



TESIS - EE185401

**PENGENALAN EMOSI MARAH PADA REMAJA
KETIKA BERMAIN VIDEO GAME BERBASIS SINYAL
PULSE RATE MENGGUNAKAN KLASIFIKASI
*SUPPORT VEKTOR MACHINE***

DHANU KURNIA UTAMA
07111550060002

DOSEN PEMBIMBING
Dr. Adhi Dharma Wibawa, S.T., M.T.
Prof. Dr. Ir. Mauridhi Hery Purnomo, M.Eng.

PROGRAM MAGISTER
BIDANG KEAHLIAN TELEMATIKA
DEPARTEMEN TEKNIK ELEKTRO
FAKULTAS TEKNOLOGI ELEKTRO DAN INFORMATIKA CERDAS
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2020



TESIS - EE185401

**PENGENALAN EMOSI MARAH PADA REMAJA
KETIKA BERMAIN VIDEO GAME BERBASIS SINYAL
PULSE RATE MENGGUNAKAN KLASIFIKASI
*SUPPORT VEKTOR MACHINE***

DHANU KURNIA UTAMA
07111550060002

DOSEN PEMBIMBING
Dr. Adhi Dharma Wibawa, S.T., M.T.
Prof. Dr. Ir. Mauridhi Hery Purnomo, M.Eng.

PROGRAM MAGISTER
BIDANG KEAHLIAN TELEMATIKA
DEPARTEMEN TEKNIK ELEKTRO
FAKULTAS TEKNOLOGI ELEKTRO DAN INFORMATIKA CERDAS
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2020

LEMBAR PENGESAHAN TESIS

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar

Magister Teknik (MT)

di

Institut Teknologi Sepuluh Nopember

Oleh:

DHANU KURNIA UTAMA

NRP: 07111550060002

Tanggal Ujian: 2 Juli 2020

Periode Wisuda: September 2020

Disetujui oleh:

Pembimbing:

1. Dr. Adhi Dharma Wibawa, S.T., M.T
NIP: 197605052008121003

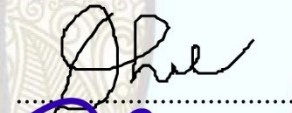


2. Prof. Dr. Ir. Mauridhi Hery Purnomo, M.Eng.
NIP: 195809161986011001



Penguji:

1. Dr. Diah Puspito Wulandari, S.T., M.Sc.
NIP: 198012192005012001



2. Eko Setijadi, ST.,MT.,Ph.D
NIP: 197210012003121002



3. Dr.Eng. Chastine Fatichah S.Kom, M.Kom
NIP: 197512202001122002



Kepala Departemen Teknik Elektro
Fakultas Teknologi Elektro dan Informatika Cerdas



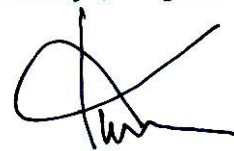
Dedet Candra Riawan, ST., M.Eng., Ph.D
NIP: 197311192000031001

PERNYATAAN KEASLIAN TESIS

Dengan ini saya menyatakan bahwa isi keseluruhan Tesis saya dengan judul **“PENGENALAN EMOSI MARAH PADA REMAJA KETIKA BERMAIN VIDEO GAME BERBASIS SINYAL *PULSE RATE* MENGGUNAKAN KLASIFIKASI *SUPPORT VEKTOR MACHINE*”** adalah benar-benar hasil karya intelektual mandiri, diselesaikan tanpa menggunakan bahan-bahan yang tidak diijinkan dan bukan merupakan 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, 4 Agustus 2020



Dhanu Kurnia Utama

NRP. 07111550060002

**PENGENALAN EMOSI MARAH PADA REMAJA KETIKA
BERMAIN VIDEO GAME BERBASIS SINYAL *PULSE RATE*
MENGUNAKAN KLASIFIKASI *SUPPORT VEKTOR*
*MACHINE***

Nama mahasiswa : Dhanu Kurnia Utama
NRP : 07111550060002
Pembimbing : 1. Dr. Adhi Dharma Wibawa, S.T., M.T.
2. Prof. Dr. Ir. Mauridhi Hery Purnomo, M.Eng.

ABSTRAK

Remaja pada saat sekarang tumbuh dalam era *video game* baik *online* maupun *offline*, sedangkan permainan *video game* lebih bersifat emosional dari pada rasional, hal ini disebabkan oleh protokol algoritma *game* yang dibuat dapat menyesuaikan terhadap perilaku pemain. Parameter seperti *reward*, tantangan baru dan bertingkat, bersosial, dan tren merupakan sedikit dari banyak cara agar *video game* menjadi lebih menarik pengguna. Salah satu akibat yang sering terjadi dari pengaruh penarik minat tersebut adalah emosi marah pada *gamer*, sedangkan emosi marah merupakan salah satu emosi negatif bagi manusia yang dapat menyebabkan kesehatan yang buruk.

Oleh karena itu, penelitian ini mencoba untuk memberikan informasi tentang pola sinyal *pulse rate* pada pemain ketika kemarahan terjadi selama bermain *video game* yang telah ditentukan. *Video game* yang dipilih telah divalidasi menggunakan survei ke komunitas pengembang *video game* sebagai *video game* yang dapat memiliki emosi marah sebagai emosi yang dominan. Data *pulse rate* dari setiap subjek memiliki dua fase: fase *baseline (pre-stimulation)* dan fase bermain (*in-stimulation*). Parameter statistik yang digunakan untuk mendapatkan pola *pulse rate* adalah *difference value of mean* dan *absolute deviation of mean*. Selain itu pengenalan emosi marah dilakukan dengan cara melakukan prediksi emosi marah dari fitur data sinyal *pulse rate* yang direkam saat bermain *video game*. Hal ini bertujuan untuk dapat memberikan informasi tentang tingkat kebenaran dari data *pulse rate* yang diasumsikan sebagai data representasi emosi marah. Metode yang digunakan dengan menerapkan algoritma klasifikasi *Support Vektor Machine (SVM)*. Terdapat permasalahan dalam data yang diperoleh pada metode klasifikasi, yaitu ketidakseimbangan antara kelas *baseline* dan *playing*. Penyelesaian permasalahan *imbalanced class* yang diusulkan menggunakan pendekatan algoritma *Bagging (bootstrap agregation)* dengan membangun beberapa model dengan algoritma yang sama (*multiple classifier*) dari sub-sampel yang berbeda.

Hasil yang didapatkan, klasifikasi menggunakan metode *Bagging* yang diusulkan mempunyai performansi dengan skor akurasi sebesar 82.98%, *Weighted Averages* dari skor PPV dan NPV sebesar 77.7%, *Weighted Averages* dari skor TPR dan TNR sebesar 83%, dan *Weighted Averages* dari skor FPR dan FNR sebesar 76.2%. Dari hasil tersebut menunjukkan bahwa prediksi emosi marah menggunakan metode klasifikasi dapat mengenali emosi marah pada remaja ketika bermain *video game* berbasis sinyal *pulse rate*.

Kata kunci: Pengenalan Emosi Marah, *Video Game*, Visualisasi Pola, Algoritma SVM, Metode *Bagging*

Halaman ini sengaja dikosongkan

RECOGNITION OF ANGRY EMOTION IN ADOLESCENTS WHEN PLAYING VIDEO GAMES BASED ON PULSE RATE SIGNALS USING CLASSIFICATION OF SUPPORT VECTOR MACHINE

By : Dhanu Kurnia Utama
Student Identity Number : 07111550060002
Supervisor(s) : 1. Dr. Adhi Dharma Wibawa, S.T., M.T.
2. Prof. Dr. Ir. Mauridhi Hery Purnomo, M.Eng.

ABSTRACT

Teenagers are now growing up in the era of video games both online and offline, while video game play is more emotional than rational, this is due to the game algorithm protocol that is made to adapt to player behavior. Parameters such as rewards, new and stratified challenges, social, and trends are just a few of the ways that games become more attractive to users. One consequence that often occurs from the influence of attracting interest is angry emotions on gamers, while angry emotions are one of the negative emotions for humans that can cause poor health.

Therefore, this study tries to provide information about the pulse rate signal patterns in players when anger occurs during playing video games that have been determined. The selected game has been validated using a survey to the video game developer community as a video game that can have angry emotions as dominant emotions. The pulse rate data for each subject has two phases: the baseline phase (pre-stimulation) and the play phase (in-stimulation). The statistical parameters used to obtain the pulse pattern are the difference value of mean and absolute deviation of mean. Besides the introduction of angry emotions is done by predicting angry emotions from the pulse rate signal feature data that is recorded while playing video games. It aims to be able to provide information about the level of truth of the pulse rate data which is assumed to be data representation of angry emotions. The method used by applying the Support Vector Machine (SVM) classification algorithm. There are problems in the data obtained in the classification method, namely the imbalance between the baseline and playing classes. Solution to the imbalanced class problem is proposed using the Bagging algorithm (bootstrap aggregation) approach by building several models with the same algorithm (multiple classifier) from different sub-samples.

The results obtained, the classification using the proposed Bagging method has performance with an accuracy score of 82.98%, Weighted Averages of the PPV and NPV scores of 77.7%, Weighted Averages of the TPR and TNR scores of 83%, and Weighted Averages of the FPR and FNR scores of 76.2%. From these results, it shows that prediction of angry emotions using the classification method can identify angry emotions in adolescents when playing video games based on pulse rate signals.

