity (Se)* dan *Specificity (Sp)* disamping menggunakan *Accuracy (Acc)*. Dengan mengacu pada Persamaan 2.11, Persamaan 2.12, dan Persamaan 2.13, maka perlu didapatkan terlebih dahulu jumlah TP, TN, FP, dan FN dari proses klasifikasi seperti yang ditunjukkan pada Tabel 3.3.

Tabel 3.2 *Confusion Matrix*

	Prediksi Positif	Prediksi Negatif
Kondisi Positif	Benar Positif (<i>TP</i>)	Salah Negatif (<i>FN</i>)
Kondisi Negatif	Salah Positif (<i>FP</i>)	Benar Negatif (<i>TN</i>)

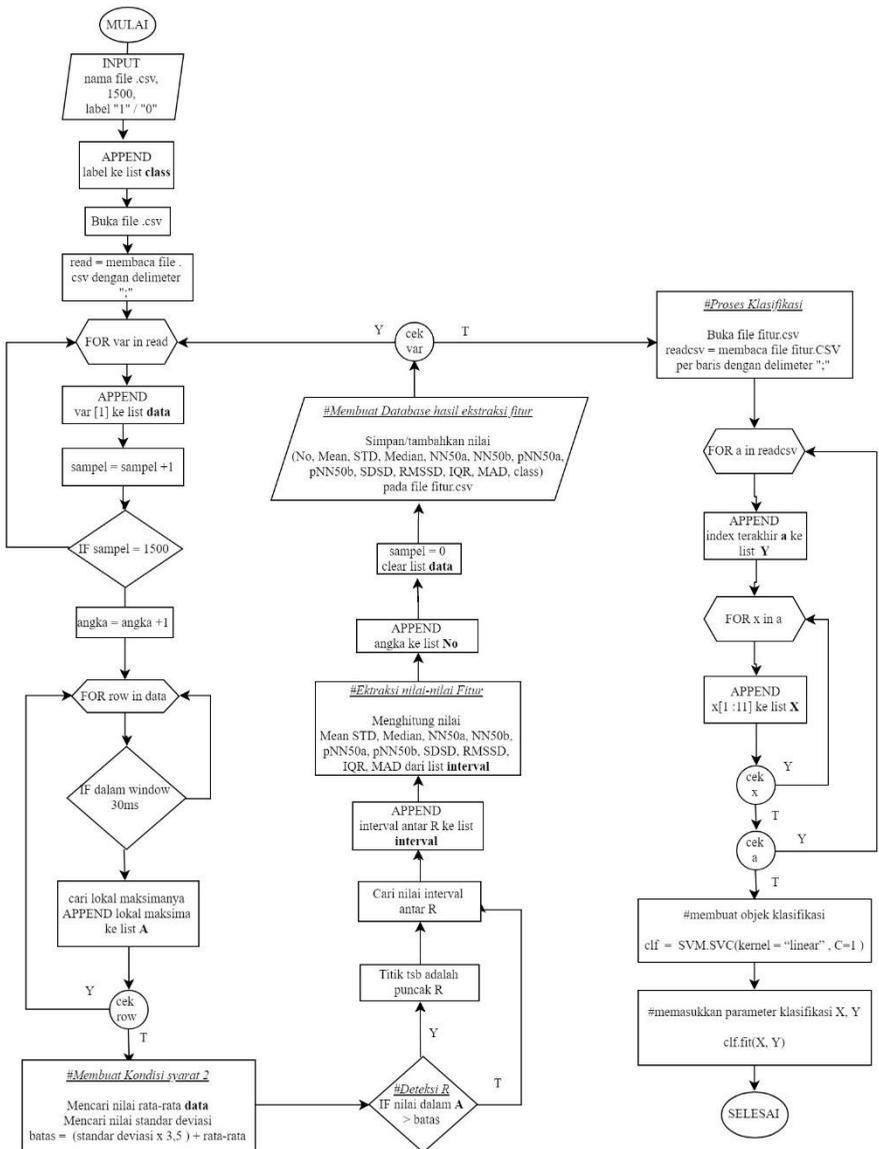
Dengan mengacu pada *confusion matrix* di Tabel 3.3, kondisi *TP* diperoleh ketika data memiliki nilai Kelas 1 mendapatkan hasil klasifikasi apnea. Sebaliknya, kondisi *TN* diperoleh ketika data dengan nilai Kelas 0 mendapat hasil klasifikasi non-apnea. Apabila data dengan nilai Kelas 1 mendapat hasil klasifikasi non-apnea, maka akan masuk dalam kondisi *FN*, dan sebaliknya juga data bernilai Kelas 0 mendapat hasil klasifikasi apnea, maka masuk dalam kondisi *FP*. Dengan demikian, apabila nilai *TP*, *TN*, *FP*, dan *FN* telah didapat *Accuracy (Acc)*, *Sensitivity (Se)*, dan *Specificity (Sp)* dapat diperoleh. Dari pengujian yang telah dilakukan dengan 5 kali percobaan, maka didapatkan hasil *Acc*, *Se*, dan *Sp* juga

sebanyak 5 nilai. Oleh karena itu, untuk mendapatkan hasil *general* dari *Acc*, *Se*, dan *Sp*, setiap nilai harus dijadikan nilai rata-rata. Hasil *Accuracy*, *Sensitivity*, dan *Specificity* rata-rata didapatkan dari Persamaan 3.1, Persamaan 3.2, dan Persamaan 3.3.

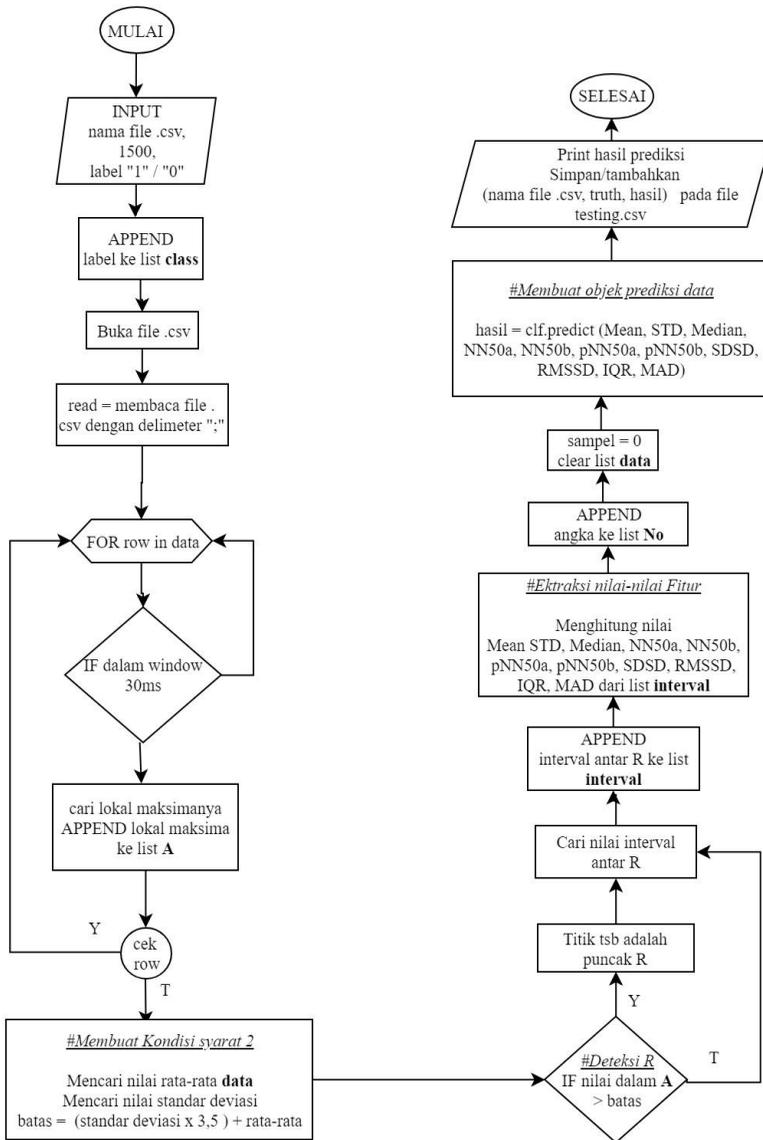
$$Accuracy = \frac{Acc1 + Acc2 + Acc3 + Acc4 + Acc5}{5} \quad (3.1)$$

$$Sensitivity = \frac{Se1 + Se2 + Se3 + Se4 + Se5}{5} \quad (3.2)$$

$$Specificity = \frac{Sp1 + Sp2 + Sp3 + Sp4 + Sp5}{5} \quad (3.3)$$



Gambar 3.10 Flowchart program Training untuk klasifikasi



Gambar 3.11 Flowchart program Testing untuk klasifikasi

BAB 4 PENGUJIAN DAN ANALISIS

Pada bab ini akan ditampilkan hasil pengumpulan data sinyal ECG, *pre-processing*, ekstraksi fitur hingga hasil proses klasifikasi. Kemudian dari data tersebut akan dipaparkan analisa penelitian dan pengujiannya.

4.1 Hasil dan Pembahasan

4.1.1 Pengumpulan Data

Seperti yang telah dijelaskan pada bab sebelumnya, data pada penelitian ini diambil dari Apnea Data pada Physionet Database. Data yang dikumpulkan untuk keperluan Tugas Akhir ini berjumlah 12 data sinyal berbeda (12 subjek) dengan variasi durasi antara 442 menit hingga 520 menit. Subjek terdiri dari 9 pria dan 3 wanita dengan umur antara 28-60 tahun, tinggi badan 168-179cm dan berat badan antara 56 hingga 103kg. Data-data yang telah dikumpulkan kemudian dipilah saat keadaan apnea dan keadaan non-apnea (normal) untuk disimpan dan digunakan pada proses *pre-processing*. Informasi subjek yang digunakan pada penelitian kali ini disajikan pada Tabel 4.1.

Tabel 4.1 Informasi subjek perekaman

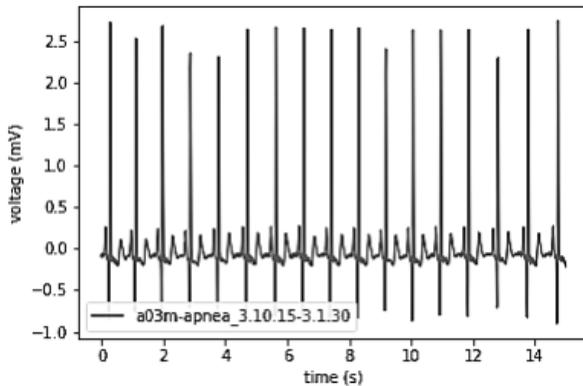
Nama Sinyal	Durasi Total (menit)	Durasi Non-apnea (menit)	Durasi Apnea (menit)	AHI	Umur (tahun)	Jenis Kelamin
a03	520	274	246	39.1	54	Pria
a05	455	179	276	41	58	Pria
a07	512	190	322	63	44	Pria
a08	502	313	189	42	51	Pria
a11	467	245	222	14	58	Pria
a13	496	252	244	42	51	Pria
a15	511	143	368	52	60	Pria
a19	503	298	205	34	55	Pria
b01	488	469	19	0.24	44	Wanita
b03	442	369	73	24	53	Pria
c06	469	468	1	0.25	28	Wanita
c07	454	450	4	0	30	Wanita

4.1.2 Pre-processing

Sinyal dengan durasi total 7-10 jam kemudian dipilih setiap 10 menit keadaan apnea dan 10 menit keadaan non-apnea. Sebagai contoh, dalam rekaman *a03* dipilih pada durasi 3:10:00.000 hingga 3:20:00.000 untuk keadaan apnea, dan untuk non-apnea dipilih pada durasi 2:28:00.000 hingga 2:38:00.000. Data diambil dan disimpan dalam format .csv yang berisi waktu dan tegangan yang terjadi setiap 10ms, seperti ditunjukkan pada Tabel 4.2.

Tabel 4.2 Contoh data yang diperoleh dari Physionet Database

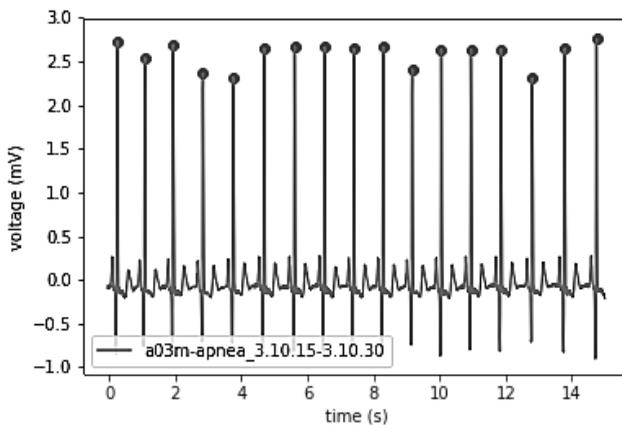
Waktu (hh:mm:ss.mmm)	ECG (mV)
03:08:45.000	-0.1
03:08:45.010	-0.125
03:08:45.020	-0.095
03:08:45.030	-0.105
03:08:45.040	-0.1
03:08:45.050	-0.11
03:08:45.060	-0.08
03:08:45.070	-0.09



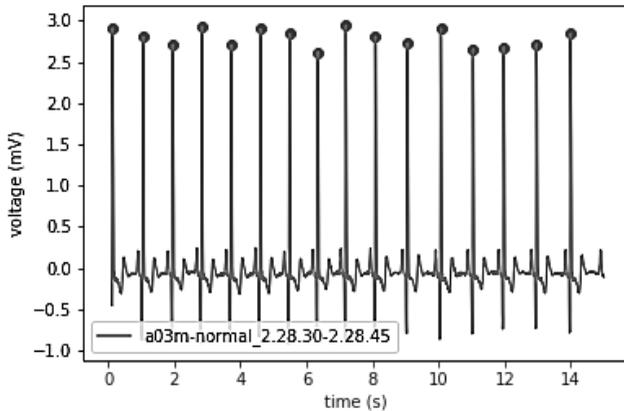
Gambar 4.1 Contoh hasil plot sinyal dengan durasi 15 detik

Dari periode 10 menit tersebut kemudian setiap sinyal disegmentasi setiap 15 detik. Angka 15 detik digunakan karena fenomena dapat disebut sebagai apnea apabila terjadi ≥ 10 detik. Selain itu, durasi ini merupakan durasi dengan tingkat akurasi terbaik menurut penelitian sebelumnya [1]. Karena frekuensi sampling sinyal adalah 100 Hz, maka dalam setiap 15 detik data terdapat 1500 *sample*. Segmentasi dilakukan secara otomatis dengan bantuan *programming* Python. Visualisasi atau plot sinyal dalam durasi 15 detik dapat dilihat pada Gambar 4.1.

Data berdurasi 15 detik yang sudah didapat selanjutnya diproses untuk dideteksi setiap gelombang R untuk dinyatakan sebagai puncak (*peak*) sinyal. Gelombang R diidentifikasi dengan memberlakukan 2 syarat yang telah dijelaskan sebelumnya pada subbab 3.2. Pada Gambar 4.2 dan Gambar 4.3, setiap gelombang R yang teridentifikasi ditandai dengan titik merah. Setelah didapatkan titik-titik puncak R, selanjutnya setiap interval atau jarak antar puncak R dihitung sebagai dasar untuk melakukan ekstraksi fitur.



Gambar 4.2 Hasil plot sinyal saat deteksi R pada keadaan apnea



Gambar 4.3 Hasil plot sinyal saat deteksi R pada keadaan non-apnea

4.1.3 Ekstraksi Fitur

4.1.3.1 Ekstraksi Fitur Sinyal

Data sinyal yang diperoleh dari *pre-processing* antara apnea dan non-apnea sulit dibedakan secara kasat mata. Oleh karena itu, untuk dapat mengetahui perbedaan keduanya dengan lebih teliti perlu dilakukan ekstraksi untuk mendapatkan nilai-nilai fiturnya. Fitur yang digunakan dalam penelitian Tugas Akhir ini berjumlah 11 fitur statistik yang lazim digunakan dalam perhitungan HRV, yaitu Mean, Standar Deviasi, Median, dua variasi nilai NN50, dua variasi nilai pNN50, SDDSD, RMSSD, kisaran interkuartil, dan MAD. Data yang diekstraksi fiturnya berjumlah 300 data hasil segmentasi dari 11 sinyal berbeda dengan rincian 150 data apnea dan 150 data non-apnea dan hasilnya disimpan dalam *file* excel sebagai *database*. Database ini selanjutnya yang akan digunakan dalam proses klasifikasi dan evaluasi kinerja. *File* terdiri dari 13 kolom dengan kolom pertama merupakan nama data yang diekstrak fiturnya, kolom 2 hingga 12 berisi nilai-nilai dari setiap fitur, sedangkan kolom terakhir adalah Kelas yang menunjukkan sinyal tersebut tergolong keadaan apnea atau non-apnea. Semua nilai memiliki satuan ms kecuali nilai NN50 yang berupa jumlah (buah) pasangan RR, dan pNN50 yang merupakan persentase (persen) dari NN50.