Keywords: Recognition of Angry Emotions, Video Games, Pattern Visualization, SVM Algorithm, Bagging Method

Halaman ini sengaja dikosongkan

KATA PENGANTAR

Alhamdulillahirobbil'alamin, segala puji dan syukur penulis panjatkan kepada Allah SWT, atas segala limpahan rahmat dan karunia-Nya yang telah memberikan nikmat kesehatan dan hikmat kepada penulis sehingga tesis ini dapat diselesaikan dengan baik sesuai dengan waktu yang telah direncanakan. Shalawat serta salam saya persembahkan kepada Nabi Besar Muhammad SAW dan kepada sahabat Rasulullah serta semua pihak yang senantiasa istiqomah berjuang dalam Iman dan Islam.

Penulisan tesis dengan judul “Pengenalan Emosi Marah Pada Remaja Ketika Bermain *Video Game* Berbasis Sinyal *Pulse Rate* Menggunakan Klasifikasi *Support Vektor Machine*” ini sebagai syarat dalam rangka menyelesaikan studi magister di Departemen Teknik Elektro, Fakultas Teknologi Elektro dan Informatika Cerdas, Institut Teknologi Sepuluh Nopember.

Dalam penyusunan tesis ini, penulis menyadari bahwa penulisan ini masih jauh dari kesempurnaan, oleh karena itu penulis sangat mengharapkan saran dan kritik yang membangun yang berguna untuk penyempurnaan penulisan selanjutnya. Tak lupa penulis ingin mengucapkan terima kasih kepada:

1. Kedua orang tua tersayang, Bapak Sumbar A Restianto bin Rasyidjan dan Ibu Jumiati binti Taslam dan kedua mertua yang tersabar, Bapak Djaenuri dan Ibu Purwanti yang senantiasa memberikan dukungan, semangat, do'a dan harapan kepada penulis saat penulisan tesis.
2. Dek Norma Yunidya Wardhani, yakni istri tercinta, selaku bidadari bagi penulis yang selalu memberikan semangat, dukungan, do'a, dan bantuannya. Terimakasih karena selalu sabar mendengarkan keluh kesah penulis saat penyusunan tesis ini. Dan ananda tersayang, Falah Hamizan Dhandya Taqwa, yang selalu menjadi penghadir senyuman ketika sampai di rumah.
3. Lembaga Pengelola Dana Pendidikan (LPDP) atas beasiswa penuh yang telah diberikan.
4. Bapak Dr. I Ketut Eddy Purnama, ST., MT. selaku Dekan Fakultas Teknologi Elektro dan Informatika Cerdas, Institut Teknologi Sepuluh Nopember.

5. Bapak Dedet Candra Riawan, ST., M. Eng., Ph.D selaku Kepala Departemen Teknik Elektro Fakultas Teknologi Elektro dan Informatika Cerdas, Institut Teknologi Sepuluh Nopember.
6. Bapak Ronny Mardiyanto S.T., M.T., Ph.D., selaku Kepala Program Studi Teknik Elektro FTE – ITS.
7. Bapak Prof. Dr. Ir. Mauridhi Hery Purnomo, M.Eng. selaku supervisor kedua yang selalu sabar membimbing, memberikan semangat, saran dan masukan dalam ranah akademis maupun nonakademis.
8. Bapak Dr. Adhi Dharma Wibawa, S.T., M.T. selaku supervisor pertama sekaligus koordinator bidang program magister telematika, Departemen Teknik Elektro, FTE, ITS yang telah meluangkan waktu, tenaga dan pikiran serta memberikan perhatian dalam bimbingan dan pengarahan selama penulisan tesis ini. Terimakasih atas motivasi dan inspirasi yang luar biasa untuk penulis.
9. Bapak Dr. Ir. Wirawan, DEA, selaku mantan Kepala Program Studi Teknik Elektro FTE – ITS penulis yang sangat sabar dan senantiasa memberikan *wejangan*, motivasi, dan saran-saran akademis.
10. Seluruh dosen program magister program Telematika, Departemen Teknik Elektro FTE - ITS yang telah memberikan ilmunya kepada penulis selama menjalani studi program magister.
11. Seluruh staff Departemen Teknik Elektro FTE-ITS, mulai dari Staff Kemahasiswaan dan Akademik, yang dengan sabar membantu dan melayani penulis mulai dari pengajuan proposal hingga yudisium.

Kiranya isi tesis ini dapat bermanfaat dalam ilmu pengetahuan dan juga dapat dijadikan sebagai salah satu sumber referensi bagi peneliti selanjutnya yang berminat meneliti hal yang sama. Akhir kata semoga Allah SWT membalas segala langkah kebaikan yang kita lakukan dan menjadi amal baik kita. Aamiin.

Surabaya, 4 Agustus 2020

Penulis

DAFTAR ISI

LEMBAR PENGESAHAN TESIS.....	iii
PERNYATAAN KEASLIAN TESIS.....	v
ABSTRAK.....	vii
ABSTRACT.....	ix
KATA PENGANTAR.....	xi
DAFTAR ISI.....	xiii
DAFTAR GAMBAR.....	xvii
DAFTAR TABEL.....	xix
BAB 1 PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	3
1.3 Tujuan.....	4
1.4 Batasan Masalah.....	5
1.5 Kontribusi Penelitian.....	5
BAB 2 KAJIAN PUSTAKA.....	7
2.1 Kajian Penelitian Terkait Prosedur Pengambilan Data.....	7
2.2 Teori Dasar.....	7
2.2.1 Pengertian <i>Video Game</i>	7
2.2.2 Pengaruh Emosi Marah terhadap Kesehatan.....	10
2.2.3 Sinyal <i>Pulse Rate</i>	13
2.2.4 Ekstraksi Fitur <i>Pulse Rate</i>	16
2.2.5 Visualisasi Data.....	18
2.2.6 <i>Support Vector Machine (SVM)</i>	21
2.2.7 <i>K-Fold Cross Validation</i>	31
2.2.8 <i>Bagging (Bootsrap Aggregating)</i>	31
2.2.9 <i>Metode Synthetic Minority Oversampling Technique</i>	33

2.2.10	Evaluasi Performansi Metode Klasifikasi	33
BAB 3	METODE PENELITIAN	377
3.1	Tahap Awal	37
3.1.1	Studi Literatur	37
3.1.2	Instalasi Perangkat	38
3.1.3	Pemilihan Subjek Penelitian	38
3.2	Pemerolehan Data	39
3.2.1	Prosedur Pengambilan Data	39
3.2.2	<i>Pre-processing</i> Data	40
3.3	Pengolahan Data	41
3.3.1	Ekstraksi Fitur untuk Visualisasi dan Klasifikasi	42
3.3.2	Normalisasi Fitur untuk Klasifikasi	45
3.3.3	Visualisasi Pola <i>Pulse Rate</i>	45
3.3.4	Klasifikasi Menggunakan SVM	46
3.4	Tahap Akhir	50
3.4.1	Skenario Pengujian	50
3.4.2	Evaluasi Performansi	51
BAB 4	HASIL DAN PEMBAHASAN	53
4.1	Hasil Pengambilan Data	53
4.2	Hasil Pengolahan Data	55
4.2.1	Hasil Ekstraksi Fitur untuk Visualisasi	56
4.2.2	Hasil Ekstraksi Fitur untuk Klasifikasi	57
4.2.3	Hasil Normalisasi Fitur Visualisasi	58
4.3	Hasil Visualisasi Data <i>Pulse Rate</i>	59
4.3.1	Hasil Visualisasi Pada Parameter Difference Value Of Mean	59
4.3.2	Hasil Visualisasi Pada Parameter Absolute Deviation Of Mean	61
4.4	Hasil Klasifikasi Menggunakan SVM	63
4.4.1	Hasil Klasifikasi Algoritma SVM Menggunakan Metode Bagging	63
4.4.2	Hasil Klasifikasi Algoritma SVM Menggunakan Metode SMOTE	68

4.4.3 Hasil Klasifikasi Algoritma SVM dengan Class Attribute Per Fase	71
4.4.4 Analisa Hasil Klasifikasi Algoritma SVM	75
BAB 5 KESIMPULAN	77
5.1 Kesimpulan	77
5.2 Saran	78
DAFTAR PUSTAKA	81
LAMPIRAN	87
BIOGRAFI PENULIS	89

DAFTAR GAMBAR

Gambar 2.1 Sensor <i>Pulse oximeter Cooking Hacks</i> dan Cara Pemasangan pada Jari Tangan	13
Gambar 2.2 Contoh Diagram Garis (Line Charts)	19
Gambar 2.3 Contoh Diagram Batang (Bar Charts)	20
Gambar 2.4 Contoh Sebaran Plot (Scatter Graph)	21
Gambar 2.5 Klasifikasi SVM: (kiri) Klasifikasi Linear Separable; (tengah) Linear Nonseparable; (kanan) Nonlinier	22
Gambar 2.6 Bidang Pemisah Terbaik dengan Margin (d) Terbesar Linier separable	22
Gambar 2.7 Bidang Pemisah Terbaik dengan Margin (d) Terbesar Linier Non-Separable	26
Gambar 2.8 Pemetaan Ruang Data Dua Dimensi (kiri) ke Dalam Ruang Feature Tiga Dimensi (kanan)	28
Gambar 3.1 Diagram Alur Penelitian	37
Gambar 3.2 Gambaran Suasana Subjek pada Saat Pengambilan Data	39
Gambar 3.3 Diagram Alir pada Prosedur Pengambilan Data	40
Gambar 3.4 Blok Diagram pada <i>Window Segmentation</i> untuk Proses Klasifikasi	41
Gambar 3.5 Blok Diagram pada <i>Window Segmentation</i> untuk Proses Visualisasi	41
Gambar 3.6 Diagram Alir Untuk Analisa Pola Menggunakan Visualisasi	46
Gambar 3.7 Flowchart Usulan Model <i>Bagging-SVM</i>	49
Gambar 3.8 Skenario Keseluruhan Pengujian	51
Gambar 4.1 Karakteristik Data Hasil Pengukuran pada Salah Satu Subjek	55
Gambar 4.2 Grafik Persebaran Data pada Fitur-fitur Sebelum Normalisasi	58
Gambar 4.3 Grafik Persebaran Data pada Fitur-fitur Sesudah Normalisasi	59
Gambar 4.4 Representasi Grafis Difference Value of Mean (DV) per Fase	60
Gambar 4.5 Representasi Grafis Absolute Deviation of Mean (AD) dari Semua Subjek	62
Gambar 4.6 Grafik Persebaran Data pada Fitur-fitur Metode SMOTE	69