Setelah nilai-nilai fitur diperoleh dan disimpan dalam satu database, dapat terlihat sedikit perbedaan antara keadaan apnea dan non-apnea. Jika diamati dari nilai *mean* interval RR, sinyal apnea memiliki rentang yang lebih luas yakni antara 670,5-1394ms (*average*: 883,5ms) dibandingkan dengan sinyal non-apnea yang memiliki rentang *mean* antara 692-1002ms (*average*: 799,6ms). Apabila ditinjau dari nilai standar deviasi RR interval, rentang nilai standar deviasi sinyal apnea juga lebih besar yakni antara 12,6-580,8ms (*average*: 69,8ms), sedangkan sinyal non-apnea memiliki nilai antara 5,6-266ms (*average*: 33,7ms). Nilai SDD, RMSDD, kisaran interkuartil, dan MAD pada kondisi apnea bahkan mencapai 2x lipat nilai pada kondisi non-apnea. Hal ini dapat artikan bahwa sinyal apnea memiliki sebaran nilai yang lebih besar dari sinyal non-apnea, dan juga memiliki jarak titik data individu ke *mean* yang lebih jauh dari sinyal non-apnea. Dari nilai-nilai diatas, dapat disimpulkan bahwa sinyal apnea memiliki jarak RR yang cenderung berubah-ubah. Sedangkan sinyal non-apnea memiliki jarak RR/RR interval yang lebih stabil.

RR interval ECG yang tidak stabil tersebut dapat dikatakan representasi atau gambaran dari detak jantung yang tidak teratur. Seperti yang disebutkan pada subbab 2.2, penderita OSA identik dengan gangguan fibrilasi atrium (tidak beraturan denyut jantung) dan bradikardia (denyut jantung lambat). Gangguan fibrilasi atrium dan bradikardia tersebut menyebabkan denyut jantung penderita OSA menjadi tidak teratur.

Sebaran nilai pada sinyal berbanding lurus tingginya nilai AHI (*Apnea-Hypopnea Indeks*) sinyal tersebut. Apabila semakin besar nilai persebaran datanya, maka semakin besar pula nilai AHI dari sinyal ECG penderita. Besar nilai AHI menentukan seberapa parah OSA yang dideritanya. Semakin tinggi nilai AHI maka semakin parah pula OSA yang diderita pasien.

Penderita OSA berat tentu saja memiliki efek yang lebih besar pada kesehatan tubuhnya dibandingkan dengan penderita OSA ringan dan sedang. Semakin berat OSA yang dialami pasien, semakin ia tidak dapat tidur dengan nyenyak karena frekuensi terjadinya apnea semakin meningkat. Kejadian apnea dalam tidurnya yang dapat berulang hingga 30 kali per jam membuatnya tidak mendapatkan tidur yang nyenyak dan berkualitas. Singkatnya, dapat dikatakan sebaran nilai data berbanding terbalik pada tingkat kenyamanan tidur penderita.

Tabel 4.3 merupakan contoh hasil ekstraksi fitur yang dihasilkan:

Tabel 4.3 Contoh hasil ekstraksi fitur

No.	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	Kelas
1	814.706	22.523	820	0	0	0	0	20.682	20.767	30	19.723	0
2	812.778	31.764	805	0	1	0	0.056	25.641	25.668	40	26.42	0
3	903.75	35.51	900	0	0	0	0	18.571	20.494	50	27.969	1
4	1171.667	138.734	1125	5	3	0.417	0.25	98.535	99.041	220	118.889	1
5	865	115.866	825	0	2	0	0.125	52.747	59.104	160	87.5	1
6	937.333	39.067	940	0	0	0	0	16.66	18.127	70	30.222	1
7	1204.167	117.08	1195	4	2	0.333	0.167	112.603	113.538	160	94.167	1
8	826.25	99.93	790	0	3	0	0.188	47.451	49.666	140	79.062	1
9	972	24	960	0	0	0	0	16.552	16.69	30	19.733	1
10	905.333	26.55	900	0	1	0	0.067	24.117	24.785	20	19.022	0
11	930.667	56.506	910	2	1	0.133	0.067	38.578	39.37	100	48.889	0
12	1015.714	119.864	1005	3	1	0.214	0.071	52.262	54.702	220	107.143	1
13	958	26.882	970	0	0	0	0	24.032	25.071	40	21.333	1
14	913.333	40.442	900	2	2	0.133	0.133	40.683	40.708	60	34.667	1
15	881.25	55.551	860	2	0	0.125	0	29.484	31.833	60	40.469	1
16	948.667	24.73	950	0	0	0	0	27.479	27.516	40	21.422	1
17	851.176	36.118	870	0	0	0	0	19.675	20.463	20	26.713	1

4.1.3.2 Analisis Statistik

Setelah didapatkan hasil ekstraksi fitur statistik dari sinyal ECG, selanjutnya perlu dilakukan analisis statistik terhadap nilai-nilai tersebut. Analisis statistik yang digunakan adalah ANOVA (*Analysis of Variance*). ANOVA merupakan prosedur uji statistik yang dapat menguji perbedaan antar kelompok. ANOVA digunakan sebagai alat analisis untuk menilai apakah ada perbedaan rerata antara kelompok.. Hasil uji ANOVA dapat dilihat pada Tabel 4.4.

Tabel 4.4. Tabel *Equality of Group Means* ANOVA

VAR	Wilks' Lambda	F	df1	df2	Sig.
VAR00001	,829	58,404	1	283	,000
VAR00002	,909	28,217	1	283	,000
VAR00003	,835	56,075	1	283	,000
VAR00004	,899	31,647	1	283	,000
VAR00005	,878	39,181	1	283	,000
VAR00006	,899	31,800	1	283	,000
VAR00007	,886	36,528	1	283	,000
VAR00008	,992	2,388	1	283	,123
VAR00009	,982	5,106	1	283	,025
VAR00010	,712	114,482	1	283	,000
VAR00011	,834	56,510	1	283	,000

Tabel 4.5 Arti nama VAR untuk masing-masing nama fitur

Nama VAR	Nama Fitur
VAR00001	Mean
VAR00002	Standar Deviasi
VAR00003	Median
VAR00004	NN50a
VAR00005	NN50b
VAR00006	pNN50a
VAR00007	pNN50b
VAR00008	SDSD
VAR00009	RMSSD
VAR00010	Kisaran interkuartil
VAR00011	MAD

Dalam ANOVA yang dilakukan, nilai α atau nilai ambang batas error yang ditentukan adalah sebesar 10% atau 0,1. Dari hasil uji analisis diskriminan yang ditunjukkan Tabel 4.4, kolom VAR00001 menunjukkan fitur Mean, VAR00002 adalah fitur standar deviasi, dan terus beurutan seperti yang telah disebutkan pada Tabel 4.5 sehingga VAR00011 adalah fitur MAD.

Sesuai tujuan dilakukannya uji diskriminan ini yaitu untuk melihat seberapa berbeda nilai fitur statistik yang dihasilkan untuk membedakan setiap kelompoknya, maka dalam kasus ini kolom yang perlu diperhatikan cukup kolom paling kiri (VAR) dan kolom paling kanan (Sig) yang merupakan hasil akhir dari perhitungan terhadap kolom-kolom disebelah kirinya. Kolom Sig. menunjukkan seberapa signifikan perbedaan setiap nilai variabel terhadap kedua kelompok. Semakin kecil nilai Sig. maka semakin signifikan perbedaannya. Sebaliknya, semakin besar nilai Sig maka semakin serupa nilai tersebut untuk kedua kelompok karena semakin tidak signifikan perbedaan nilainya dari masing-masing kelompok.

Dari Tabel 4.4, dapat dilihat bahwa hampir seluruh nilai fitur yang dihasilkan memiliki nilai Sig 0,000 yang artinya nilai fitur signifikan berbeda antara masing-masing kelompok. Nilai fitur yang hampir mirip adalah VAR00010 yang merupakan nilai fitur SDSD

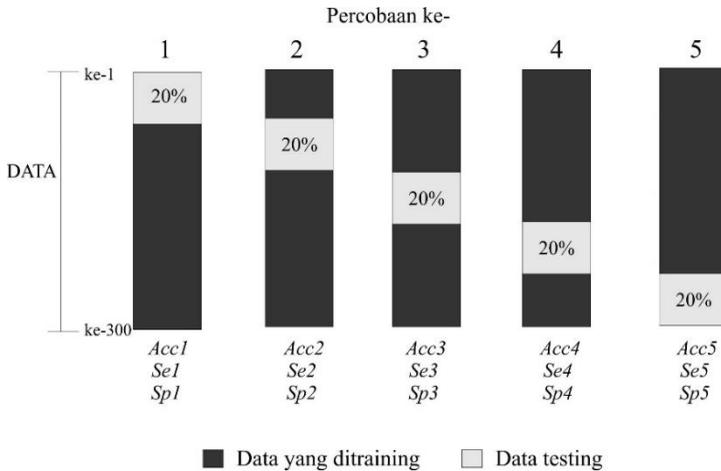
dengan nilai α sebesar 0,025 atau 2%. Sedangkan perbedaan yang paling tidak signifikan adalah VAR00009 yang merupakan fitur RMSSD dengan nilai Sig sedikit melebihi nilai α yaitu sebesar 0,123 atau 12,3%. Hasil analisis diskriminan selengkapnya dapat dilihat pada Lampiran V.

4.2 Klasifikasi dan Evaluasi Kinerja

4.2.1 Klasifikasi

Setelah seluruh data dihitung nilai fiturnya untuk tiap periode seperti pada contoh Tabel 4.3 diatas, kemudian data ini digunakan sebagai input untuk proses klasifikasi. Pada penelitian ini digunakan perbandingan persentase data *training* dan testing sebesar 80:20 persen. Jadi, dari 12 rekaman yang ada, 10 digunakan untuk training dan 2 untuk testing. Dari 12 rekaman yang telah disegmentasi menjadi menghasilkan 300 data periode singkat, sebanyak 240 data digunakan sebagai data *training* dan 60 data sisanya akan digunakan sebagai data *testing*. Pengklasifikasian ini dilakukan untuk mengetahui apakah sistem sudah memiliki kemampuan generalisasi yang baik dalam mengklasifikasikan data.

Karena data *testing* yang digunakan sebanyak 20% dari seluruh data, maka pada satu kali pengujian dilakukan 5 kali percobaan dengan menggunakan data-data yang berbeda. Karena penelitian ini menggunakan 300 data, maka percobaan pertama adalah data ke-1 sampai data ke-60, percobaan kedua adalah data ke-61 sampai data 120, dan seterusnya. Untuk percobaan 1 hingga percobaan 5, setiap kelompok data *training* memiliki rasio data apnea dan non-apnea yang sama yakni 50:50. Untuk data *testing* memiliki rasio bebas dengan ketentuan data tersebut bukan termasuk data *training* demi menghindari *overfitting*. Cara pengujian yang telah dilakukan diatas divisualisasikan pada Gambar 4.4 agar lebih mudah dipahami.



Gambar 4.4 Visualisasi teknik pengujian yang dilakukan

Balok-balok tersebut diibaratkan seperti tumpukan data 1 hingga data ke-300. *Testing* percobaan pertama dilakukan dengan mengambil 20% data pertama, dengan 80% lainnya digunakan sebagai data *training*. Percobaan kedua adalah 20% data kedua (data ke 61 – 120), dan seterusnya. Pada setiap percobaan kemudian menghasilkan nilai *Acc*, *Se*, dan *Sp*, dimana *Acc1* adalah *Accuracy* hasil percobaan 1, *Acc2* adalah hasil *Accuracy* dari percobaan 2, dan seterusnya. Begitu pula dengan *Se1*, *Se2*, dan *Sp1*, *Sp2*, dan sebagainya. Masing-masing nilai tersebut kemudian dihitung rata-ratanya untuk dijadikan sebagai nilai akurasi sistem.

Data-data yang termasuk dalam data *testing* juga tersimpan dalam file format *.csv* sehingga memudahkan dalam proses klasifikasi. Hasil klasifikasi dari data testing dapat dilihat seperti pada Tabel 4.6.

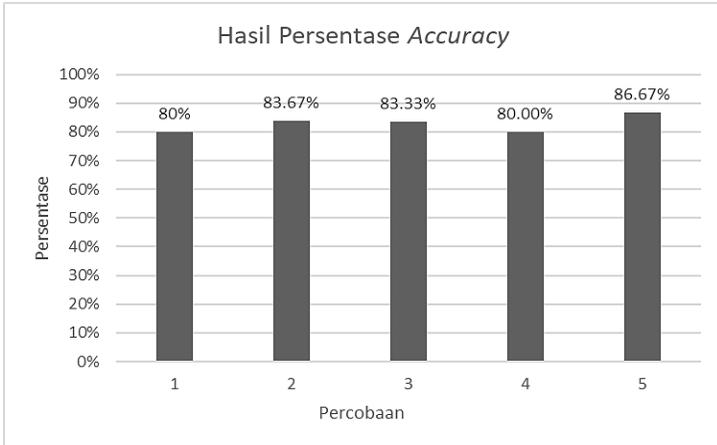
Tabel 4.6 Contoh tabel data *testing* hasil klasifikasi

Nama	Truth	Hasil
nb03-6	0	1
nb03-7	0	0
nb03-8	0	0
nb03-9	0	0
nb03-10	0	0
ab03-1	1	0
ab03-2	1	1

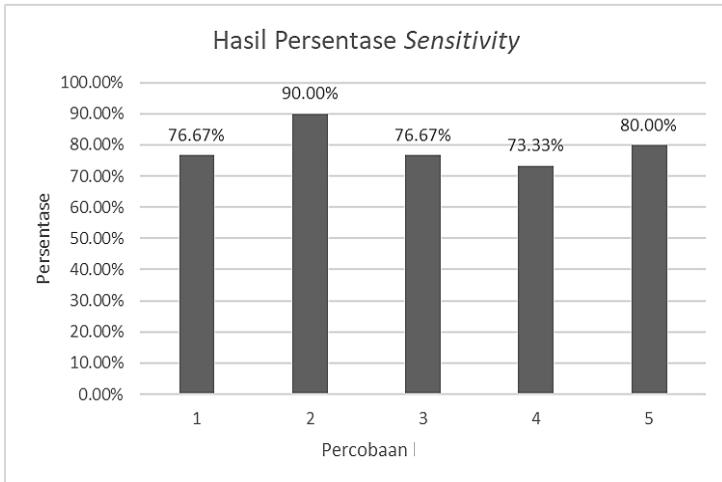
Pada Tabel 4.6 disajikan contoh hasil klasifikasi yang telah direkap seperti file sebelumnya. Daa kelas hasil klasifikasi seperti yang ditunjukkan pada tabel pada kolom Hasil, akan dibandingkan dengan nilai pada kolom *Kelas* dan digunakan untuk menghitung nilai *Accuracy* (*Acc*), *Sensitivity*, dan *Specificity*. Pada kolom Hasil, angka satu menyatakan apnea dan angka nol menunjukkan non-apnea.

4.2.2 Evaluasi Kinerja

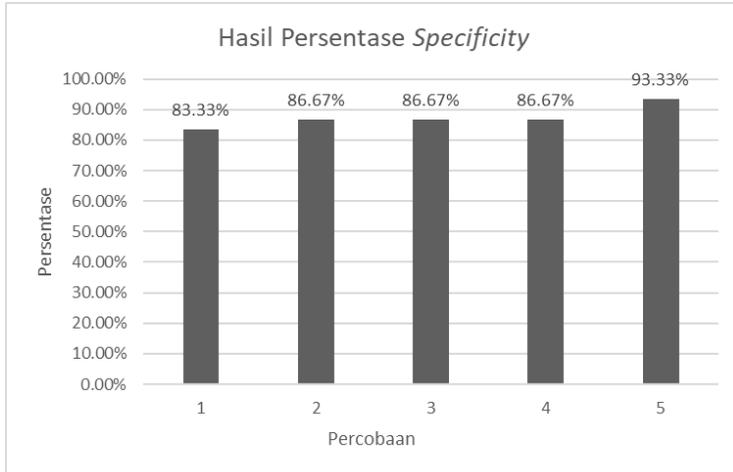
Salah satu metode yang sering digunakan untuk pengukuran kinerja adalah *Acc*, *Se*, dan *Sp* yang telah dijelaskan pada bab sebelumnya. Nilai *Acc*, *Se*, dan *Sp* diperoleh dari masing-masing jumlah *TP*, *FP*, *TN* dan *FN*. Nilai *TP*, *FP*, *TN* dan *FN* didapat dari berapa jumlah prediksi salah dan benar dengan kondisi seperti penjelasan pada subbab 3.2.5 yang diperoleh data file .csv hasil *testing*. Grafik dari hasil persentase *Acc*, *Se*, dan *Sp* pada setiap percobaan yang dilakukan dapat diamati pada Gambar 4.5, Gambar 4.6, dan Gambar 4.7.



Gambar 4.5 Grafik persentase nilai *Accuracy* pada setiap percobaan



Gambar 4.6 Grafik persentase nilai *Sensitivity* pada setiap percobaan



Gambar 4.7 Grafik persentase nilai Specificity pada setiap percobaan

Seperti yang telah dijelaskan pada subbab 4.2.1, setiap pengujian dilakukan sebanyak 5 kali dengan mengambil 20% data pertama, dengan 80% lainnya digunakan sebagai data *training*. Percobaan kedua adalah 20% data kedua (data ke 61 – 120), dan seterusnya. Pada setiap percobaan kemudian menghasilkan nilai *Acc*, *Se*, dan *Sp* seperti yang disajikan pada Gambar 4.5, Gambar 4.6, dan Gambar 4.7.

Dari data diatas, terlihat bahwa setiap percobaan memiliki nilai yang fluktuatif, baik untuk *Acc*, *Se*, dan *Sp*. Namun cenderung stabil. Hal ini dapat terjadi karena dari data yang diujikan memiliki nilai fitur yang jelas dengan rentang yang dekat dari rata-rata nilai pada kelas masing-masing, mengingat masing-masing nilai fitur pada tiap kelas memiliki rata-rata yang berbeda seperti yang telah dijelaskan pada subbab 4.1.3. Sehingga, baik data apnea maupun data non-apnea dapat terklasifikasi secara jelas. Pada percobaan 5 memiliki nilai yang relatif tinggi dibandingkan dengan percobaan lainnya, hal ini dikarenakan pada kelompok data yang diujikan pada percobaan 5, tidak terdapat data yang merupakan ECG yang tidak biasa yakni memiliki detak yang tidak beraturan meskipun dalam kondisi normal. Detak jantung yang tidak beraturan dari subjek menyebabkan RR interval juga menjadi tidak beraturan meskipun subjek tersebut bukan

pengidap apnea. Karena hal demikian, data tersebut lantas menghasilkan nilai fitur yang samar untuk dikelompokkan. Nilai samar yang dimaksud adalah ketika data tersebut susah untuk dikelompokkan karena memiliki nilai fitur yang mirip dengan kelompok lainnya. Semisal, data dengan label non-apnea yang diujikan ternyata memiliki nilai fitur seperti data apnea pada data *training*. Dengan adanya kasus demikian menjadikan terdapat kemungkinan sistem mengenali data sebagai kelas yang sama sehingga mengakibatkan hasil prediksi yang salah. Hal tersebut juga yang menjadikan nilai *Acc* pada percobaan 5 menjadi paling tinggi diantara percobaan yang lain karena pada komposisi data yang diujikan tidak banyak data ‘samar’ seperti yang telah dijelaskan. Sedangkan pada percobaan lain, jumlah data ‘samar’ tersebut lebih banyak.

Namun jika ditinjau secara keseluruhan, sebenarnya sistem sudah dapat menjalankan proses klasifikasi dan pengenalan pola dengan benar meski belum optimal, dikarenakan hampir seluruh percobaan sistem dapat memprediksi lebih dari 50% data dengan tepat. Untuk pengujian ini, didapatkan hasil rata-rata *Acc* sistem sebesar **83.67%**, *Se* **79.33%**, dan *Sp* **87.33%**.

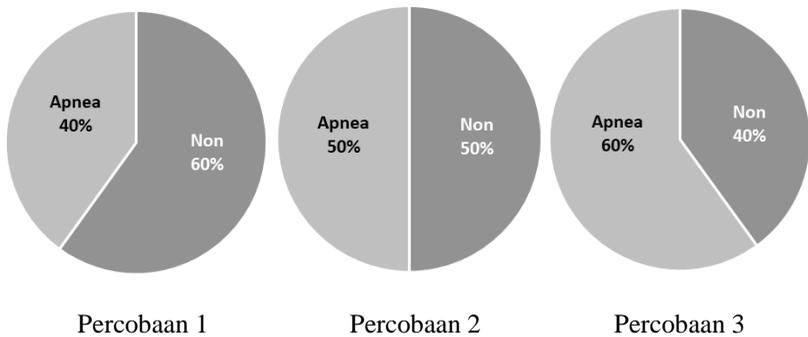
4.2.3 Pengujian terhadap Komposisi Data

Untuk mengetahui sejauh mana performansi metode, perlu dilakukan variasi proses pengujian. Seperti yang telah dijelaskan, tugas akhir ini menggunakan metode SVM menggunakan penghitungan akurasi positif-negatif. Pada tugas akhir ini dilakukan dua skenario pengujian yakni dengan data dan parameter yang berbeda-beda. Pengujian pertama yang dilakukan adalah pengujian terhadap komposisi data apnea dan non-apnea pada saat *training*. Pengujian ini dilakukan agar dapat diketahui komposisi data seperti apakah yang paling optimal dan menghasilkan nilai akurasi terbaik.

Perlu diingat kembali bahwa pada pengujian sebelumnya, rasio data yang digunakan untuk *training* dan *testing* adalah 80:20 dengan komposisi apnea dan normal 50:50. Dari 300 total data (150 data apnea dan 150 data non-apnea), 240 data *training* terdiri dari 120 data apnea dan 120 data non-apnea. Pada pengujian kali ini dilakukan variasi komposisi data apnea dan non-apnea dengan jumlah yang berbeda namun tetap dengan rasio *training* dan *testing* sama yakni 80% dan 20%. Jumlah data

apnea pun dibuat tetap agar perubahan parameter lebih mudah diamati. Pengujian dilakukan dengan tiga kali percobaan. Percobaan pertama adalah melakukan uji data dengan komposisi data non-apnea yang lebih banyak dari data apnea. Pada percobaan ini dilakukan komposisi dengan perbandingan apnea:non-apnea sebesar 40:60. Dari total 354 data, sebanyak 300 data yang terdiri dari 120 data apnea dan 180 data non-apnea digunakan sebagai data *training*.

Percobaan ketiga dilakukan perbandingan data apnea dan non-apnea adalah 60:40. Visualisasi variasi komposisi data pengujian dapat dilihat pada Gambar 4.8.

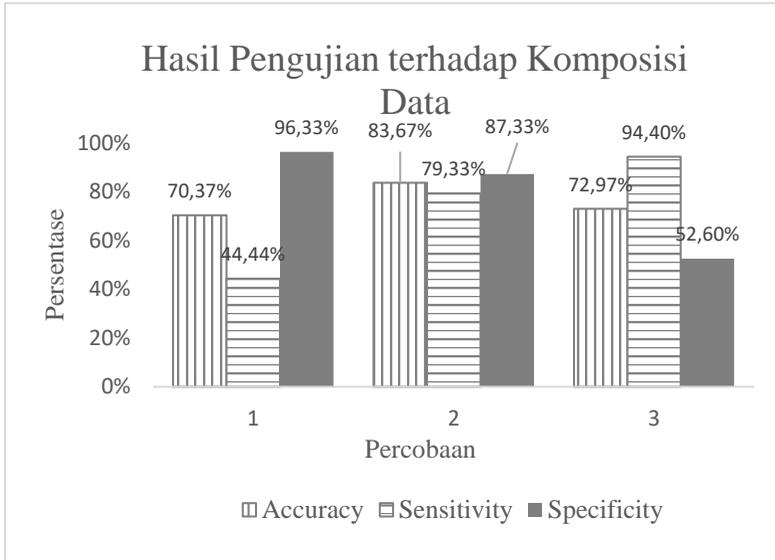


Gambar 4.8 Visualisasi variasi komposisi data apnea dan non-apnea dalam pengujian

Sebagaimana yang dilakukan pada pengujian sebelumnya dan mengacu pada subbab 4.2.1, pengujian dilakukan berulang sebanyak 5 kali untuk masing-masing percobaan. Pada pengujian ini juga berlaku hal yang sama pada setiap variasi percobaan. Jadi, setiap percobaan dilakukan perulangan *testing* sebanyak 5 kali dengan data yang berbeda-beda pada setiap *testing* untuk kemudian diambil nilai rata-rata dari setiap hasil *Acc*, *Se*, dan *Sp*.

Setelah dilakukan kedua percobaan diatas, tentu saja didapatkan hasil *Acc*, *Se*, dan *Sp* yang berbeda-beda pula. Pada percobaan pertama, dengan komposisi data non-apnea yang lebih banyak dari data apnea,

didapatkan hasil rata-rata *Acc* sebesar 70,37%, *Se* 44,44%, dan *Sp* 96,33%. Pada percobaan kedua dengan komposisi 50:50, didapatkan hasil



Gambar 4.9 Hasil akurasi terhadap variasi komposisi data apnea dan non-apnea sebagai data *training*

rata-rata akurasi *Acc* sebesar 83.67%, *Se* 79.33%, dan *Sp* 87.33%. Sedangkan pada percobaan ketiga dengan jumlah data apnea lebih banyak dari data non-apnea, didapatkan hasil rata-rata akurasi *Acc* sebesar 72,97%, *Se* 94,40%, dan *Sp* 52,60%. Hasil rata-rata akurasi pada setiap proses pengujian dapat diamati pada Gambar 4.9.

Gambar 4.9 merupakan hasil rata-rata akurasi dari setiap percobaan pengklasifikasian mulai dari percobaan pertama hingga percobaan ketiga. Dari grafik tersebut terlihat bahwa *Se* dan *Sp* adalah nilai yang saling berlawanan. Ketika nilai *Se* rendah, maka nilai *Sp* tinggi, begitu juga sebaliknya. Apabila nilai *Sp* meningkat, nilai *Se* menjadi turun. Hal ini sesuai dengan yang telah dijelaskan sebelumnya pada Bab 2. Grafik diatas juga menunjukkan bahwa sistem nilai akurasi terbaik dapat diperoleh ketika komposisi data apnea dan non-apnea seimbang.

Diluar rasio seimbang, nilai Sp dan Se akan berubah linear dengan perubahan rasio. Hasil tersebut berkaitan dengan banyaknya variasi data yang dilatih. Apabila rasio data negatif (non-apnea) lebih banyak, maka nilai Sp akan meningkat. Hal tersebut dikarenakan data yang dilatih lebih variatif karena jumlah data yang lebih banyak sehingga referensi yang didapatkan juga lebih luas sehingga ketepatan dalam pengklasifikasian menjadi lebih besar. Apabila rasio data negatif lebih sedikit dari data positif (apnea) maka nilai Sp juga akan menurun.

Hal tersebut juga secara otomatis berlaku pada Se . Apabila rasio data positif (apnea) lebih banyak, maka nilai Se akan meningkat. Sedangkan apabila rasio data positif lebih sedikit dari data negatif (non-apnea) maka nilai Se juga akan menurun.

4.2.4 Pengujian terhadap Durasi Segmentasi Data

Selain dilakukan pengujian terhadap komposisi data, pengujian kedua dilakukan terhadap durasi periode data yang akan diproses. Tujuan pengujian ini adalah agar dapat diketahui durasi segmentasi manakah yang memiliki akurasi paling baik.

Pada proses sebelumnya, segmentasi data dilakukan setiap 15 detik untuk setiap kelas apnea dan non-apnea. Pada pengujian ini dilakukan 4 percobaan lainnya, yaitu segmentasi setiap 10 detik, 20 detik, 25 detik, dan 30 detik, seperti pada Tabel 4.7.

Tabel 4.7 Skenario pengujian terhadap durasi segmentasi data

Percobaan ke-	Durasi Segmentasi Data
1	10 detik
2	15 detik
3	20 detik
4	25 detik
5	30 detik

Dengan frekuensi sampling 100Hz, maka periode 10 detik terdiri dari 1000 sampel, 20 detik 2000 sampel, periode 25 detik terdiri dari 2500 sampel, dan 30 detik terdiri dari 3000 sampel untuk setiap data. Proses selanjutnya dilakukan sama persis seperti sebelumnya yakni deteksi RR interval, ekstraksi fitur, dan klasifikasi. Dari total data tersegmentasi pada

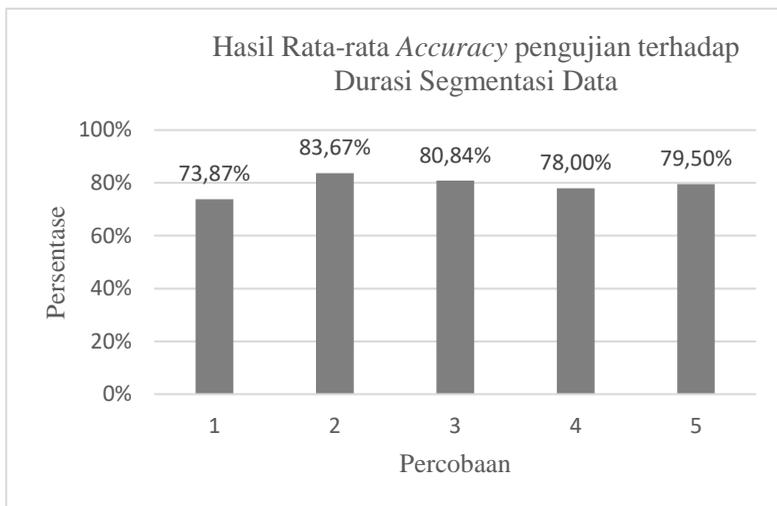
setiap durasinya, dibagi menjadi data *training* dan data *testing*. Data diambil dari durasi perekaman yang sama yang kemudian disegmentasi menurut periode durasi tertentu. Sehingga, total data pengujian untuk setiap periode tidak sama. Data yang disegmentasi dengan periode semakin kecil akan memiliki jumlah data semakin banyak, begitu juga sebaliknya. Jumlah data yang dijadikan sebagai data *training* pada setiap rekaman untuk setiap kelas apnea dan non-apnea memiliki jumlah yang sama. Dengan mengacu Tabel 4.7, pembagian data *training* dan data *testing* pada setiap percobaan disajikan pada Tabel 4.8.

Tabel 4.8 Skenario pembagian data *training* dan *testing*

Percobaan	Jumlah Data		
	Total	<i>Training</i>	<i>Testing</i>
1	480	384	96
2	300	240	60
3	260	208	52
4	240	192	48
5	200	160	40

Pada percobaan pertama sistem dilakukan proses pengenalan dengan data *training* yang berjumlah 384 data terhadap data uji berjumlah 96 data. Jumlah data hasil segmentasi yang digunakan sebagai data *training* pada percobaan ini sebanyak 192 data pada setiap kelasnya. Pada setiap kelas, diambil 24 data dari masing-masing rekaman ECG (8 rekaman). Sehingga total data pada data *training* adalah 384 sampel (192 data dikali 2 kelas). Selanjutnya sistem digunakan untuk mengenali data *training* yang berjumlah 240 data terhadap data uji berjumlah 60 data. Jumlah data yang digunakan sebagai data *training* berjumlah 120 data pada setiap kelasnya dengan setiap kelas terdiri dari 15 data dari masing-masing rekaman sinyal ECG. Pada percobaan ketiga, jumlah data yang digunakan sebagai data *training* sebanyak 104 data pada setiap kelasnya dengan setiap kelas terdiri dari 13 data dari masing-masing rekaman sinyal ECG. Pada percobaan keempat, jumlah data yang digunakan sebagai data *training* sebanyak 96 data pada setiap kelasnya dengan setiap kelas terdiri dari 12 data dari masing-masing rekaman sinyal ECG. Pada percobaan kelima, jumlah data yang digunakan sebagai data *training*

sebanyak 80 data pada setiap kelasnya dengan setiap kelas terdiri dari 10 data dari masing-masing rekaman sinyal ECG. Seperti yang dilakukan pada pengujian sebelumnya dan sesuai dengan subbab 4.2.1, pengujian juga dilakukan 5 kali untuk setiap percobaan. Hasil rata-rata akurasi dari setiap percobaan dapat dilihat pada Gambar 4.10.