Halaman ini sengaja dikosongkan

DAFTAR TABEL

Tabel 2.1 Reaksi Psikologi dan Perilaku Ekspresi Emosi	11
Tabel 2.2 Tabel <i>Confusion Matrix</i>	34
Tabel 3.1 Struktur Data pada Fitur <i>Difference Value Of Mean</i> untuk Visualisasi	43
Tabel 3.2 Struktur Data pada Fitur <i>Absolute Deviation Of Mean</i> untuk Visualisasi	44
Tabel 3.3 Struktur Data pada Ekstraksi Fitur untuk Klasifikasi	44
Tabel 3.4 Tabel Persentase <i>Class Attribute</i>	45
Tabel 4.1 Hasil Survey <i>Video Game</i>	53
Tabel 4.2 Daftar Subjek Penelitian	54
Tabel 4.3 Hasil Ekstraksi Fitur <i>Difference Value Of Mean</i> untuk Visualisasi.....	56
Tabel 4.4 Hasil Ekstraksi Fitur <i>Absolute Deviation Of Mean</i> untuk Visualisasi..	57
Tabel 4.5 Hasil Ekstraksi Fitur untuk Klasifikasi SVM	57
Tabel 4.6 Hasil Pengambilan Data <i>Pulse Rate</i> Per Subjek	61
Tabel 4.7 <i>Confusion Matrix</i> dari SVM Menggunakan Metode <i>Bagging</i>	66
Tabel 4.8 Hasil Klasifikasi Algoritma SVM Menggunakan Metode <i>Bagging</i>	67
Tabel 4.9 <i>Confusion Matrix</i> dari SVM Menggunakan Metode <i>SMOTE</i>	70
Tabel 4.10 Hasil Klasifikasi Algoritma SVM Menggunakan Metode <i>SMOTE</i> ...	71
Tabel 4.11 <i>Confusion Matrix</i> dari SVM dengan <i>Class Attribute</i> per Fase	73
Tabel 4.12 Hasil Klasifikasi Algoritma SVM dengan <i>Class Attribute</i> per Fase...	75
Tabel 4.13 Hasil <i>Weighted Average</i> pada Parameter Metode <i>Bagging</i> dan <i>SMOTE</i>	76

Halaman ini sengaja dikosongkan

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Generasi Y atau disebut juga generasi milenium memiliki ciri menggunakan teknologi komunikasi instan seperti *email*, SMS, *instant messaging* dan media sosial seperti *facebook* dan *twitter*. Mereka juga tumbuh dalam era *game online* [1]. Bermain *game* komputer lebih bersifat emosional dari pada rasional. *Game* komputer juga disesuaikan dengan mengukur perilaku pemain dan mengenali keadaan emosi pemain untuk menyesuaikan fitur permainan tertentu secara dinamis agar dapat meningkatkan tantangan dan keterlibatan yang lebih [2].

Setiap *video game* memiliki emosi yang dominan saat memainkannya. Doumbouya *et al* [3] telah melakukan studi tentang emosi dominan di setiap *video game*. Penelitian mereka menemukan bahwa emosi kemarahan dan surprise adalah emosi dominan dari beberapa permainan yang mereka selidiki. Penelitian tentang pengenalan emosi menarik untuk dilakukan, karena emosi terbagi kedalam emosi positif dan negatif yang akan berpengaruh pada sistem kekebalan tubuh dan ada dampak pada kesehatan manusia [4], [5]. Oleh karena itu, penelitian tentang pengenalan emosi harus dilakukan secara berkelanjutan sehingga pengetahuan tentang dampaknya dapat menjadi informasi yang berguna.

Beberapa penelitian sebelumnya yang menggunakan *electroencephalogram*, *electrocardiogram*, dan lainnya untuk pengenalan emosi pada manusia. Pengenalan pola fisiologis berdasarkan *pulse rate* menggunakan video sebagai stimulus telah dilakukan oleh Wibawa *et al* [6] untuk mengetahui emosi dasar manusia. Analisis perubahan *pulse rate* adalah cara yang handal untuk menilai perubahan emosional, karena dalam beberapa penelitian menyatakan keandalannya sebagai indeks aktivitas jantung noninvasif [7], [8], [9]. Sementara penelitian yang menggunakan *video game* sebagai stimulus telah dilakukan oleh Chanel *et al* [10] dan Ambinder [11] untuk mengenali beberapa emosi yang terjadi saat bermain *video game*. Tujuan pengenalan emosi saat bermain *video game*

elektronik menurut Mandryk *et al* [12] adalah untuk memberikan evaluasi dari perspektif *gamer* dan memberikan penilaian tentang kesulitan *video game* sehingga penyesuaian pengguna dapat dinilai oleh pengembang *video game*.

Selain menimbulkan emosi negatif pada pengguna *video game* (*gamer*), efek lain yang disebabkan oleh bermain *video game* adalah menjadi kecanduan [13]. Menurut Organisasi Kesehatan Dunia (WHO), dalam edisi ke-11 dari Klasifikasi Penyakit Internasional, kecanduan *video game* diklasifikasikan sebagai salah satu gangguan kesehatan mental, kelainan dalam permainan *video game* ditandai dengan pola perilaku permainan yang terus-menerus atau berulang, yang terjadi secara *online* (yaitu, melalui *internet*) atau *offline* [14]. Hal ini juga disampaikan oleh praktisi kesehatan jiwa, dr. Kristiana Siste, SpKJ (K) dari Departemen Psikiatri Fakultas Kedokteran Universitas Indonesia, bahwa bermain *video game* apabila sampai menyebabkan distress dan disfungsi, maka termasuk kategori gangguan mental [15]. Saat ini Rumah Sakit Umum di Indonesia telah membuka klinik adiksi untuk menangani pasien-pasien yang mengalami ketergantungan (adiksi) [16]. Pasien yang bermasalah dengan kecanduan *gadget* dan *game online* terus mengalami peningkatan [17]. Tidak hanya Rumah Sakit Umum yang mendapat pasien tersebut, namun beberapa Rumah Sakit Jiwa pun mendapat kunjungan pasien yang mengalami kecanduan *game online* [18]. Kontribusi penelitian ilmiah telah dilakukan studi kasus di beberapa daerah di Indonesia yang menerangkan bahwa terdapat dampak negatif *game online* terhadap prokrastinasi akademik [19], perilaku negatif dan *impulsive* [20], [21] [22], kecerdasan sosial [23], dan prestasi belajar [24]. Namun pada penelitian lain, terdapat hubungan yang tidak begitu signifikan (kategori sedang), pada pengukuran tingkat kontrol diri dari pemain *game* DotA 2 di Kota Malang, dengan persentase sebesar 69% dan untuk tingkat kecanduan pada kategori sedang dengan persentase 70%. Stimulan yang sama yaitu *game online* DotA 2, juga telah dilakukan oleh Yohanes dkk [25] menghasilkan hubungan negatif signifikan antara kecanduan *game online* DotA 2 terhadap penyesuaian sosial pada remaja di Kota Salatiga. Penelitian pengaruh *game online* didominasi pada subjek remaja, hal ini didukung dengan penelitian Mihara *et al* [26] yang mengatakan bahwa pengaruh *Internet Gaming Disorder* (IGD) lebih tinggi di antara orang yang lebih

muda daripada orang tua. Kemudian, hubungan IGD terhadap kontrol kemarahan pada remaja dan dewasa mempunyai hubungan yang signifikan secara statistik [27], [28] dan hasil penyelidikan hubungan antara IGD dan *alexithymia* (sulit menggambarkan emosi dalam diri) dari celine *et al* [56] adalah: *gamer* pria menjadi *alexithymia* memiliki skor kecemasan dan depresi yang tinggi jika dikaitkan dengan IGD. Bahkan terdapat studi kohor [29], [30], [31] dalam menyelidiki pengaruh emosi marah pada kesehatan, yaitu tentang hubungan penyakit kardiovaskular dan karakteristik psikologis, seperti kecemasan dan kemarahan.

Pengenalan pola fisiologi berdasarkan denyut nadi (*pulse rate*) telah dilakukan oleh Adhi *et al* [32] untuk mengetahui emosi dasar manusia. Analisis perubahan denyut nadi adalah cara yang ampuh untuk menilai perubahan emosi, karena dalam beberapa penelitian menyatakan keandalannya sebagai indeks noninvasif dari aktivitas jantung [33], [34], [35].