Gambar 4.10 Hasil persentase rata-rata Accuracy pada pengujian terhadap durasi segmentasi data

Gambar 4.10 merupakan hasil akurasi dari setiap percobaan pengklasifikasian mulai dari percobaan pertama hingga percobaan kelima. Mengacu pada Tabel 4.7, percobaan 1 adalah percobaan dengan durasi segmentasi 10 detik dan percobaan kedua adalah durasi segmentasi 15 detik. Percobaan ketiga adalah durasi segmentasi 20 detik, percobaan keempat adalah durasi segmentasi 25 detik dan percobaan kelima adalah percobaan dengan durasi segmentasi 30 detik.

Pada percobaan pertama, pengklasifikasian menggunakan segmentasi 10 detik memiliki akurasi yang paling rendah, yaitu 73,87%. Besar akurasi terus meningkat secara signifikan pada percobaan 2 yakni dengan durasi segmentasi 15 detik. Selanjutnya akurasi cenderung turun pada percobaan 3 dan konstan pada percobaan-percobaan berikutnya.

Pada percobaan tersebut akurasi telah mencapai tingkat yang optimal dimana akurasi yang diperoleh tidak mengalami peningkatan, yaitu pada segmentasi data dengan durasi 15 detik.

4.2.5. Pengujian Terhadap Kernel SVM yang Digunakan

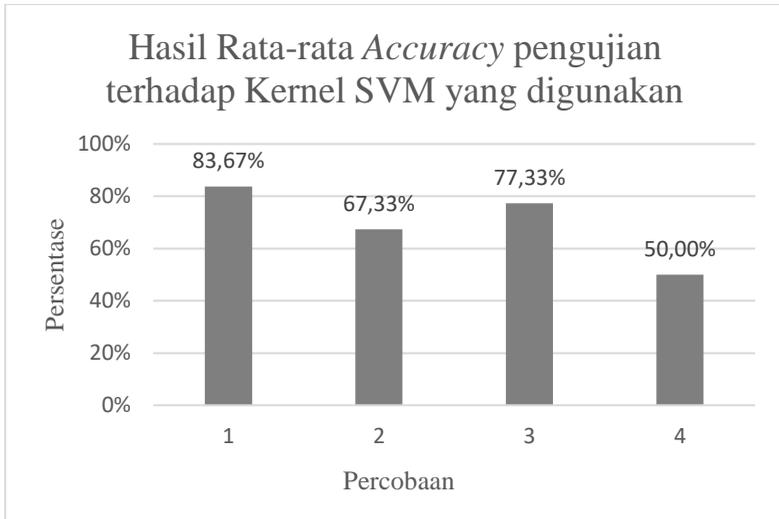
Disamping melakukan pengujian terhadap variasi data, dilakukan juga pengujian terhadap kernel SVM yang digunakan untuk *learning* untuk mengetahui kernel manakah yang dapat memberikan hasil optimal.

Seperti yang telah disampaikan pada subbab 3.2.4, seluruh pengujian dilakukan dengan kernel linear untuk proses *learning* data. Pada pengujian kali ini dilakukan percobaan dengan melakukan *learning* dengan kernel yang berbeda-beda. Variasi kernel yang digunakan adalah fungsi kernel *Radial Basis Function (RBF)*, kernel polinomial berderajat 3 dan kernel sigmoid, seperti yang disajikan pada Tabel 4.9.

Tabel 4.9 Skenario pengujian terhadap kernel SVM yang digunakan

Percobaan ke-	Kernel yang digunakan
1	Linear
2	RBF
3	Polinomial berderajat 3
4	Sigmoid

Pada setiap percobaan, data yang digunakan untuk *training* dan *testing* adalah sama persis baik dari segi komposisi maupun jumlahnya. Data yang digunakan adalah data dengan durasi segmentasi 15 detik. Data yang digunakan untuk *training* berjumlah 240 data dan data uji berjumlah 60 data. Jumlah data yang digunakan sebagai data *training* berjumlah 120 data pada setiap kelasnya dengan setiap kelas terdiri dari 15 data dari masing-masing rekaman sinyal ECG. Seperti yang dilakukan pada pengujian sebelumnya dan sesuai dengan subbab 4.2.1, pengujian juga dilakukan 5 kali untuk setiap percobaan untuk kemudian didapatkan hasil akurasi rata-ratanya. Hasil rata-rata akurasi dari setiap percobaan dapat dilihat pada Gambar 4.11.



Gambar 4.11 Hasil persentase rata-rata Accuracy pada pengujian terhadap kernel SVM yang digunakan.

Gambar 4.11 merupakan hasil akurasi dari setiap percobaan pengklasifikasian mulai dari percobaan pertama hingga percobaan keempat. Mengacu pada Tabel 4.9, percobaan 1 adalah percobaan dengan menggunakan fungsi kernel linear dan percobaan kedua menggunakan fungsi kernel RBF. Percobaan ketiga menggunakan fungsi kernel polinomial berderajat 3 dan percobaan keempat menggunakan fungsi kernel sigmoid.

Dari Gambar 4.11, terlihat bahwa *learning* dengan menggunakan kernel linear memiliki hasil yang lebih optimal dibandingkan dengan menggunakan kernel yang lain. Hal ini dikarenakan data yang dimodelkan berupa pemrosesan teks yang dapat dipisahkan hanya dengan garis linear.

4.2.6. Pengujian Pengurangan Fitur

Untuk mengetahui bagaimana pengaruh fitur terhadap akurasi, maka dilakukan percobaan pengurangan fitur. Pengurangan fitur dilakukan secara bertahap dengan mengurangi 1 fitur pada setiap pengujian. Urutan dari fitur manakah yang dihilangkan terlebih dahulu mengacu pada hasil uji analisa ANOVA seperti yang ditunjukkan pada Tabel 4.10.

Tabel 4.10 *Structure Matrix*

VAR	Function
	1
VAR00010	,822
VAR00001	,587
VAR00011	,578
VAR00003	,575
VAR00005	,481
VAR00007	,464
VAR00006	,433
VAR00004	,432
VAR00002	,408
VAR00009	,174
VAR00008	,119

Tabel 4.10 merupakan tabel *structure matrix* dari analisa ANOVA. Tabel tersebut berisi informasi tentang seberapa kuat pengaruh sebuah variabel terhadap pengambilan keputusan pengelompokkan. Seperti yang telah dijelaskan pada Tabel 4.5, kolom VAR pada tabel ini juga berlaku demikian. Kemudian kolom sebelah kanan menunjukkan seberapa besar pengaruh fitur tersebut dalam pengambilan keputusan untuk mengelompokkan menjadi apnea dan non-apnea. Semakin besar nilai *Function*, maka semakin besar pula pengaruh fitur tersebut. Dari Tabel 4.10 terlihat bahwa fitur yang paling berpengaruh adalah VAR00010 atau fitur kisaran interkuartil, sedangkan fitur yang paling kecil pengaruhnya adalah fitur SDDS.

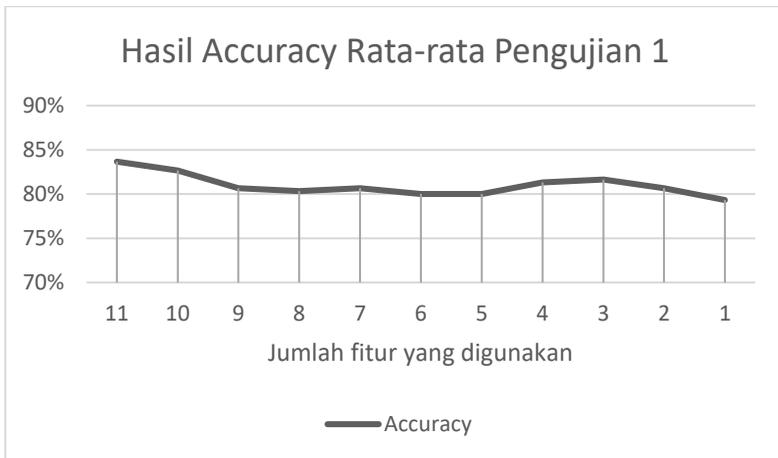
Pengujian pengurangan fitur dilakukan sebanyak 2 kali pengujian. Pengujian pertama dimulai dari mengurangi fitur yang paling kecil pengaruhnya yakni SDDS (VAR00008), dilanjutkan mengurangi fitur SDDS dan RMSDD, dan seterusnya pengurangan terus dilakukan pada fitur urutan di atasnya hingga menyisakan 1 buah fitur saja yakni VAR00010 atau kisaran interkuartil.

Sebaliknya, untuk pengujian kedua pengurangan dilakukan mulai dari urutan teratas, yakni kisaran interkuartil. Dilanjutkan dengan pengurangan fitur kisaran interkuartil dan Mean, dan seterusnya hingga hanya menyisakan fitur SDDS.

Dengan kata lain, setiap pengujian terdiri dari 10 percobaan. Percobaan pertama menggunakan 10 fitur, percobaan kedua menggunakan 9 fitur, dan seterusnya hingga menggunakan 1 fitur saja. Seperti pengujian-pengujian yang telah dilakukan sebelumnya, setiap percobaan pengurangan fitur akan dilakukan dengan 5 kali percobaan sehingga didapat nilai akurasi rata-ratanya. Hasil akurasi rata-rata dari masing-masing percobaan pada pengujian pertama dapat dilihat pada Tabel 4.11 dan grafik perubahan akurasi terhadap pengurangan fitur dapat diamati pada Gambar 4.12. Sedangkan hasil akurasi rata-rata pada pengujian kedua dapat dilihat pada Tabel 4.12 dan grafik perubahan akurasinya dapat diamati pada Gambar 4.13

Tabel 4.11 Hasil *accuracy* rata-rata pengujian pertama

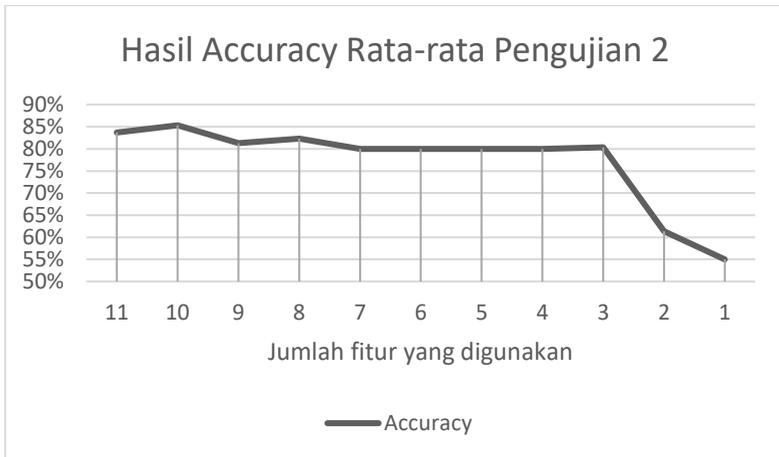
Jumlah fitur yang Digunakan	Accuracy
11 Fitur	83,67%
10 Fitur	82,67%
9 Fitur	80,67%
8 Fitur	80,33%
7 Fitur	80,67%
6 Fitur	80%
5 Fitur	80%
4 Fitur	81,33%
3 Fitur	81,66%
2 Fitur	80,67%
1 Fitur	79,33%



Gambar 4.12 Grafik tren hasil *accuracy* rata-rata pada pengujian pertama

Tabel 4.12 Hasil *accuracy* rata-rata pengujian kedua

Jumlah fitur yang Digunakan	Accuracy
11 Fitur	83,67%
10 Fitur	85,33%
9 Fitur	81,33%
8 Fitur	82,33%
7 Fitur	80%
6 Fitur	80%
5 Fitur	80%
4 Fitur	80%
3 Fitur	80,34%
2 Fitur	61,33%
1 Fitur	55%



Gambar 4.13 Grafik tren hasil *accuracy* rata-rata pada pengujian kedua

Dari Tabel 4.11 dan Tabel 4.12 disajikan akurasi rata-rata masing-masing percobaan dan pada Gambar 4.12 dan 4.13 disajikan visualisasi tren perubahan akurasi terhadap jumlah fitur yang digunakan. Pada kedua pengujian didapatkan kesimpulan bahwa semakin sedikit fitur yang digunakan maka akan semakin berkurang pula akurasinya. Selain itu, jika dihubungkan kembali dengan nilai *structure matrix* pada Tabel 4.11, fitur yang memiliki nilai *Function* yang hampir sama juga menghasilkan akurasi yang tidak jauh berbeda apabila fitur tersebut dihilangkan berurutan. Terlihat dari hasil pengujian kedua pada penggunaan 7 fitur hingga 3 fitur memiliki nilai akurasi yang hampir sama. Namun pada penggunaan 2 fitur nilai *Function* jauh berbeda sehingga akurasi menjadi turun cukup drastis.

Sedangkan pada pengujian pertama nilai akurasi cenderung turun namun masih fluktuatif. Hal ini dikarenakan pengurangan dilakukan dari fitur yang memiliki pengaruh paling kecil dalam pengambilan keputusan pengelompokkan, sehingga meskipun jumlah fiturnya berkurang namun besarnya pengaruh fitur-fitur yang digunakan masih mempengaruhi hasil akurasi. Dari kedua pengujian tersebut didapatkan hasil bahwa dengan menghilangkan fitur Kisaran interkuartil maka akurasi rata-rata bisa meningkat mencapai 85,33%. Apabila dilakukan pengurangan fitur yang lain baik 1 fitur atau lebih maka akurasinya akan menurun.

BAB 5

PENUTUP

5.1 Kesimpulan

Dalam tugas akhir ini dibahas mengenai klasifikasi *obstructive sleep apnea* dari irama jantung atau sinyal ECG dengan menggunakan klasifikasi *Support Vector Machine* dimana fitur statistik didapatkan berdasarkan pada RR interval yang terjadi. Hasil dari klasifikasi tersebut akan dihitung nilai *Accuracy*, *Sensitivity*, dan *Specificity* untuk mengetahui performansi metode SVM dalam tugas akhir ini. Berdasarkan hasil pengujian dan performa yang didapatkan, kesimpulan yang dapat diambil adalah sebagai berikut :

1. Pengklasifikasian sinyal untuk deteksi OSA dari ekstraksi fitur sinyal ECG berdasar RR interval dengan menggunakan 11 fitur statistik yaitu Mean, Standar deviasi, Median, 2 nilai NN50, 2 nilai pNN50, SDD, RMSD, Kisaran interkuartil, dan MAD dapat bekerja dengan baik, dibuktikan dengan hasil performansi sistem meraih *Accuracy* 83.67%, *Sensitivity* 79.33% dan *Specificity* 87.33%.
2. Pada pengujian terhadap komposisi data *training*, didapatkan hasil :
 - a. Dengan rasio yang seimbang memiliki tingkat *Accuracy*, *Sensitivity*, dan *Specificity* yang paling optimal.
 - b. Diluar rasio seimbang, nilai *Sp* akan meningkat apabila data *training* negatif (non-apnea) memiliki jumlah lebih banyak dan menurun apabila jumlah data training negatif lebih sedikit dari data positif. (apnea), seperti yang dapat dilihat pada Gambar 4.8.
3. Pada pengujian terhadap durasi segmentasi data, didapatkan hasil optimal yang digunakan untuk klasifikasi adalah dengan melakukan segmentasi data dengan durasi 15 detik karena menunjukkan nilai akurasi yang paling baik yaitu rata-rata mencapai 83.67% seperti yang ditunjukkan pada Gambar 4.10.
4. Pada pengujian terhadap kernel SVM yang digunakan, didapatkan hasil yakni fungsi kernel yang paling baik digunakan untuk pelatihan klasifikasi pada tugas akhir adalah menggunakan kernel linear,

dengan hasil akurasi rata-rata yang paling baik yaitu 83,67% seperti yang ditunjukkan pada Gambar 4.11.

5. Pada pengujian pengurangan fitur didapatkan hasil bahwa dengan menghilangkan fitur Kisaran interkuartil dapat meningkatkan akurasi rata-rata sistem. Nilai akurasi rata-rata yang dihasilkan mencapai 85,33%. Apabila dilakukan pengurangan fitur yang lain baik 1 fitur saja atau lebih maka akurasinya akan menurun, seperti yang ditunjukkan pada Tabel 4.13 dan Tabel 4.14.

5.2 Saran

Demi pengembangan lebih lanjut mengenai tugas akhir ini, disarankan beberapa langkah lanjutan sebagai berikut :

1. Penambahan jumlah sampel subjek untuk meningkatkan variasi data.
2. Mengoptimalkan pemilihan fitur-fitur yang ada pada sinyal ECG yang belum digunakan pada tugas akhir ini.
3. Proses klasifikasi dilakukan secara *real-time* sehingga dapat secara langsung digunakan untuk mendeteksi penyakit Apnea.

DAFTAR PUSTAKA

- [1] L. Almazaydeh, K. Elleithy and M. Faezipour, "Detection of Obstructive Sleep Apnea through ECG Signal Features", IEEE International Conference on Electro/Information Technology. 2012. (Dikutip pada halaman 1, 2, 6, 16, 25, 30, 35).
- [2] PhysioNet, *www.physionet.org*. (Dikutip pada halaman 2, 24).
- [3] Sleep Apnea Overview. *www.ncbi.nlm.nih.gov*. (Dikutip pada halaman 1, 7).
- [4] S. Bambang and D. Rusmala, "Obstructive sleep apnea syndrome pada Anak", Sari Pediatri, Vol. 7, No. 2, pp. 77 – 84, September 2005. (Dikutip pada halaman 5).
- [5] B. Antariksa, EDITORIAL "Obstructive Sleep Apnea (OSA)", Jurna Respirology Indonesia, 30, 2010. (Dikutip pada halaman 1, 5).
- [6] C. V. Silva, A. Philominraj and C. del Río, "A DSP Practical Application: Working on ECG Signal", Applications of Digital Signal Processing, 10.5772/25499, 2011. (Dikutip pada halaman 8).
- [7] Modul SkillabA-Jilid 1, "Elektrokardiografi", Lab. Ketrampilan Medik PPD Unsoed. (Dikutip pada halaman 7).
- [8] ATS Patient Education Series, "Obstructive Sleep Apnea and Heart Disease", American Thoracic Society, Am J Respir Crit Care Med Vol 188, P1-P2, 2013. (Dikutip pada halaman 1, 6,7).
- [9] H. Jambukia, Shweta, D. Vipul and P. Harshadkumar, "Classification of ECG signals using Machine Learning Techniques: A Survey," in International Conference on Advances in Computer Engineering and Applications (ICACEA), IMS Engineering College, Ghaziabad, India, 2015. (Dikutip pada halaman 8, 9).

- [10] P. Siwindarto, "Heart Rate Variability (HRV)" <URL: <http://instrumentasi.lecture.ub.ac.id/heart-rate-variability-hrv>>, Februari, 2012. (Dikutip pada halaman 10).
- [11] A.S. Nugroho, A.B. Witarto, and D. Handoko, "Application of Support Vector Machine in Bioinformatics", Proceeding of Indonesian Scientific Meeting in Central Japan, 2003. (Dikutip pada halaman 13, 15,16,17,18,19).
- [12] M. Khalil and O.A. Rifaie, "Electrocardiographic changes in obstructive sleep apnoea syndrome," in *Respiratory Medicine* 92, pp. 25-27, 1998. (Dikutip pada halaman 8).
- [13] Task Force of The European Society of Cardiology and The North American Society of Pacing and Electrophysiology, "Heart Rate Variability—Standard of Measurement, Physiological Interpretation, and Clinical Used". *Circulation*, Maret 1996. (Dikutip pada halaman 10, 13).
- [14] K. Sembiring, "Penerapan Teknik Support Vector Machine untuk Pendeteksian Intrusi pada Jaringan", Tugas Akhir S1 Teknik Informatika, Sekolah Teknik Elektro dan Informatika, ITB, 2007. (Dikutip pada halaman 14).
- [15] A. Baratloo, M. Hosseini, A. Negida, and G. El Ashal. "Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity." *Emergency* 3(2), pp. 48-49, 2015. (Dikutip pada halaman 19).
- [16] I. Cholissodin, "Pengenalan Pola/Pattern Recognition : Support Vector Machine (SVM)" <URL: http://imamcs.lecture.ub.ac.id/files/2015/02/Support-Vector-Machines_EJ_v5.06.pptx >, Februari, 2015. (Dikutip pada halaman 16).
- [17] Wang Juan, "Parameter SVM" <URL: <http://mypages.iit.edu/~jwang134/post/Parameter-in-SVM.html> >, Februari, 2015. (Dikutip pada halaman 26).

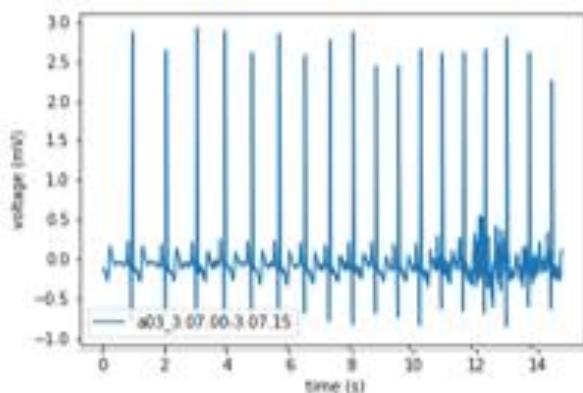
LAMPIRAN

Lampiran I.

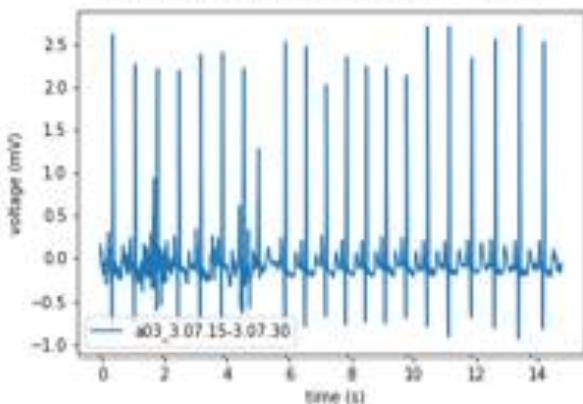
Plot sinyal ECG pada Durasi 15 Detik Apnea dan Non-apnea

Pada Lampiran I, akan ditampilkan plot rekaman dari masing-masing subjek untuk dapat dilihat perbedaan bentuk sinyal saat terjadi apnea dan saat dalam keadaan normal. Pada halaman sisi sebelah kiri (halaman buku gasal) berisi plot pada periode apnea, sedangkan pada halaman sisi kanan (halaman buku genap) berisi plot pada periode non-apnea atau normal.