Oleh karena itu, pada tulisan ini turut mengisi kontribusi pengetahuan untuk mengetahui pola emosi marah pada remaja ketika bermain *video game* dengan mengusulkan ukuran secara objektif menggunakan grafik visualisasi dan klasifikasi. Jenis *video game* yang di pilih telah tervalidasi oleh survey sebagai *video game* yang mempunyai emosi marah sebagai emosi yang dominan. Sinyal *pulse rate* telah diinvestigasi untuk mengenali berbagai emosi dasar, seperti senang, sedih dan marah [32], [36]. Hasilnya menunjukkan bahwa terdapat *physiological pattern* dengan karakteristik khusus pada sinyal *pulse rate* tersebut. Sedangkan pada penelitian ini bertujuan untuk memberikan informasi pola dan prediksi emosi marah pada remaja saat bermain *video game*. Selain itu, masih langkanya penelitian yang memanfaatkan data *pulse rate* untuk pengenalan pola emosi marah pada remaja dengan menggunakan stimulan *video game* yang tervalidasi. Hal tersebut merupakan perihal yang melatarbelakangi tujuan penelitian ini.

1.2 Rumusan Masalah

Tren perilaku remaja yang memiliki kebiasaan bermain *video game* pada saat sekarang, hingga terdapat klinik khusus pada Rumah Sakit di Indonesia yang

menangani pasien yang kecanduan *video game*, karena akibat yang ditimbulkan seperti ketidakmampuan untuk mengontrol diri, menjalankan kebutuhan diri, dan terdapat sikap agresif hingga kematian yang berarti kecanduan *video game* mempunyai dampak yang buruk. Bahkan lembaga kesehatan dunia *World Health Organization* (WHO) telah menetapkan perilaku kecanduan terhadap *video game* sebagai kelainan mental. Oleh karena itu, dirumuskan permasalahan sebagai berikut:

1. Bagaimana pola *pulse rate* yang terjadi pada pemain *video game* (*gamer*) ketika sebelum dan sedang bermain *video game*?
2. Bagaimana cara melakukan prediksi emosi marah yang terjadi saat bermain *video game* berbasis sinyal *pulse rate*?
3. Dari proses ekstraksi fitur sinyal *pulse rate* yang dipergunakan untuk klasifikasi dan pengelompokan *class attribute*, maka diperoleh dataset yang *imbalanced class*. Hal ini merupakan permasalahan yang kerap terjadi dalam proses klasifikasi karena algoritma *classifier* akan menghasilkan akurasi prediksi yang lebih tinggi terhadap kelas mayoritas dari pada kelas minoritas. Dari permasalahan tersebut, bagaimana usulan metode yang akan dilakukan agar algoritma klasifikasi dapat memberikan hasil prediksi dengan maksimal pada kondisi *imbalanced class* ?
4. Bagaimana cara untuk mengetahui bahwa model klasifikasi yang diusulkan telah memberikan hasil yang maksimal?

1.3 Tujuan

Bedasarkan rumusan permasalahan, maka tujuan pada penelitian ini yaitu:

1. Untuk memperoleh informasi dan wawasan baru mengenai kondisi *pulse rate* pada *gamer* ketika sebelum dan sedang bermain *video game*. Sehingga diharapkan tulisan ini turut mengisi kontribusi pengetahuan tentang pengaruh permainan *video game* dengan mengusulkan ukuran secara objektif.

2. Mengusulkan algoritma klasifikasi *Support Vector Machine* (SVM) untuk mengetahui prediksi dari sampel data *pulse rate* yang dikategorikan sebagai data emosi marah.
3. Mengusulkan metode Bagging (*Bootstrap Aggregating*), yaitu metode dengan pendekatan algoritma untuk menyelesaikan permasalahan *imbalanced class* dengan cara membangun beberapa model pada algoritma *classifier* yang sama (*multiple classifier*) dari sub-sampel yang berbeda.
4. Membandingkan metode Bagging dengan metode yang mempunyai pendekatan data, yaitu SMOTE (*Synthetic Minority Oversampling Technique*) yang memiliki cara kerja menambah kelas minoritas agar jumlahnya sebanding dengan kelas mayoritas.

1.4 Batasan Masalah

Batasan masalah pada penelitian ini adalah:

1. Subjek penelitian adalah remaja dengan kategori tidak memiliki riwayat penyakit kronis.
2. Stimulan yang dipilih adalah *video game* online DotA 2 (*Defense of the Ancients*) yang telah tervalidasi oleh survey di lingkungan tempat mengakusisi data.
3. Perangkat keras yang dipergunakan adalah sensor *pulse rate* dari *Cooking Hacks*.
4. Menggunakan Python dengan matplotlib sebagai *plotting library* untuk menggambarkan grafik visualiasi dengan *input* dari beberapa parameter yang statistik.

1.5 Kontribusi Penelitian

Kontribusi yang diharapkan dari hasil penelitian tesis ini adalah

1. Selanjutnya, penulis berharap penelitian ini dapat memberikan informasi bahwa sinyal *pulse rate* dapat menjadi tolok ukur untuk melihat bagaimana perubahan dalam pola emosi marah yang terjadi saat bermain *video game*. Selain itu informasi ini juga dapat digunakan sebagai

informasi lebih lanjut di bidang kesehatan dan pengembang teknologi *video game*.

2. Menambah informasi tentang metode pengolahan data yang efisien pada klasifikasi untuk permasalahan *imbalance class*.

BAB 2

KAJIAN PUSTAKA

Bab ini membahas mengenai kajian pustaka yang menjadi dasar dari penelitian ini, diantaranya kajian penelitian yang terkait dan teori dasar dari penelitian ini. Berikut uraian lengkap dari bab kajian pustaka ini.

2.1 Kajian Penelitian Terkait Prosedur Pengambilan Data

Penelitian tentang pengaruh *game online* terhadap perilaku remaja hingga saat ini masih menggunakan kuesioner sebagai metode observasi pengumpulan data dan menggunakan analisa statistik inferensial sederhana. Pengukuran sinyal *pulse rate* pada penelitian Lutfi *et al* [36] untuk mengenali pola fisiologis sebagai 3 jenis emosi dasar (senang, sedih dan marah) pada manula yang berusia lebih dari 55 tahun menggunakan algoritma klasifikasi SVM.

Pada penelitian yang dilakukan oleh Marzuki *et al* [37], menggunakan metode *Learning Vector Quantization* (LVQ) untuk mengklasifikasi sinyal *pulse rate* dan SpO₂ untuk mengenali pola fisiologi sebanyak 6 emosi dasar yaitu senang, marah, sedih, kaget, takut dan jijik. Selanjutnya, pada penelitian yang dilakukan oleh Khoirunnisaa dkk [38] menggunakan metode *Correlation Feature Selection* (CFS). Untuk mengidentifikasi *Cybersickness* pada *Electroencephalograph* (EEG) dimana identifikasi *cybersickness* dilakukan menggunakan dua metode pengukuran yaitu berdasarkan *Simulator Sickness Questionnaire* (SSQ) dan EEG 14 kanal.

2.2 Teori Dasar

2.2.1 Pengertian Video Game

Video game adalah *game* elektronik yang melibatkan interaksi dengan antarmuka pengguna atau perangkat *input*, seperti *joystick*, *controller*, *keyboard*, atau perangkat pengindra gerak, untuk menghasilkan umpan balik visual pada perangkat tampilan video dua atau tiga dimensi seperti perangkat televisi, monitor, layar sentuh, atau *headset* realitas virtual. Permainan video ditambah dengan

umpan balik audio dari *speaker* atau *headphone*, dan secara opsional dengan jenis sistem umpan balik lainnya termasuk teknologi *haptic*.

Video game dikelompokkan berdasarkan *platform*, yang meliputi *game arcade*, *video game* berjenis *consoles*, konsol *game* genggam, dan komputer, termasuk *mainframe* dan komputer pribadi. Baru-baru ini, industri *game* telah berkembang ke *game mobile* melalui *smartphone* dan komputer tablet, dan membuat *remoting game cloud*.

Video game pertama merupakan pengembangan dari *game* elektronik menggunakan *output* seperti *video* dari komputer dengan ukuran yang besar pada 1950 dan 1960, sedangkan *video game* pertama yang tersedia bagi konsumen muncul pada 1972 melalui *Magnavox Odyssey*, yaitu *video game consoles* berbasis rumahan komersial pertama yang dikembangkan oleh Ralph H. Selanjutnya pada tahun 1971 rilis *game arcade*, yaitu *game* dalam bentuk mesin yang dioperasikan dengan koin yang biasanya dipasang di bisnis publik seperti restoran, bar, dan arena hiburan. Sebagian besar *game arcade* adalah mesin *pinball*, *game* elektro-mekanis, dan *game* penebusan atau *merchandiser*.

Saat ini, pengembangan *video game* membutuhkan banyak keterampilan untuk membawa *game* ke pasaran, termasuk pengembang, penerbit, distributor, pengecer, dan produsen pihak ketiga lainnya, dan peran lainnya. Sejak 2010-an, kepentingan komersial dari industri *video game* telah meningkat. Pemasaran yang dapat mendorong pertumbuhan industri di negara-negara benua Asia yang sedang berkembang adalah *game mobile* pada *smartphone*. Pada 2020, *video game* menghasilkan penjualan sebanyak 2.7 miliar dollar Amerika setiap tahun di seluruh dunia, dan merupakan segmen terbesar ketiga di pasar hiburan di Amerika Serikat, setelah segmen dari siaran televisi berbasis internet [39].

2.2.1.1 Jenis Aliran (Genre) dari Video Game

Genre *video game* adalah pengelompokan *video game* berdasarkan beberapa kriteria, diantaranya adalah [40].:

- a. Berdasarkan pada *gameplay* (jenis interaksi) terhadap fitur visual atau naratif. Jenis *game* berdasarkan kategori ini diantaranya: *Action game*,

Adventure game, Massively Multiplayer Online Game (MMOG), Fighting game, Platform game, Puzzle game, Racing game, Role-playing game, Shooter game, Simulation game, Sports game, Strategy game dan lainnya.