Plot sinyal ECG pada Durasi 15 Detik Apnea dan Non-apnea

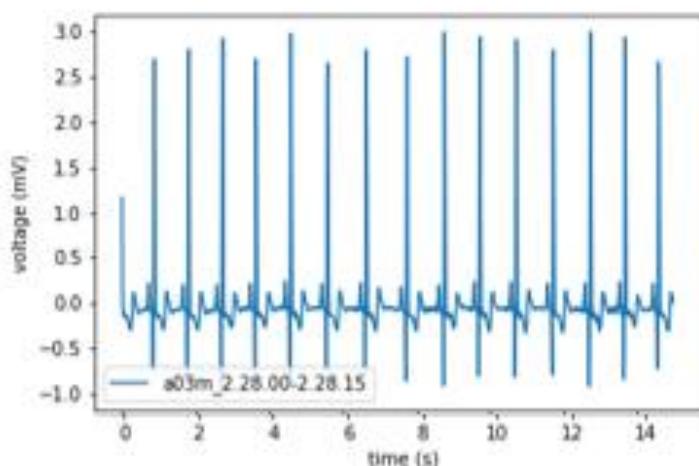


(a) Sampel waktu ke 03:07:00 – 03:07:15

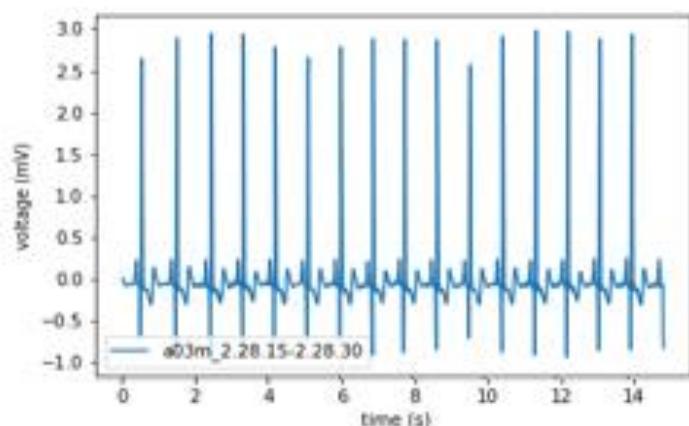


(b) Sampel waktu ke 03:07:15 – 03:07:30

Gambar 1 Bentuk sinyal rekaman a03 periode apnea durasi 15 detik

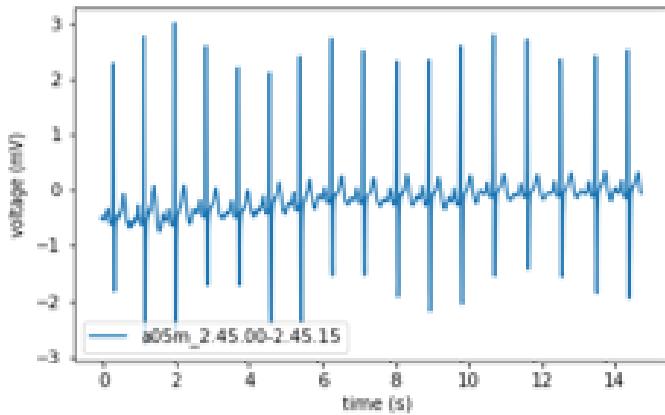


(a) Sampel waktu ke 02:28:00 – 02:28:15

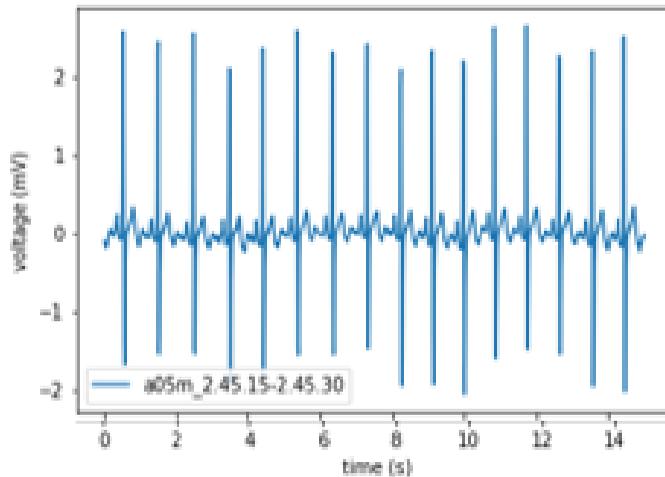


(b) Sampel waktu ke 02:28:15 – 02:28:30

Gambar 2 Bentuk sinyal rekaman a03 periode non-apnea durasi 15 detik

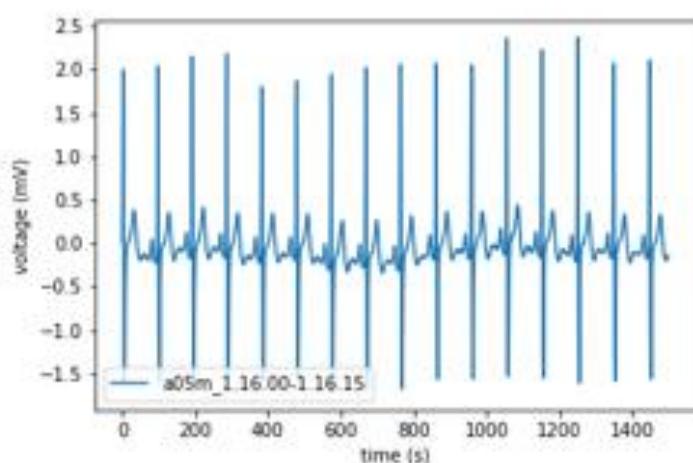


(a) Sampel waktu ke 02:45:00 – 02:45:15

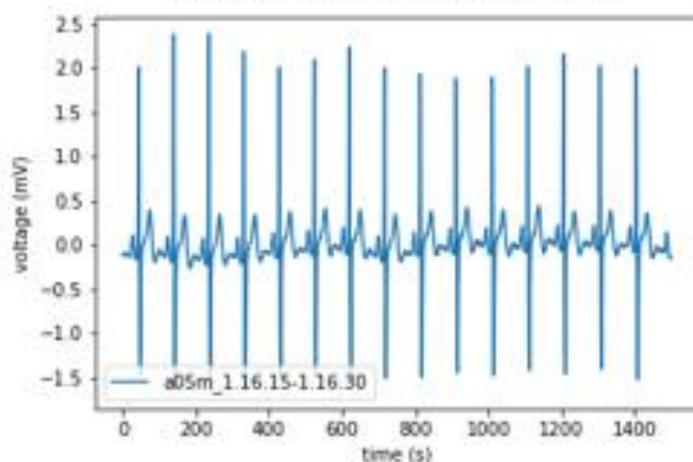


(b) Sampel waktu ke 02:45:15 – 02:45:30

Gambar 3 Bentuk sinyal rekaman a05 periode apnea durasi 15 detik

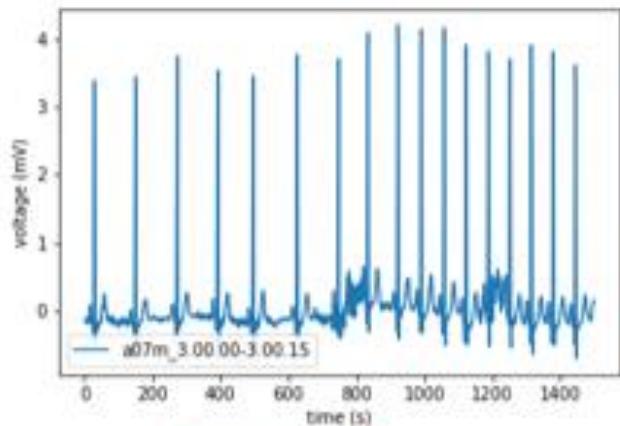


(a) Sampel waktu ke 01:16:00 – 01:16:15

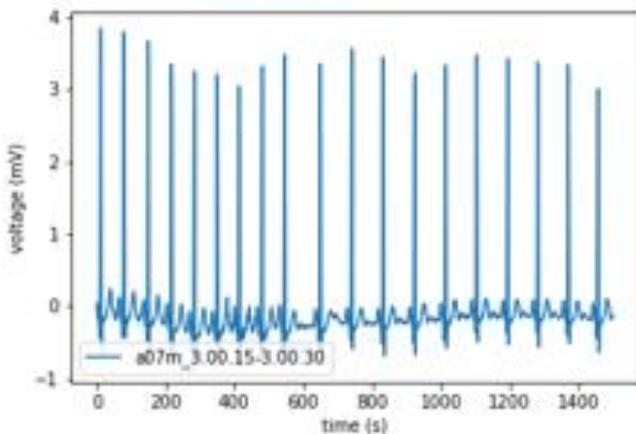


(b) Sampel waktu ke 01:16:15 – 01:16:30

Gambar 4 Bentuk sinyal rekaman a05 periode non- apnea durasi 15 detik

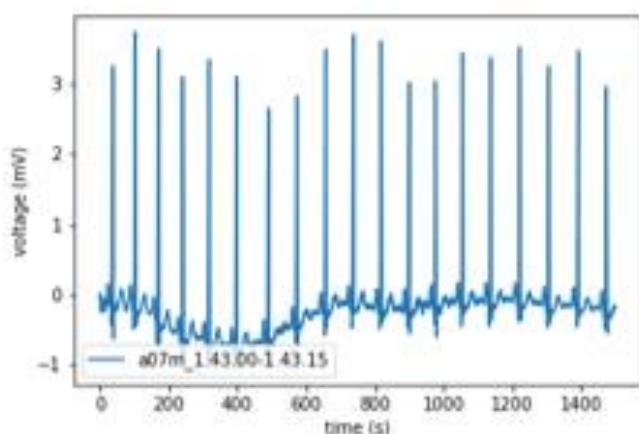


(a) Sampel waktu ke 03:00:00 – 03:00:15

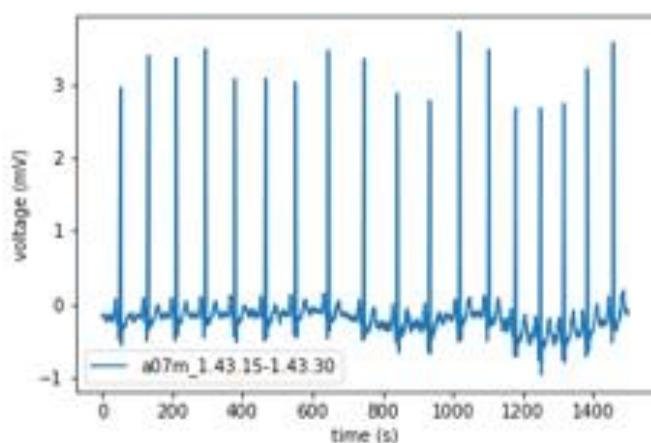


(b) Sampel waktu ke 03:00:15 – 03:00:30

Gambar 5 Bentuk sinyal rekaman a07 periode apnea durasi 15 detik

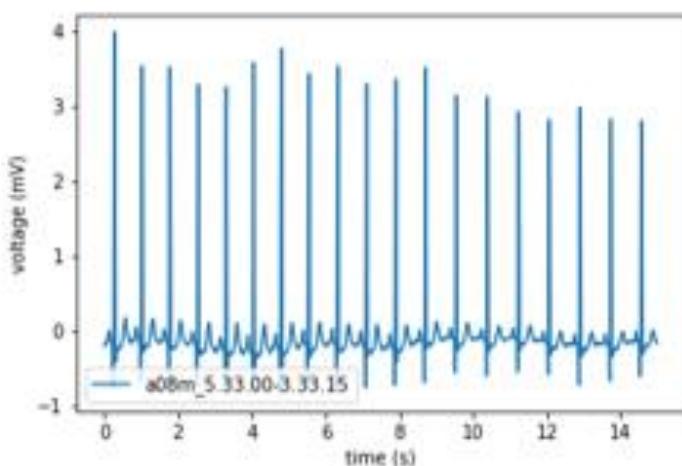


(a) Sampel waktu ke 01:43:00 – 01:43:15

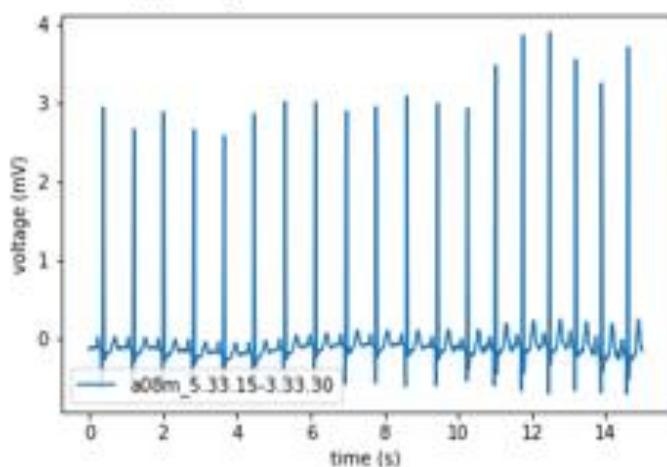


(b) Sampel waktu ke 01:43:15 – 01:43:30

Gambar 6 Bentuk sinyal rekaman a07 periode non- apnea durasi 15 detik

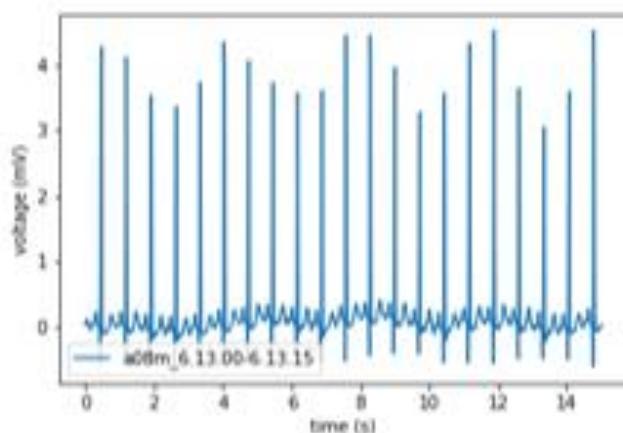


(a) Sampel waktu ke 05:33:00 – 05:33:15

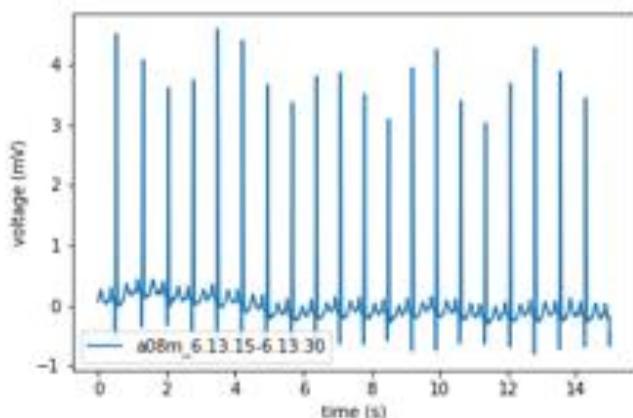


(b) Sampel waktu ke 05:33:15 – 05:33:30

Gambar 7 Bentuk sinyal rekaman a08 periode apnea durasi 15 detik

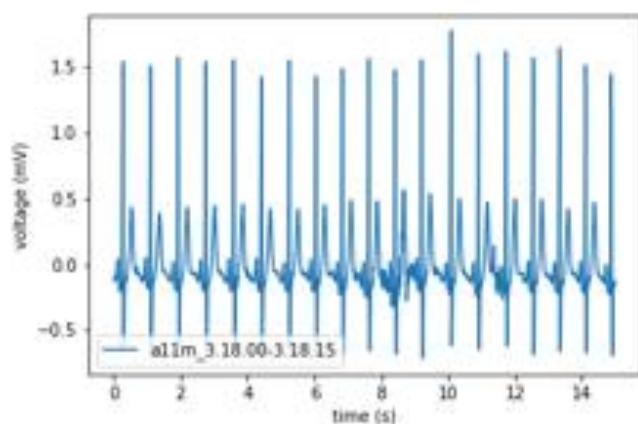


(a) Sampel waktu ke 06:13:00 – 06:13:15

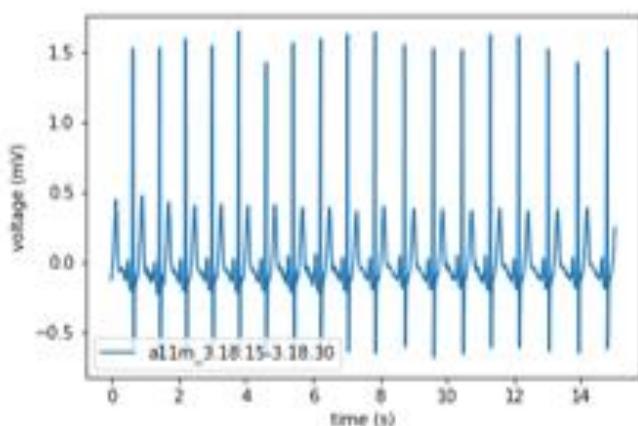


(b) Sampel waktu ke 06:13:15 – 06:13:30

Gambar 8 Bentuk sinyal rekaman a08 periode non-apnea durasi 15 detik

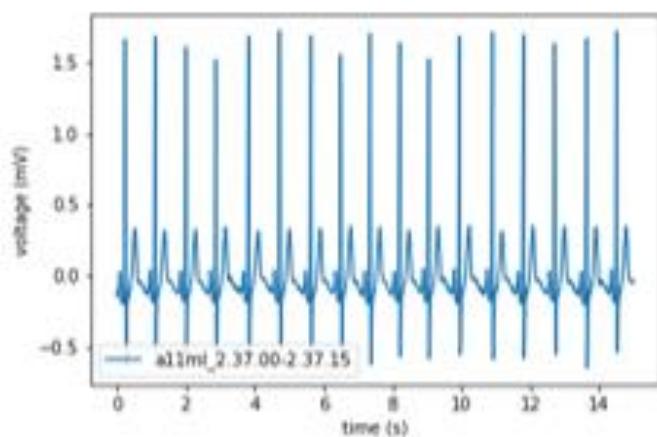


(a) Sampel waktu ke 03:18:00 – 03:18:15

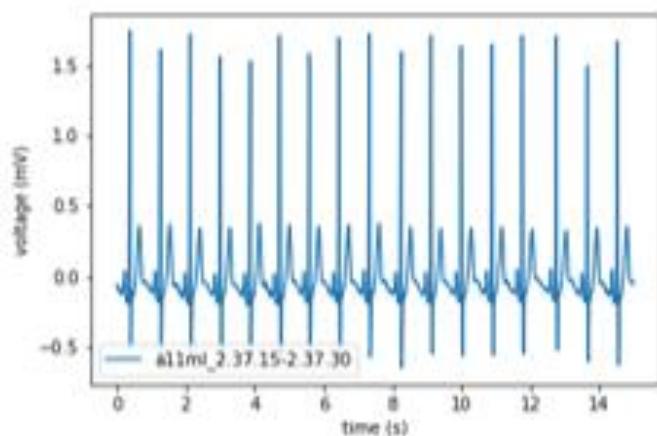


(b) Sampel waktu ke 03:18:15 – 03:18:30

Gambar 9 Bentuk sinyal rekaman all periode apnea durasi 15 detik.

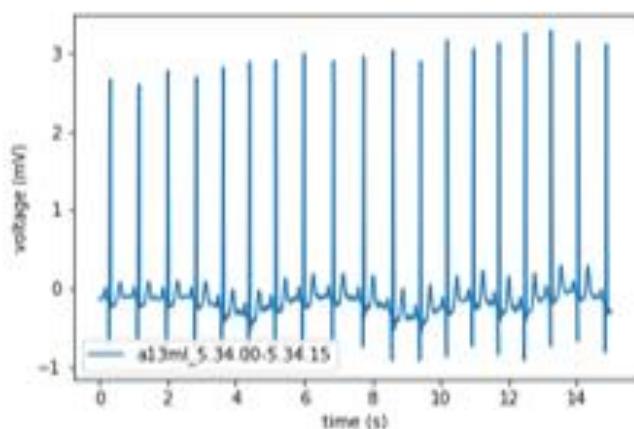


(a) Sampel waktu ke 02:37:00 – 02:37:15

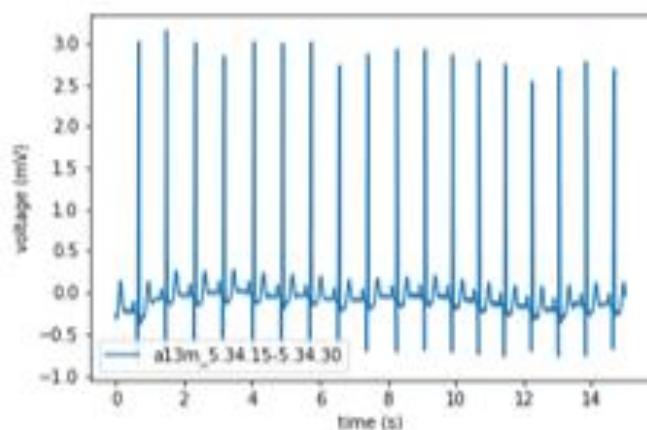


(b) Sampel waktu ke 02:37:15 – 02:37:30

Gambar 10 Bentuk sinyal rekaman a11 periode non-apnea durasi 15 detik

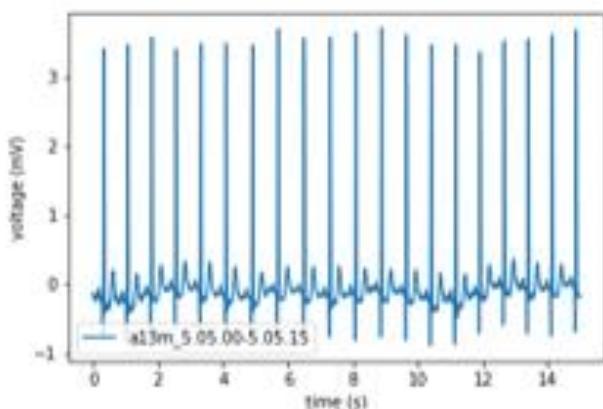


(a) Sampel waktu ke 05:34:00 – 05:34:15

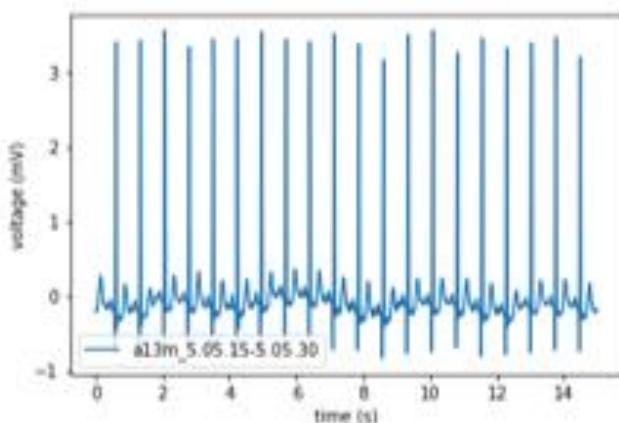


(b) Sampel waktu ke 05:34:15 – 05:34:30

Gambar 11 Bentuk sinyal rekaman a13 periode apnea durasi 15 detik

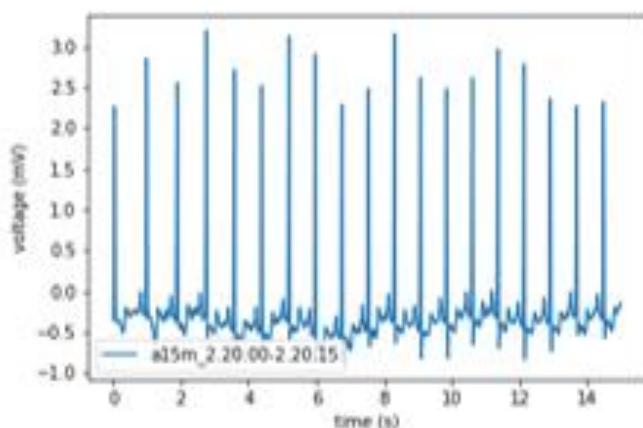


(a) Sampel waktu ke 05:05:00 – 05:05:15

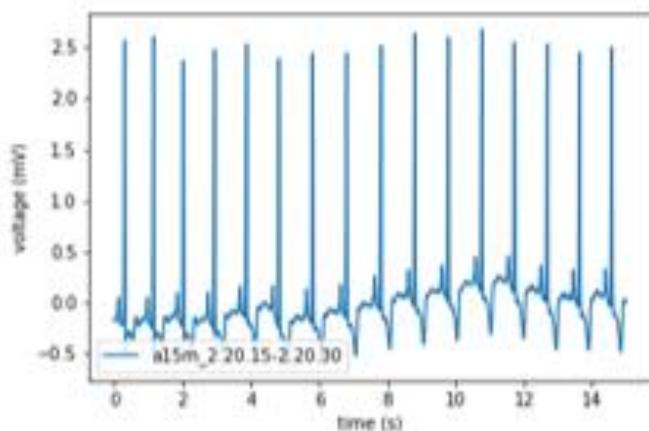


(b) Sampel waktu ke 05:05:15 – 05:05:30

Gambar 12 Bentuk sinyal rekaman a13 periode non-agnea durasi 15 detik

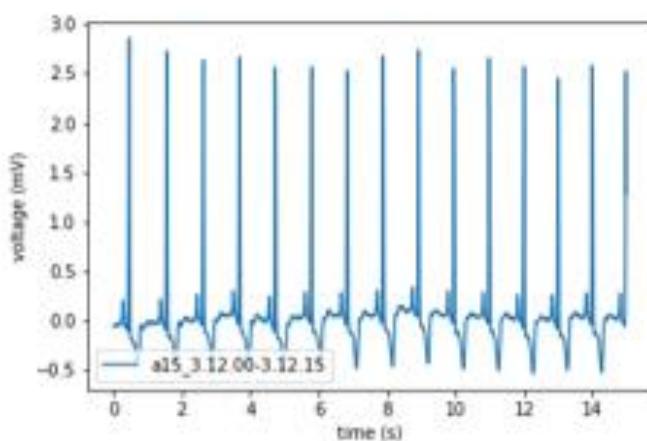


(a) Sampel waktu ke 02:20:00 – 02:20:15

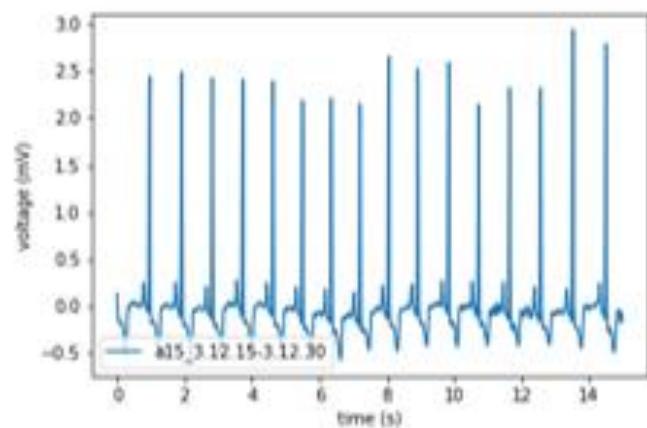


(b) Sampel waktu ke 02:20:15 – 02:20:30

Gambar 13 Bentuk sinyal rekaman a15 periode apnea durasi 15 detik

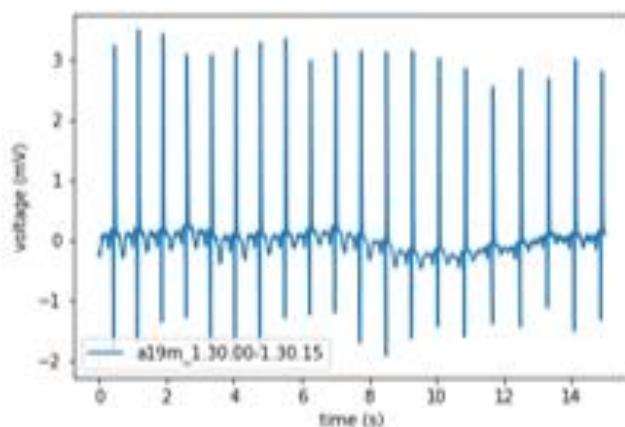


(a) Sampel waktu ke 03:12:00 – 03:12:15

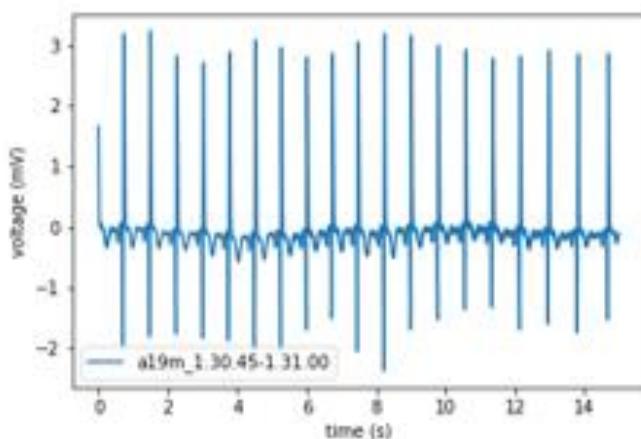


(b) Sampel waktu ke 03:12:15 – 03:12:30

Gambar 14 Bentuk sinyal rekaman a15 periode non-apnea durasi 15 detik

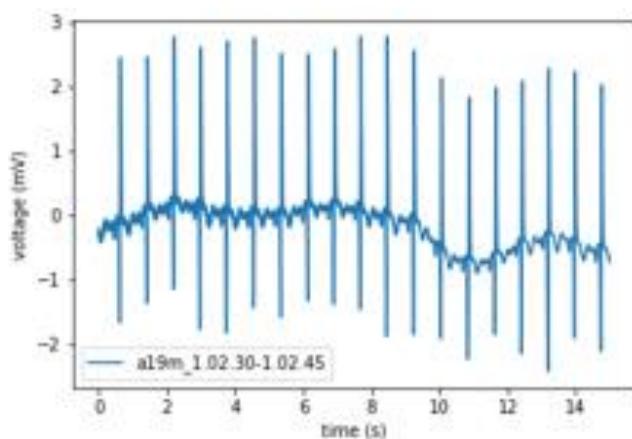


(a) Sampel waktu ke 01:30:00 – 01:30:15

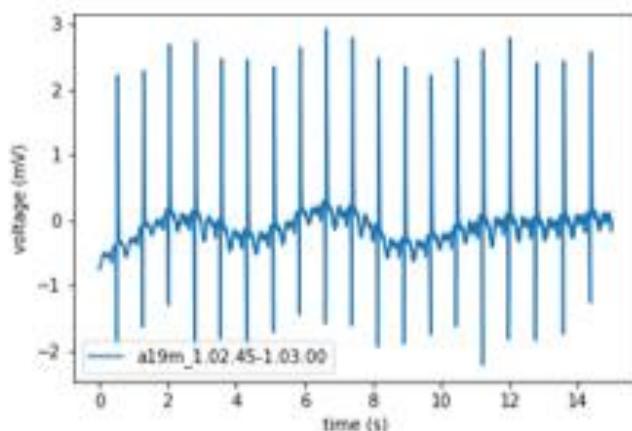


(b) Sampel waktu ke 01:30:45 – 01:31:00

Gambar 15 Bentuk sinyal rekaman a19 periode apnea durasi 15 detik

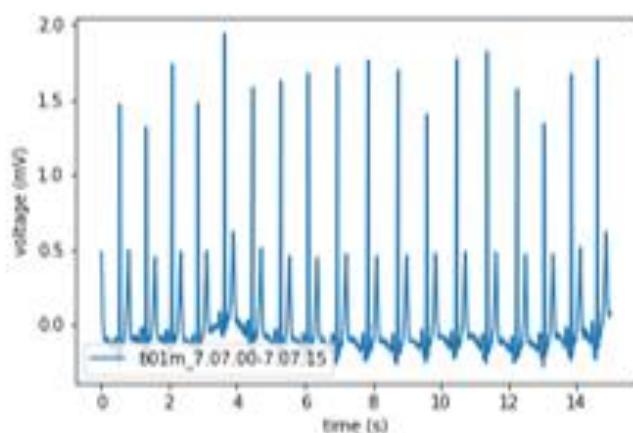


(a) Sampel waktu ke 01:02:30 – 01:02:45

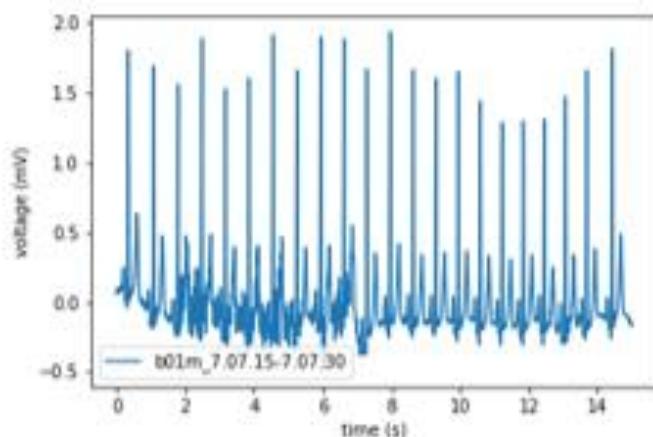


(b) Sampel waktu ke 01:02:45 – 01:03:00

Gambar 16 Bentuk sinyal rekaman a19 periode non-apnea durasi 15 detik

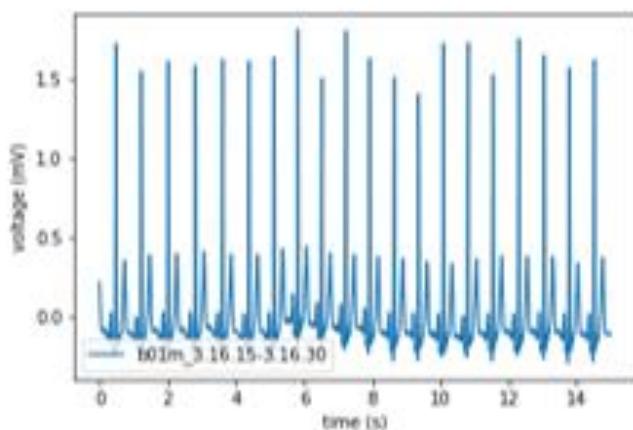


(a) Sampel waktu ke 07:07:00 – 07:07:15

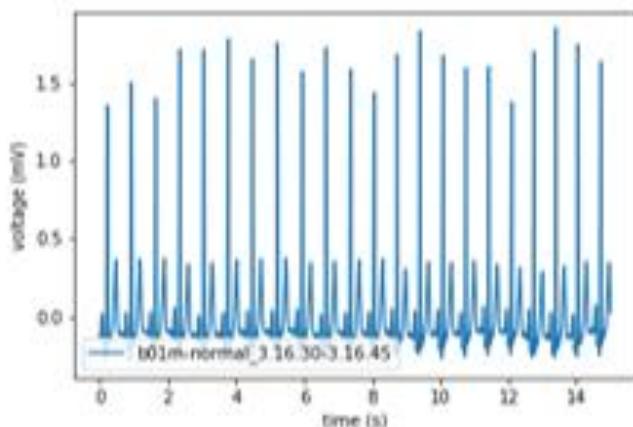


(b) Sampel waktu ke 07:07:15 – 07:07:30

Gambar 17 Bentuk sinyal rekaman b01 periode apnea durasi 15 detik

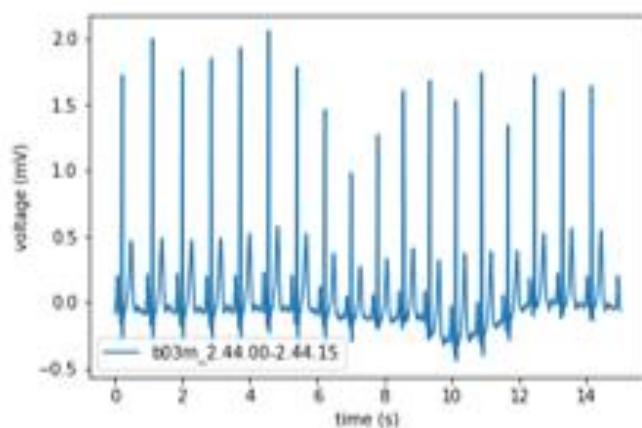


(a) Sampel waktu ke 03:16:15 – 03:16:30

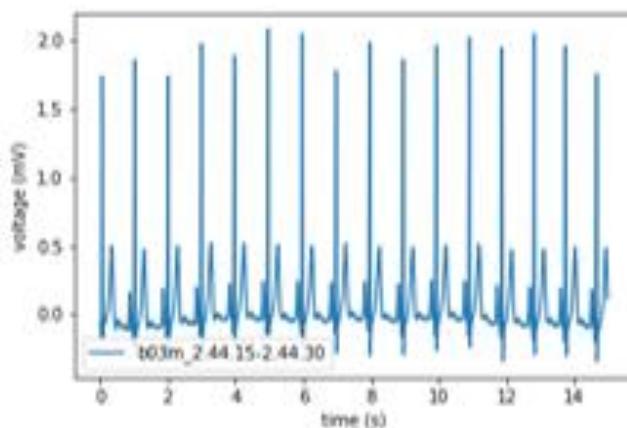


(b) Sampel waktu ke 03:16:30 – 03:16:45

Gambar 18 Bentuk sinyal rekaman b01 periode non-apnea durasi 15 detik

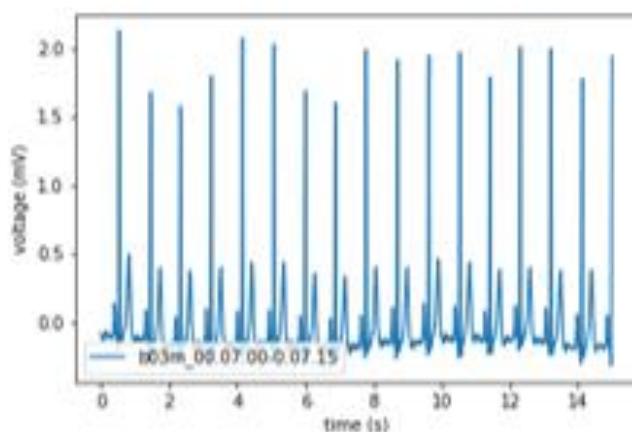


(a) Sampel waktu ke 02:44:00 – 02:44:15

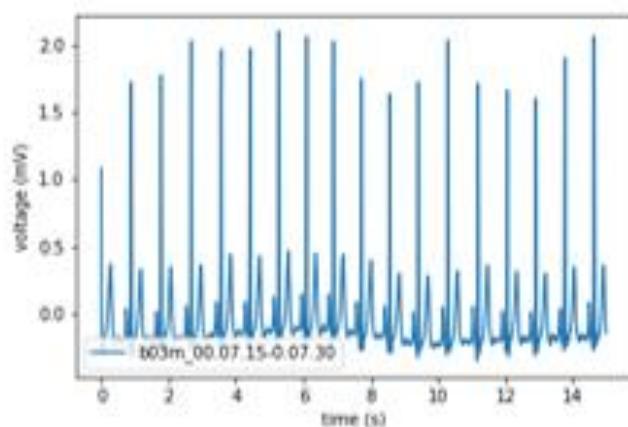


(b) Sampel waktu ke 02:44:15 – 02:44:30

Gambar 19 Bentuk sinyal rekaman b03 periode apnea durasi 15 detik

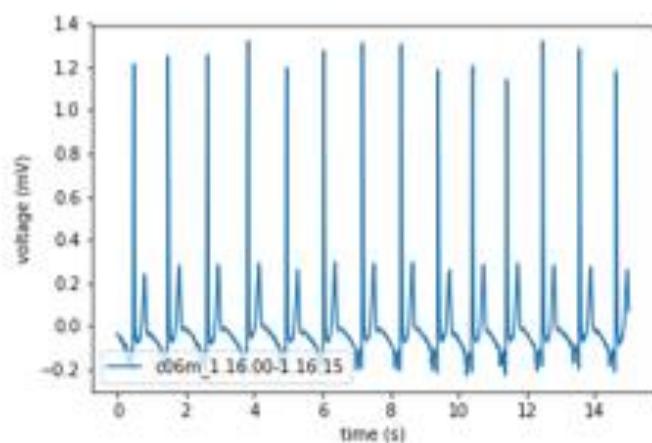


(a) Sampel waktu ke 00:07:00 – 00:07:15

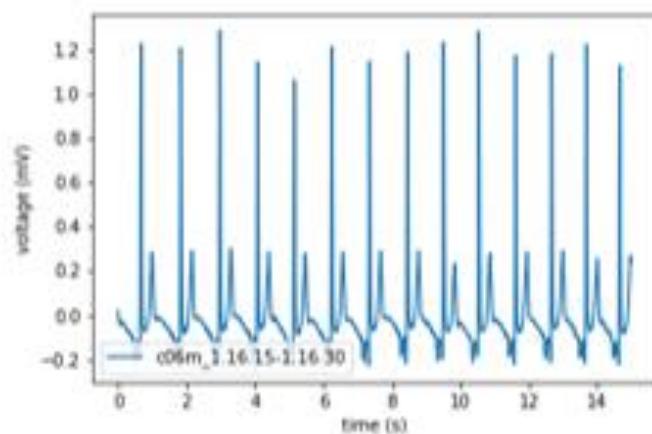


(b) Sampel waktu ke 00:07:15 – 00:17:30

Gambar 20 Bentuk sinyal rekaman b03 periode non-apnea durasi 15 detik

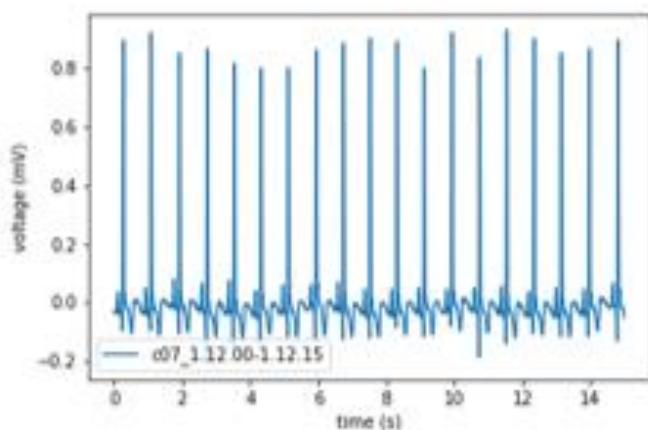


(a) Sampel waktu ke 01:16:00 – 01:16:15

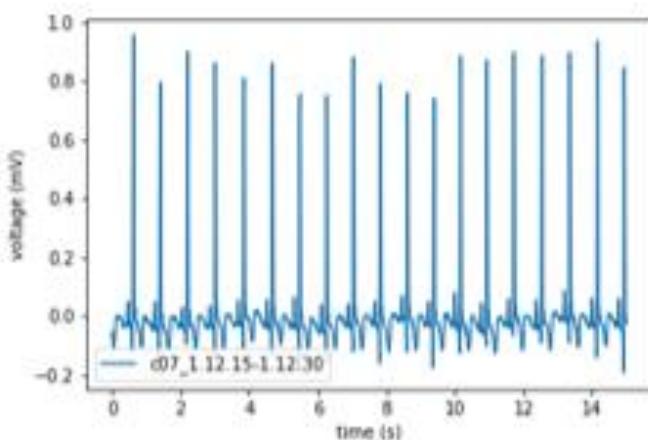


(b) Sampel waktu ke 01:16:15 – 01:16:30

Gambar 21 Bentuk sinyal rekaman c06 periode non-apnea durasi 15 detik



(a) Sampel waktu ke 01:12:00 – 01:12:15



(b) Sampel waktu ke 01:12:15 – 01:12:30

Gambar 22 Bentuk sinyal rekaman c07 periode non-apnea durasi 15 detik

Lampiran II.