- b. Berdasarkan *platform*, yaitu umumnya berdasarkan perangkat keras berbasis komputer tempat *game* dimainkan. Hal ini akan mencakup jenis *game* seperti *arcade game, console game, electronic game, handheld game (handheld console, game mobile* untuk *smartphone*, komputer tablet), dan *online game*.
- c. Berdasarkan *mode*, yaitu merujuk pada sebuah *game*, apakah memiliki pemain tunggal, multipemain, atau variasi. Hal ini termasuk *game massively multiplayer online (MMO)*, *game kooperatif, Player versus environment (PvE)* atau *Player versus player (PvP)*, dan sebagainya.
- d. Berdasarkan narasi, yaitu kategori *video game* dengan berdasarkan narasi yang dimilikinya, seperti fiksi ilmiah atau fantasi, dan fiksi horor.
- e. Berdasarkan tujuan, yaitu *game* yang dirancang dengan tujuan tertentu, dalam hal ini terbagi kedalam tiga kategori, yaitu: *Entertaining games, Serious games, dan Educational games*. *Entertaining games* adalah *game* yang memiliki tujuan sebagai hiburan. *Serious games* adalah *game* yang dirancang untuk menyampaikan informasi atau pengalaman belajar kepada pemain. *Game* jenis ini dirancang untuk dimainkan oleh para profesional sebagai bagian dari pekerjaan tertentu atau untuk peningkatan keterampilan dalam berbagai aspek seperti simulasi penerbangan, simulasi medis, eksplorasi ilmiah, manajemen darurat, perencanaan kota, teknik, dan politik. Selanjutnya adalah *Educational games*, yaitu *video game* yang menggambarkan penggabungan dari *video game* dan perangkat lunak pendidikan.

2.2.1.2 Aspek Sosial dari *Video Game*

Berdasarkan dari aspek demografi, menurut ESA (*Entertainment Software Association*) yaitu asosiasi perdagangan industri *video game* di Amerika Serikat menyediakan ringkasan pada bulan juli tahun 2020

berdasarkan survey ke 4000 responden di Amerika Serikat [41] sebagai berikut:

- a. *Gamer* rata-rata berusia 35 sampai 44 tahun. Dimana 21% berusia dibawah 18 tahun, 38% berusia 18 sampai 34 tahun, 26% berusia 35 sampai 54 tahun, 9% berusia 55 sampai 64 tahun dan 6% bersuaia lebih dari 65 tahun.
- b. Pemain *game* dewasa menghabiskan waktu 6.6 jam perminggu bermain secara *online*, dan 4.3 jam perminggu bermain langsung bersama pemain lain.
- c. Sebanyak 61% *game* dimainkan menggunakan *smartphone*, 52% menggunakan *console*, dan 49% menggunakan komputer.
- d. Sebanyak 71% janis *game* yang dimainkan adalah jenis *casual games*, 53% berjenis *action games*, dan 48% berjenis *shooter games*

2.2.2 Pengaruh Emosi Marah Terhadap Kesehatan

Sundari [42] memaparkan bahwa perasaan manusia menurut modalitasnya terbagi menjadi tiga, yaitu suasana hati, perasaan, dan emosi. Emosi merupakan bagian dari perasaan dalam arti luas. Emosi dapat terbentuk karena rasa yang bergejolak sehingga yang bersangkutan mengalami perubahan dalam situasi tertentu mengenai perasaan, namun seluruh pribadi menanggapi situasi tersebut. Pada akhirnya, individu dapat menyesuaikan diri dengan lingkungannya dan menentukan respons yang sesuai terkait situasi yang dihadapi.

Scherer dan Wallbott [43] melakukan riset mengenai emosi yang paling sering dipakai di 37 negara berbeda dengan pendekatan universal dan faktor budaya. Dari penelitian tersebut ditemukan bahwa terdapat 7 emosi manusia dan dapat dijelaskan melalui reaksi psikologi dan perilaku ekspresi [44].

Tabel 2.1 Reaksi Psikologi dan Perilaku Ekspresi Emosi

Emosi	Reaksi Psikologis	Perilaku Ekspresi
Gembira (<i>Joy</i>)	<ul style="list-style-type: none"> • Suhu tubuh menjadi hangat • Detak jantung meningkat 	<ul style="list-style-type: none"> • Tertawa/tersenyum • Pendekatan • Tutur kata menjadi panjang • Perubahan nada suara
Takut (<i>Fear</i>)	<ul style="list-style-type: none"> • Denyut nadi meningkat • Otot menegang • Perubahan nafas • Suhu tubuh menjadi dingin • Kalap/bingung 	<ul style="list-style-type: none"> • Diam • Ekspresi wajah
Marah (<i>Anger</i>)	<ul style="list-style-type: none"> • Denyut nadi meningkat • Otot menegang • Perubahan nafas • Suhu tubuh menjadi panas 	<ul style="list-style-type: none"> • Ekspresi wajah • Perubahan suara • Tutur kata menjadi panjang • Tindakan menyerang • Perubahan tempo
Sedih (<i>Sadness</i>)	<ul style="list-style-type: none"> • Kalap/bingung • Otot menegang 	<ul style="list-style-type: none"> • Diam • Menangis/tersedu-sedu • Perilaku menghindar • Ekspresi wajah
Muak (<i>Disgust</i>)	<ul style="list-style-type: none"> • Tidak ada gejala tertentu 	<ul style="list-style-type: none"> • Diam • Ekspresi wajah
Malu (<i>Shame</i>)	<ul style="list-style-type: none"> • Suhu tubuh menjadi panas • Denyut nadi meningkat 	<ul style="list-style-type: none"> • Diam • Ekspresi wajah
Bersalah (<i>Guilt</i>)	<ul style="list-style-type: none"> • Kalap/bingung 	<ul style="list-style-type: none"> • Diam

Dalam tulisannya, Duffy [45] mengungkapkan bahwa marah adalah sesuatu yang sangat normal dan merupakan perasaan yang sehat. Namun sangatlah penting untuk membedakan antara marah, agresi dan kekerasan yang sering kali disamakan. Marah merupakan potensi perilaku, yakni emosi yang dirasakan dalam diri seseorang. Sedangkan agresi atau kekerasan merupakan perilaku yang muncul akibat emosi tertentu, khususnya marah. Emosi marah tidak harus berujung pada perilaku agresi, marah yang dikelola dengan baik akan

memunculkan perilaku yang dapat diterima norma sosial seperti perilaku asertif, namun jika marah tidak mampu dikelola dengan baik, maka marah dapat berdampak pada munculnya perilaku agresi atau kekerasan yang tidak diterima norma sosial.

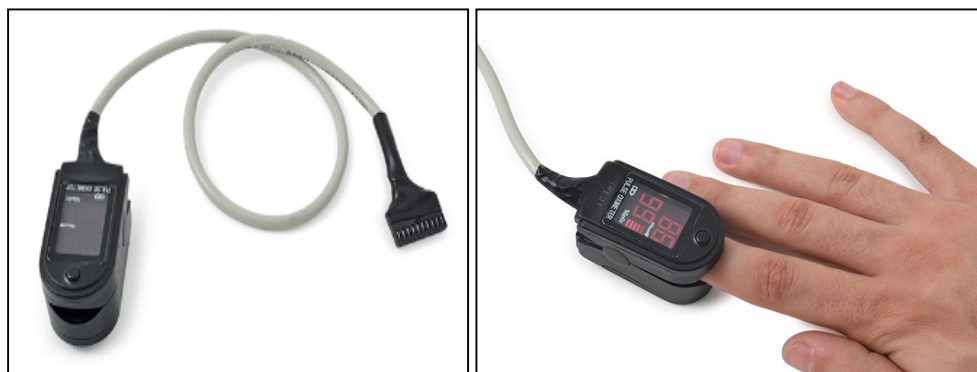
Penelitian Cautin *et al* [46] terhadap 92 remaja menunjukkan bahwa marah mempunyai peran yang sangat penting bagi timbulnya depresi dan menjadi salah satu faktor yang menyumbangkan risiko bunuh diri bagi remaja. Mereka membagi ekspresi marah menjadi dua golongan yaitu diinternalisasi atau dipendam sendiri dan dieksternalisasi atau diekspresikan pada lingkungannya. Hasilnya menunjukkan bahwa remaja yang menginternalisasi marahnya mempunyai kecenderungan terhadap depresi, dan terlebih lagi mengarah pada kemungkinan bunuh diri.

Lebih lanjut jika dikaitkan dengan kesehatan tubuh manusia, Sylvia D. Kreibig *et al* [47] menyatakan bahwa aktivitas sistem syaraf otonom atau *Autonomic Nervous System* (ANS) ditunjukkan sebagai komponen utama yang berhubungan dengan aktivitas emosi manusia. Aktivitas ANS yang ditimbulkan dari respon emosi ini dapat mempengaruhi perubahan fisiologis tubuh. ANS berperan untuk terus menyempurnakan fungsi dan sistem organ sesuai dengan rangsangan baik internal maupun eksternal, membantu mempertahankan stabilitas internal dan keseimbangan melalui koordinasi berbagai aktivitas seperti aktivitas pada sistem kardiovaskular, proses respirasi, sekresi hormon, sirkulasi, pencernaan dan ekskresi [47], [48]. Selain itu, ANS juga berfungsi menghubungkan sistem saraf pusat dengan kelenjar endokrin, yaitu kelenjar yang mengontrol detak jantung dan organ-organ internal lainnya, seperti perut, intestin, hati, dan paru-paru secara psikologis, serta sangat mempengaruhi dinamika emosi, perasaan dan suasana hati (*mood*) [47], [51]. Pada saat emosi, sistem syaraf simpatik juga membuat perubahan pada aktivitas sistem kardiovaskular, aktivitas jantung, proses pernafasan dan kandungan oksigen dalam darah, konduktansi kulit [47], [48], [49]. Kinerja syaraf simpatik yang meningkat secara terus menerus akan mengakibatkan terjadinya serangan penyakit, terutama penyakit kronis.

Pada penelitian ini mengidentifikasi emosi marah yang timbul pada remaja yang mempunyai kebiasaan bermain *game online* dengan mengusulkan pengukuran yang dilakukan secara objektif untuk mengetahui variabilitas kondisi *pulse rate*.

2.2.3 Sinyal *Pulse Rate*

Pulse rate (PR) atau denyut nadi merupakan salah satu parameter yang mengekspresikan denyut jantung (*heart rate*), yang menggambarkan jumlah siklus jantung yang terjadi setiap menitnya atau *beats per minute* (bpm) [51]. Meskipun diekspresikan dalam bpm, *pulse rate* tidak selalu dibutuhkan untuk menilai frekuensi denyut selama full satu menit. Sebagai contoh, satu metode untuk menilai *pulse rate* yang melibatkan perhitungan jumlah beat selama 15 detik kemudian mengalikan nilai ini dengan 4, atau menghitung jumlah beat selama 30 detik, kemudian mengalikan nilai tersebut dengan 2, untuk mendapatkan pengukuran *pulse rate* dalam *beats per minute* (bpm). Level normal *pulse rate* pada rata-rata orang dewasa diantara 60-100 bpm. Namun jika diklasifikasikan lagi, level normal wanita dewasa diantara 72-80 bpm dan laki-laki diantara 64-72 bpm [51]. Jika level *pulse rate* seseorang lebih dari 100 bpm, maka dapat didiagnosa orang tersebut mengalami gejala *tachycardia* [55]. Kondisi lain juga dikatakan seseorang mengalami gejala *bachycardia* jika level *pulse rate* kurang dari 60 bpm [49].



Gambar 2.1 Sensor *Pulse oximeter* Cooking Hacks dan cara pemasangan pada jari tangan [50]

Selanjutnya, Sinyal SpO_2 dan *Pulse rate* direpresentasikan sebagai kumpulan nilai SpO_2 dan *Pulse rate* yang direkam dalam rentang waktu dan berisi informasi dari sifat fisiologis seseorang berdasarkan metode *photoplethysmogram* (PPG). Untuk mengukur sinyal SpO_2 dan *Pulse rate* dalam rentang waktu, penelitian ini menggunakan perangkat Sensor Pulse Oximeter. Pulse Oximeter adalah perangkat yang bekerja berdasarkan metode PPG seperti yang dijelaskan di atas, dimana perangkat ini digunakan secara universal untuk monitoring pasien dalam pengaturan perawatan kritis [52]. Penelitian ini menggunakan sensor *Pulse Oximeter* yang dikembangkan oleh Cooking Hacks. Sensor ini dapat mengukur level SpO_2 dan *Pulse rate* seseorang dengan memasang perangkat pada jari tangan (Gambar 2.1). Untuk dapat memonitor dan merekam SpO_2 dan *Pulse rate* pada perangkat komputer, diperlukan perangkat pendukung seperti E-health Platform V2.0, Arduino Uno Rev.3, Kabel USB dan Aplikasi Coolterm yang dijelaskan secara detail pada Bab 3 tentang eksplorasi *hardware*.

Pulse oximeter yang merupakan pengembangan dari Cooking Hacks berprinsip pada sifat hemoglobin yang mampu menyerap cahaya dan denyut alami aliran darah di dalam arteri. Alat ini dilengkapi dengan sumber cahaya, pendeteksi cahaya dan mikroprosesor. Terdapat komponen dioda yang memancarkan dua jenis sinar yakni merah dan inframerah dengan panjang gelombang tertentu yang akan diserap oleh darah. Kedua jenis sinar tersebut disebarkan melalui jaringan tubuh menuju pendeteksi cahaya yang terdapat pada sisi lain dari alat ini. Hemoglobin yang mengandung oksigen lebih banyak akan menyerap lebih banyak cahaya infra merah, sedangkan yang tidak memiliki oksigen akan menyerap cahaya merah. Lebih lanjut, gelombang yang diserap oleh darah, selanjutnya dipantulkan kembali dalam panjang gelombang tertentu yang kemudian akan dideteksi oleh pulse oximeter dan menginterpretasikan nilai SpO_2 dan *Pulse rate*. Selain itu, Pancaran cahaya melalui bagian tubuh yang relatif tembus dan memiliki aliran darah arteri yang berdenyut bagus, seperti jari, kaki dan daun telinga. Rasio cahaya merah terhadap cahaya inframerah yang melewati situs pengukuran dan diterima oleh detektor oximeter berdasarkan pada persentase dari oksigen terhadap hemoglobin teroksigenasi melalui cahaya yang lewat. Sementara itu, mikroprosesor pada alat ini berguna untuk menghitung

perbedaan kadar oksigen dan mengubah informasi tersebut ke dalam nilai digital. Nilai tersebut selanjutnya ditaksir untuk menentukan jumlah oksigen yang dibawa oleh darah.

Meskipun sensor *pulse oximeter* memiliki banyak kelebihan, termasuk presisi yang tinggi dan kemudahan penggunaan untuk semua orang, ada beberapa faktor yang menyebabkan terjadinya kesalahan pengukuran. Berdasarkan literatur yang ditulis oleh kontributor Konika Minolta [53] memberikan poin-poin yang perlu diperhatikan oleh pengguna tentang faktor-faktor penyebab kesalahan penggunaan dan langkah-langkah pencegahannya, diantaranya:

- 1) Hemoglobin abnormal. Darah mungkin mengandung hemoglobin abnormal seperti *karbonsihemoglobin* dan *methemoglobin* yang tidak berkontribusi terhadap pengiriman oksigen. *Pulse oximeter* dengan panjang gelombang ganda dapat dipengaruhi oleh hemoglobin abnormal ini.
- 2) Pewarna medis. Jika pewarna seperti kardio hijau, intravaskular dan hijau indosianin telah disuntikkan ke dalam darah, dapat mempengaruhi tingkat transmisi cahaya merah dan inframerah.
- 3) *Manicure* dan *pedicure*. Jika pengguna memakai cat kuku, mungkin akan menyerap cahaya yang dipancarkan oleh LED, mengubah cahaya yang ditransmisikan melalui tubuh, dan mempengaruhi nilai yang dihitung.
- 4) Gerakan tubuh utama. Gerakan tubuh dapat menyebabkan noise yang mempengaruhi nilai yang dihitung. Saat *noise*, termasuk yang disebabkan oleh gerakan tubuh, mengurangi reliabilitas nilai yang dihitung, maka *pulse oximeter* akan menampilkan peringatan.
- 5) Aliran darah terblokir karena tekanan pada lengan atau jari. *Pulse oximeter* mengukur saturasi atau kadar oksigen berdasarkan perubahan aliran darah. Oleh karena itu, jika aliran darah terhambat, pengukuran yang benar mungkin tidak terjadi. Selain itu, jika jari-jari dilipat dengan kecepatan yang seragam, *pulse oximeter* dapat menafsirkan tekanan sebagai perubahan *pulse rate* yang menyebabkan kesalahan.
- 6) Kegagalan peredaran darah *peripheral*. *Pulse oximeter* menggunakan aliran darah untuk memantau perubahan jumlah cahaya yang ditransmisikan untuk menghitung nilai. Jika aliran darah perifer berkurang, data yang memadai

mungkin tidak didapatkan dan hasilnya adalah pengukuran yang tidak akurat. Dalam hal ini, perlu untuk mendorong aliran darah dengan memijat atau menghangatkan jari atau mengukur jari lainnya dengan aliran darah yang lebih teratur.

- 7) Lampu *ambient* yang berlebihan. *Pulse oximeter* biasanya bisa menghilangkan efek cahaya sekitar. Namun, jika lampu sekitar terlalu kuat, perangkat tidak akan bisa menghilangkan efeknya dan hal ini bisa menyebabkan kesalahan pengukuran.
- 8) Gelombang elektromagnetik *ambient*. Jika peralatan listrik seperti televisi, telepon genggam atau perangkat medis yang menghasilkan gelombang elektromagnetik tingkat tinggi digunakan di dekat *pulse oximeter*, gelombang elektromagnetik dari perangkat tersebut dapat mengganggu pengukuran yang akurat.
- 9) *Probe* tidak terpasang dengan benar. Jika *probe* tidak terpasang dengan benar, mungkin mendeteksi berbagai *noise*, sehingga pengukurannya tidak akurat.