Tingkat Keberhasilan Klasifikasi pada Komposisi Data yang Berbeda

Tabel 13 Tingkat keberhasilan klasifikasi pada komposisi data Apnea:Non 40:60

Komposisi	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
Apnea:Non 40:60	1	74.67%	40.33%	95.33%
	2	68.67%	45.87%	98.67%
	3	65.43%	43.67%	97.67%
	4	72.76%	37.33%	93.67%
	5	70.33%	55.00%	96.33%
Rata-rata		70.37%	44.44%	96.33%

Tabel 14 Tingkat keberhasilan klasifikasi pada komposisi data Apnea:Non 50:50

Komposisi	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
Apnea:Non 50:50	1	80%	76.67%	83.33%
	2	88.33%	90%	86.67%
	3	83.33%	76.67%	86.67%
	4	80%	73.33%	86.67%
	5	86.67%	80%	93.33%
Rata-rata		83.67%	79.33%	87.33%

Tabel 15 Tingkat keberhasilan klasifikasi pada komposisi data Apnea:Non 60:40

Komposisi	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
Apnea:Non 60:40	1	66.05%	94.00%	50.33%
	2	78.05%	95.33%	46.33%
	3	85.87%	95.00%	43.67%
	4	75.83%	92.67%	57.33%
	5	59.05%	95.00%	65.33%
Rata-rata		72.97%	94.40%	52.60%

Lampiran III.

Tingkat Keberhasilan Klasifikasi pada Durasi Segmentasi yang Berbeda

Tabel 16 Tingkat keberhasilan klasifikasi durasi segmentasi 10 detik

Durasi	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
10 detik	1	73.73%	60.42%	83.33%
	2	68.64%	45.83%	68.75%
	3	74.43%	43.75%	97.91%
	4	72.34%	47.08%	93.55%
	5	80.23%	85.42%	66.25%
Rata-rata		73.87%	56.50%	81.96%

Tabel 17 Tingkat keberhasilan klasifikasi durasi segmentasi 15 detik

Durasi	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
15 detik	1	80%	76.67%	83.33%
	2	88.33%	90%	86.67%
	3	83.33%	76.67%	86.67%
	4	80%	73.33%	86.67%
	5	86.67%	80%	93.33%
Rata-rata		83.67%	79.33%	87.33%

Tabel 18 Tingkat keberhasilan klasifikasi durasi segmentasi 20 detik

Durasi	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
20 detik	1	85.00%	83.07%	86.92%
	2	78.56%	84.33%	72.39%
	3	81.15%	76.15%	96.15%
	4	77.80%	55.01%	100.00%
	5	81.69%	88.21%	55.13%
Rata-rata		80.84%	77.35%	82.12%

Tabel 19 Tingkat keberhasilan klasifikasi durasi segmentasi 25 detik

Durasi	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
25 detik	1	87.08%	89.16%	85.00%
	2	82.50%	86.67%	78.33%
	3	75.41%	40.00%	95.83%
	4	75.83%	50.83%	95.83%
	5	69.58%	80.83%	11.33%
Rata-rata		78%	69.50%	73.26%

Tabel 20 Tingkat keberhasilan klasifikasi durasi segmentasi 30 detik

Durasi	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
30 detik	1	81.00%	73.73%	73.73%
	2	90.00%	68.64%	68.64%
	3	80.33%	71.43%	71.43%
	4	70.00%	62.34%	62.34%
	5	76.33%	70.23%	70.23%
Rata-rata		79.5%	69.27%	69.27%

Lampiran IV.

Tingkat Keberhasilan Klasifikasi menggunakan Kernel SVM yang berbeda

Tabel 21 Tingkat keberhasilan klasifikasi menggunakan kernel Linear

Kernel	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
Linear	1	80%	76.67%	83.33%
	2	88.33%	90%	86.67%
	3	83.33%	76.67%	86.67%
	4	80%	73.33%	86.67%
	5	86.67%	80%	93.33%
Rata-rata		83.67%	79.33%	87.33%

Tabel 22 Tingkat keberhasilan klasifikasi menggunakan kernel RBF

Kernel	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
RBF	1	75.00%	93.33%	56.67%
	2	50.00%	63.33%	36.67%
	3	56.67%	63.33%	50.00%
	4	83.33%	73.33%	93.33%
	5	71.67%	86.67%	56.67%
Rata-rata		67.33%	76 %	58.67%

Tabel 23 Tingkat keberhasilan klasifikasi menggunakan kernel Polinomial derajat 3

Kernel	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
Polinomial derajat 3	1	71.67%	73.33%	70.00%
	2	76.67%	83.337%	70.00%
	3	86.67%	73.33%	100%
	4	71.67%	70.00%	73.33%
	5	80.00%	66.67%	93.33%
Rata-rata		77.33%	73.33%	81.33%

Tabel 24 Tingkat keberhasilan klasifikasi menggunakan kernel Sigmoid

Kernel	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
Sigmoid	1	50%	100%	0%
	2	50%	100%	0%
	3	50%	100%	0%
	4	50%	100%	0%
	5	50%	100%	0%
Rata-rata		50%	100%	0%

Lampiran V.

Hasil uji analisis statistik ANOVA

Tabel 25 *Analysis Case Processing Summary*

Unweighted Cases		N	Percent
Valid		285	100,0
Excluded	Missing or out-of-range group codes	0	,0
	At least one missing discriminating variable	0	,0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	,0
	Total	0	,0
Total		285	100,0

Table 26 *Tests of Equality of Group Means*

	Wilks' Lambda	F	df1	df2	Sig.
VAR00001	,829	58,404	1	283	,000
VAR00002	,909	28,217	1	283	,000
VAR00003	,835	56,075	1	283	,000
VAR00004	,899	31,647	1	283	,000
VAR00005	,878	39,181	1	283	,000
VAR00006	,899	31,800	1	283	,000
VAR00007	,886	36,528	1	283	,000
VAR00008	,992	2,388	1	283	,123
VAR00009	,982	5,106	1	283	,025
VAR00010	,712	114,482	1	283	,000
VAR00011	,834	56,510	1	283	,000

Summary of Canonical Discriminant Functions

Tabel 27 *Wilks' Lambda*

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	,626	130,140	11	,000

Tabel 28 *Functions at Group Centroids*

	Function
VAR00012	1
,00	-,774
1,00	,768

Lampiran VI.

Tingkat Keberhasilan Masing-masing Percobaan terhadap Pengurangan Fitur

A. Pengujian pertama

Tabel 29 Hasil akurasi tiap percobaan menggunakan 10 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
10	1	80%	76.67%	83.33%
	2	88.33%	90%	86.67%
	3	83.33%	76.67%	86.67%
	4	80%	73.33%	86.67%
	5	86.67%	80%	93.33%
Rata-rata		83.67%	79.33%	87.33%

Tabel 30 Hasil akurasi tiap percobaan menggunakan 9 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
9	1	85%	80%	90%
	2	85%	90%	80%
	3	83.33%	66.67%	100%
	4	82.67%	69.67%	96.67%
	5	86.67%	80%	93.33%
Rata-rata		85.33%	76.67%	92%

Tabel 31 Hasil akurasi tiap percobaan menggunakan 8 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
8	1	80.00%	76.67%	83.33%
	2	80.00%	80.00%	80.00%
	3	75.00%	53.33%	96.67%
	4	78.33%	76.67%	80.00%
	5	88.33%	86.67%	90.00%
Rata-rata		80.33%	74.67%	86.00%

Tabel 32 Hasil akurasi tiap percobaan menggunakan 7 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
7	1	80.00%	76.67%	83.33%
	2	81.67%	83.33%	80.00%
	3	76.67%	56.67%	96.67%
	4	76.67%	73.33%	80.00%
	5	88.33%	90.00%	86.67%
Rata-rata		80.67%	76.00%	85.33%

Tabel 33 Hasil akurasi tiap percobaan menggunakan 6 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
6	1	80.00%	76.67%	83.33%
	2	78.33%	76.67%	80.00%
	3	76.67%	56.67%	96.67%
	4	76.67%	73.33%	80.00%
	5	88.33%	90.00%	86.67%
Rata-rata		80.00%	74.67%	85.33%

Tabel 34 Hasil akurasi tiap percobaan menggunakan 5 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
5	1	80.00%	76.67%	83.33%
	2	80.00%	80.00%	80.00%
	3	75.00%	56.67%	96.67%
	4	78.33%	73.33%	83.33%
	5	86.67%	86.67%	86.67%
Rata-rata		80.00%	74.67%	86.00%

Tabel 35 Hasil akurasi tiap percobaan menggunakan 4 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
4	1	80.00%	76.67%	83.33%
	2	78.33%	76.67%	80.00%
	3	73.33%	50.00%	96.67%
	4	81.67%	76.67%	86.67%
	5	93.33%	96.67%	90.00%
Rata-rata		81.33%	75.34%	87.33%

Tabel 36 Hasil akurasi tiap percobaan menggunakan 3 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
3	1	80.00%	76.67%	83.33%
	2	78.33%	73.33%	83.33%
	3	73.33%	50.00%	96.67%
	4	83.33%	80.00%	86.67%
	5	93.33%	96.67%	90.00%
Rata-rata		81.66%	75.33%	88.00%

Tabel 37 Hasil akurasi tiap percobaan menggunakan 2 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
2	1	80.00%	76.67%	83.33%
	2	78.33%	66.67%	90.00%
	3	75.00%	50.00%	100.00%
	4	78.33%	73.33%	83.33%
	5	91.67%	96.67%	86.67%
Rata-rata		80.67%	72.67%	88.67%

Tabel 38 Hasil akurasi tiap percobaan menggunakan 1 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
1	1	80.00%	76.67%	83.33%
	2	75.00%	56.00%	93.33%
	3	73.33%	46.67%	100.00%
	4	76.67%	63.33%	90.00%
	5	91.67%	90.00%	93.33%
Rata-rata		79.33%	66.53%	92.00%

B. Pengujian kedua

Tabel 39 Hasil akurasi tiap percobaan menggunakan 10 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
10	1	80%	76.67%	85%
	2	88%	90%	85%
	3	83.67%	76.67%	83.67%
	4	80%	73.33%	82,33%
	5	86.33%	80%	86.67%
Rata-rata		82.67%	79.33%	85.33%

Tabel 40 Hasil akurasi tiap percobaan menggunakan 9 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
9	1	83.33%	83.33%	83.33%
	2	85.00%	90.00%	80.00%
	3	80.00%	63.33%	96.67%
	4	75.00%	73.33%	76.67%
	5	83.33%	73.33%	93.33%
Rata-rata		81.33%	76.66%	86.00%

Tabel 41 Hasil akurasi tiap percobaan menggunakan 8 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
8	1	83.33%	76.67%	90.00%
	2	83.33%	86.67%	80.00%
	3	80.00%	63.33%	96.67%
	4	81.67%	80.00%	83.33%
	5	83.33%	73.33%	93.33%
Rata-rata		82.33%	76.00%	88.67%

Tabel 42 Hasil akurasi tiap percobaan menggunakan 7 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
7	1	80.00%	73.33%	86.67%
	2	83.33%	86.67%	80.00%
	3	80.00%	63.33%	96.67%
	4	76.67%	73.33%	80.00%
	5	80.00%	63.33%	96.67%
Rata-rata		80.00%	72.00%	88.00%

Tabel 43 Hasil akurasi tiap percobaan menggunakan 6 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
6	1	81.67%	76.67%	86.67%
	2	83.33%	86.67%	80.00%
	3	80.00%	63.33%	96.67%
	4	76.67%	73.33%	80.00%
	5	78.33%	60.00%	96.67%
Rata-rata		80.00%	72.00%	88.00%

Tabel 44 Hasil akurasi tiap percobaan menggunakan 5 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
5	1	80.00%	73.33%	86.67%
	2	83.33%	86.67%	80.00%
	3	80.00%	63.33%	96.67%
	4	76.67%	73.33%	80.00%
	5	80.00%	63.33%	96.67%
Rata-rata		80.00%	72.00%	88.00%

Tabel 45 Hasil akurasi tiap percobaan menggunakan 4 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
4	1	80.00%	73.33%	86.67%
	2	83.33%	86.67%	80.00%
	3	80.00%	63.33%	96.67%
	4	76.67%	73.33%	80.00%
	5	80.00%	63.33%	96.67%
Rata-rata		80.00%	72.00%	88.00%

Tabel 46 Hasil akurasi tiap percobaan menggunakan 3 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
3	1	76.67%	73.33%	80.00%
	2	81.67%	83.33%	80.00%
	3	80.00%	63.33%	96.67%
	4	81.67%	83.33%	80.00%
	5	81.67%	66.67%	96.67%
Rata-rata		80.34%	74.00%	86.67%

Tabel 47 Hasil akurasi tiap percobaan menggunakan 2 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
2	1	98.33%	100.00%	96.67%
	2	50.00%	10.00%	90.00%
	3	58.33%	20.00%	96.67%
	4	50.00%	100.00%	0.00%
	5	50.00%	100.00%	0.00%
Rata-rata		61.33%	66.00%	56.67%

Tabel 48 Hasil akurasi tiap percobaan menggunakan 1 fitur

Jumlah Fitur	Percobaan	<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
1	1	63.33%	30.00%	96.67%
	2	50.00%	10.00%	90.00%
	3	58.33%	20.00%	96.67%
	4	48.33%	3.00%	93.33%
	5	55.00%	10.00%	100.00%
Rata-rata		55.00%	14.60%	95.33%

BIODATA PENULIS



Desanti Nurma Risalah, lahir pada 1 Desember 1996 berasal dari kota kecil Tulungagung, Jawa Timur. Penulis lulus dari SMP Islam Al Azhaar pada tahun 2012 kemudian melanjutkan pendidikan ke MAN 2 Tulungagung melalui program akselerasi hingga akhirnya lulus pada tahun 2014. Tertarik dengan dinamika perkembangan Teknologi, penulis kemudian melanjutkan pendidikan Strata satu ke Jurusan Teknik Multimedia dan Jaringan, FTI-ITS yang kemudian berganti nama menjadi Teknik Komputer, FTE-ITS Surabaya. Pada tahun ketiga penulis mengambil bidang studi Telematika. Saat di kuliah penulis aktif menjadi asisten laboratorium B401. Selain itu, penulis juga aktif mengikuti organisasi dimulai dengan menjadi staff Departemen Dalam Negeri BEM FTI ITS periode 2015/2016, dan staff divisi Edukasi UKM Click ITS pada periode yang sama. Pada tahun berikutnya, penulis menjadi Sekretaris I BEM FTI-ITS periode 2016/2017, sekaligus menjadi staff divisi Eksebsi UKM Click ITS pada periode yang sama. Selain itu penulis juga aktif dalam berbagai kepanitiaan seperti sebagai sie acara GERIGI ITS 2015, Koor administrasi TRF 2016, dan bagian dari event MAGE 2016 hingga 2017, dll. Pada tahun keempatnya, penulis bersama tim merintis sebuah *startup* dibidang *IoT* yang rencananya akan masuk menjadi bagian dari anak perusahaan PT. ITS Tekno Sains.

Halaman ini sengaja dikosongkan