2.2.4 Ekstraksi Fitur *Pulse Rate*

Ekstraksi fitur merupakan bagian yang paling penting dalam penelitian pengenalan pola. Kinerja dari algoritma klasifikasi dipengaruhi dari pemilihan metode ekstraksi fitur yang digunakan. Fitur didefinisikan sebagai sebuah variabel *input* atau atribut. Sehingga proses ekstraksi fitur merupakan proses pencarian ciri-ciri unik dari sebuah variabel *input* atau atribut yang berupa sinyal fisiologis, dalam hal ini sinyal SpO₂ dan *Pulse rate*, agar masing-masing kelas dapat dikenali oleh algoritma klasifikasi. Pada penelitian ini metode ekstraksi fitur yang digunakan adalah metode ekstraksi fitur statistik pada domain waktu. Penelitian ini memanfaatkan parameter statistik yang dideskripsikan pada bagian dibawah ini:

- 1) Rata-rata (*mean*) merupakan ukuran tendensi sentral yang didefinisikan sebagai nilai rata-rata dari suatu distribusi. Nilai rata-rata dinyatakan dengan persamaan berikut :

$$X_{mean} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.1)$$

- 2) Standar deviasi merupakan ukuran sebaran (*dispersi*) suatu distribusi. Jika

suatu distribusi flat atau rata, maka nilai standar deviasinya kecil sedangkan jika suatu distribusi berfluktuasi (tidak rata) maka nilai standar deviasinya besar. Standar deviasi dirumuskan oleh persamaan berikut:

$$X_{std} = \sqrt{\frac{\sum_{i=1}^n (x_i - MS)^2}{n-1}} \quad (2.2)$$

- 3) Nilai maksimum merupakan nilai maksimal dari level SpO2 dan *Pulse rate* dalam rentang waktu. Persamaan berikut digunakan untuk mendapatkan nilai maksimum dari sinyal SpO2 dan *pulse rate*.

$$X_{max} = Max [x_n] \quad (2.3)$$

- 4) Nilai minimum dari level SpO2 dan *pulse rate* dalam rentang waktu. Berikut persamaan untuk mendapatkan nilai minimum.

$$X_{min} = Min [x_n] \quad (2.4)$$

- 5) *Varians* merupakan salah satu ukuran dispersi atau ukuran variasi yang menggambarkan bagaimana berpencarnya suatu data kuantitatif

$$X_{var} = \sum_{i=1}^n (x_i - MS) \frac{2}{n-1} \quad (2.5)$$

- 6) *Skewness* merupakan pengukuran tingkat ketidaksimetrisan (kecondongan) sebaran data di sekitar rata-ratanya. Distribusi normal merupakan distribusi yang simetris dan nilai *skewness* adalah 0. *Skewness* yang bernilai positif menunjukkan ujung dari kecondongan menjulur ke arah nilai positif (ekor kurva sebelah kanan lebih panjang). *Skewness* yang bernilai negatif menunjukkan ujung dari kecondongan menjulur ke arah nilai negatif (ekor kurva sebelah kiri lebih panjang).

$$X_{skew} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - MS)^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - MS)^2\right)^{\frac{3}{2}}} \quad (2.6)$$

- 7) *Kurtosis* menunjukkan keruncingan (*peakedness*) atau kerataan (*flatness*) suatu distribusi data dibandingkan dengan distribusi normal. Pada distribusi normal, nilai kurtosis sama dengan 0. Nilai *kurtosis* yang positif menunjukkan distribusi yang relatif runcing, sedangkan nilai *kurtosis* yang negatif menunjukkan distribusi yang relatif rata.

$$X_{kurt} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i + x^{\wedge})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i + x^{\wedge})^2\right)^2} \quad (2.7)$$

2.2.5 Visualisasi Data

Visualisasi adalah suatu proses mengubah data kedalam format visual atau tabel sehingga karakteristik dari data dan relasi di antara item data atau atribut dapat dengan mudah dianalisis dan dipahami oleh pengguna. Visualisasi biasanya bersifat interaktif dan seringkali ditampilkan dalam bentuk animasi. Teknik visualisasi dapat dikelompokkan menjadi beberapa cara seperti geometrik, simbolis, stimulus 2D, 3D, n-dimensi, atau tampilan statis atau dinamis.

Visualisasi data merupakan salah satu dari teknik yang paling baik dan menarik untuk menampilkan sekumpulan data. Visualisasi data dirancang berdasarkan proses visual dan analitik yang dikembangkan dalam berbagai disiplin ilmu termasuk visualisasi ilmiah, grafik komputer, penggalian data, statistik, dan pembelajaran mesin dengan ekstensi khusus yang menjadikan kumpulan data multidimensi sangat besar secara interaktif. Metodologi didasarkan pada kedua fungsi yang mencirikan struktur dan menampilkan data untuk mengetahui pola, pengecualian, tren, dan hubungan. Tujuan utama dari visualisasi data adalah untuk memberikan pemahaman dan mengkomunikasikan informasi secara jelas dan lebih efisien kepada pengguna lewat grafik informasi yang digunakan, seperti tabel dan grafik.

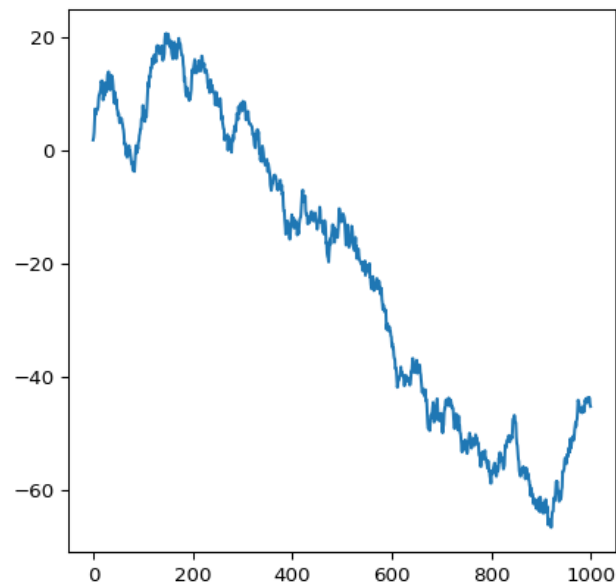
Menurut [57], terdapat dua macam teknik visualisasi data, yaitu visualisasi ilmiah (*Scientific Visualization*) dan visualisasi informasi (*Information Visualization*). Visualisasi ilmiah berfokus terutama pada data fisik seperti tubuh manusia, bumi, dan molekul. Visualisasi ilmiah juga dapat berhubungan dengan data multidimensi, tetapi sebagian besar data yang digunakan dalam bidang ini menggunakan atribut spasial untuk tujuan visualisasi, misalnya, Sistem Informasi Geografis (SIG). Sedangkan visualisasi informasi lebih dikhususkan pada data abstrak dan nonfisik seperti teks, hierarki, dan data statistik. Teknik visualisasi data yang digunakan pada penelitian ini sebagai berikut:

2.2.5.1 Informasi Grafik

Grafik adalah suatu rangka untuk membentuk objek visualisasi tabel, dimana tabel terdiri atas angka-angka yang bisa ditampilkan ke dalam bentuk gambar, bisa berbentuk diagram garis (*line chart*), diagram batang (*bar chart*), dan sebaran plot (*scatter plots*) dan lain-lainnya. Berikut ini adalah teknik yang digunakan dalam visualisasi informasi grafik antara lain [57]:

a. Diagram Garis (*Line Charts*)

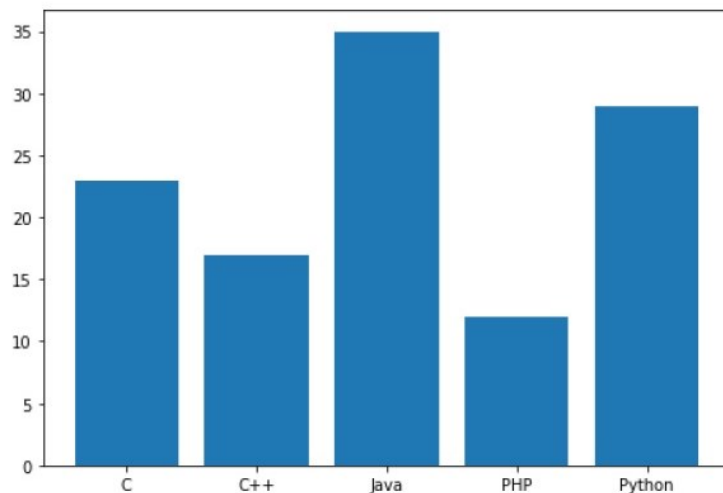
Diagram garis (*line charts*) merupakan teknik visualisasi yang digunakan untuk menunjukkan suatu hubungan satu variabel dengan variabel yang lain. Diagram ini paling sering digunakan untuk menunjukkan perubahan atau tren dari waktu ke waktu. Diagram garis juga berguna untuk membandingkan beberapa item selama periode waktu yang sama. Garis susun digunakan untuk membandingkan tren atau nilai individual untuk beberapa variabel. Contoh diagram garis dapat dilihat pada Gambar 2.2.



Gambar 2.2 Contoh Diagram Garis (*Line Charts*).

b. *Diagram Batang (Bar Charts)*

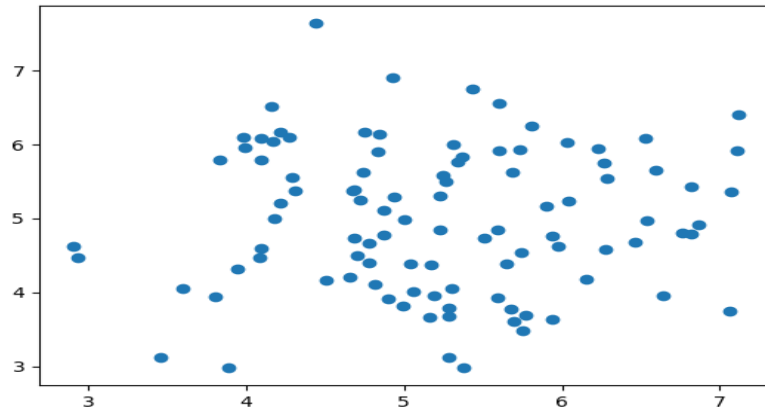
Diagram Batang (*Bar Charts*) merupakan teknik visualisasi yang paling sering digunakan untuk membandingkan jumlah kategori atau kelompok yang berbeda. Nilai kategori diwakili menggunakan bar, dan dapat dikonfigurasi dengan bar vertikal atau horizontal, dengan panjang atau tinggi setiap batang yang menunjukkan suatu nilai. Nilai yang diberikan harus jelas, sehingga perbedaan dalam bar dapat dideteksi oleh mata manusia. Namun, ketika nilai-nilai (bar) sangat berdekatan atau ada sejumlah besar nilai (bar) yang perlu ditampilkan, itu akan menjadi lebih sulit untuk membandingkan bar satu dengan yang lain. Untuk membandingkan biasanya memberikan variasi visual dengan warna yang berbeda untuk menunjukkan hal-hal seperti status atau rentang tertentu. Contoh grafik batang dapat dilihat pada Gambar 2.3.



Gambar 2.3 Contoh Diagram Batang (*Bar Charts*).

c. *Sebaran Plot (Scatter Plots)*

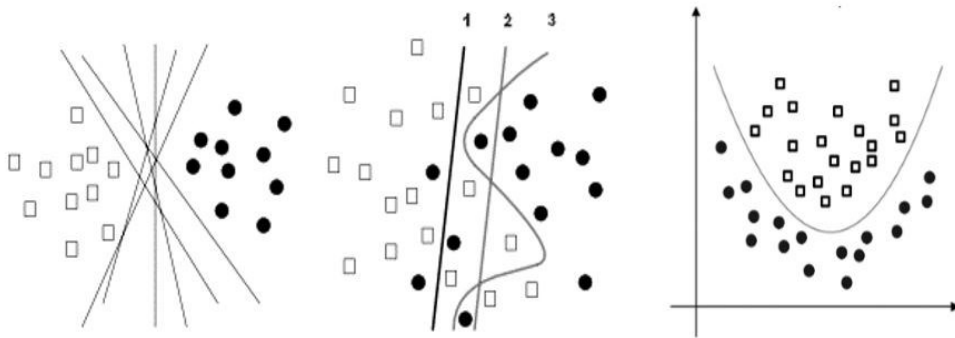
Sebaran plot merupakan teknik yang digunakan untuk analisis bivariat. Sebaran plot (atau plot X-Y) adalah plot dua dimensi yang menunjukkan variasi gabungan dari dua item data. Dalam plot pencar, setiap penanda (simbol seperti titik, kotak dan tanda tambah) mewakili pengamatan. Posisi penanda menunjukkan nilai untuk setiap observasi. Contoh sebaran plot dapat dilihat pada Gambar 2.4.



Gambar 2.4 Contoh Sebaran Plot (*Scatter Plots*).

2.2.6 *Support Vector Machine (SVM)*

Support vector machine (SVM) adalah metode pembelajaran *supervised* yang diperkenalkan pertama kali oleh Vapnik pada tahun 1995 dan sangat berhasil dalam melakukan prediksi, baik dalam kasus regresi maupun klasifikasi. SVM didasarkan pada prinsip minimalisasi resiko struktural / *structural risk minimization (SRM)*. Prinsip induksi ini berbeda dari prinsip minimalisasi resiko empirik (ERM) yang hanya meminimalkan kesalahan pada proses pelatihan. Pada SVM, fungsi tujuan dirumuskan sebagai masalah optimisasi konveks berbasis *quadratic programming*, untuk menyelesaikan *dual problem*. Menurut Tan, Steinbach dan Kumar [58], *Support Vector Machine (SVM)* adalah metode klasifikasi yang bekerja dengan cara mencari *hyperplane* dengan margin optimum. *Hyperplane* adalah garis batas pemisah data antar kelas. *Margin (d)* adalah jarak antara *hyperplane* dengan data terdekat pada masing-masing kelas. Bidang pembatas pertama membatasi kelas pertama dan bidang pembatas kedua membatasi kelas kedua sedangkan data yang berada pada bidang pembatas merupakan vektor-vektor yang terdekat dengan *hyperplane* terbaik disebut dengan *Support Vector*. SVM untuk klasifikasi dapat bekerja pada kasus klasifikasi linier maupun *nonlinier*. Pada klasifikasi linier, SVM dapat dibedakan menjadi dua yaitu *linierly separable* dan *linierly nonseparable* [59]. Gambar 2.5 merupakan gambar ilustrasi dari klasifikasi linier.



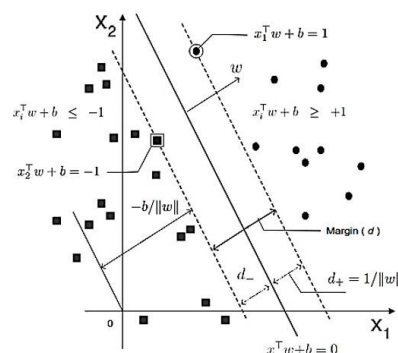
Gambar 2.5 Klasifikasi SVM: (kiri) Klasifikasi Linear Separable; (tengah) Linear Nonseparable; (kanan) Nonlinear [59]

2.2.6.1 SVM Linier Separable

Menurut Chen [59], setiap observasi terdiri dari sepasang p prediktor $\mathbf{X}_i = (x_{i1}, x_{i2}, \dots, x_{ip}) \in \mathbf{R}^p$, $i = 1, 2, \dots, n$ dan dihubungkan dengan $y_i \in \mathbf{y} = \{-1, 1\}$ maka dapat dinyatakan dalam himpunan berikut:

$$D_n = \{(x_1, y_1), \dots, (x_n, y_n)\} \in X \times \{-1, 1\} \quad (2.8)$$

Jika \mathbf{X}_i adalah anggota kelas (+1) maka \mathbf{X}_i diberi label (target) $y_i = +1$ dan jika tidak maka diberi label (target) $y_i = -1$ sehingga data yang diberikan berupa pasangan $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$ merupakan himpunan data *training* dari dua kelas yang akan diklasifikasi dengan SVM [59]. Pada Gambar 2.6, dapat dilihat bahwa berbagai alternatif bidang pemisah yang dapat memisahkan semua dataset sesuai dengan kelasnya namun bidang pemisah terbaik tidak hanya dapat memisahkan data tetapi juga memiliki margin paling besar [59].



Gambar 2.6 Bidang Pemisah Terbaik dengan Margin (d) Terbesar Linier Separable [59]

Pertama akan dijelaskan SVM pada kasus linier *separable*. Sebuah konsep utama untuk menetapkan pemisah yang bersifat linier adalah *dot product*, juga disebut sebagai *inner product* atau *scalar product*, antara dua vektor yang ditetapkan sebagai $\mathbf{x}^T \mathbf{w} = \sum_i x_i w_i$. Keluarga \mathcal{F} dari fungsi klasifikasi yang terdapat pada ruang data diberikan sebagai:

$$\mathcal{F} = \mathbf{x}^T \mathbf{w} + b, \mathbf{w} \in \mathbf{R}^p, b \in \mathbf{R}, \quad (2.9)$$

dimana \mathbf{w} diketahui sebagai vektor pembobot dan b disebut dengan bias. Bidang pemisah (*separating hyperplane*):

$$f(x) = \mathbf{x}^T \mathbf{w} + b = 0 \quad (2.10)$$

yang membagi ruang (*space*) menjadi dua daerah seperti yang terdapat pada Gambar 2.6. Bentuk pada $f(x)$ adalah sebuah garis dalam dua dimensi, sebuah bidang pada tiga dimensi, dan secara umum berupa *hyperplane* pada dimensi yang lebih tinggi. *Hyperplane* dikatakan linier jika merupakan fungsi linier dalam *input* x_i . Data yang berada pada *margin* (d) disebut dengan *support vector*. Fungsi pemisah untuk kedua kelas adalah sebagai berikut:

$$\begin{aligned} \mathbf{x}_i^T \mathbf{w} + b &\geq 1 \text{ untuk } y_i = +1 \\ \mathbf{x}_i^T \mathbf{w} + b &\leq -1 \text{ untuk } y_i = -1 \end{aligned} \quad (2.11)$$

dimana \mathbf{w} adalah vektor bobot (*weight vector*) yang berukuran ($p \times 1$), b adalah posisi bidang relatif terhadap pusat koordinat atau lebih dikenal dengan bias yang bernilai skalar.

$$\mathbf{X} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \mathbf{x}_i^T = [x_{i1} \quad x_{i2} \quad \cdots \quad x_{ip}] \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \end{bmatrix} \mathbf{y}_i = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad (2.12)$$

Pada Gambar 2.6 menunjukkan $\frac{|b|}{\|\mathbf{w}\|}$ adalah jarak bidang pemisah yang tegak lurus dari titik pusat koordinat dan $\|\mathbf{w}\|$ adalah jarak Euclidean (*norm Euclidean*) dari \mathbf{w} . Panjang *vector* \mathbf{w} adalah *norm* $\|\mathbf{w}\| = \sqrt{\mathbf{w}^T \mathbf{w}} = \sqrt{w_1^2 + w_2^2 + \cdots + w_p^2}$. Bidang batas pertama membatasi kelas (+1) sedangkan bidang pembatas kedua membatasi kelas (-1). Bidang pembatas pertama $\mathbf{x}_i^T \mathbf{w} +$

$b=1$ mempunyai bobot \mathbf{w} dan jarak tegak lurus dari titik asal sebesar $\frac{|1-b|}{\|\mathbf{w}\|}$, sedangkan bidang pembatas kedua $\mathbf{x}_i^T \mathbf{w} + b = -1$ mempunyai bobot \mathbf{w} dan jarak tegak lurus dari titik asal sebesar $\frac{|-1-b|}{\|\mathbf{w}\|}$. Jarak antara margin dan bidang pemisah (*separating hyperplane*) adalah $d_+ = d_- = \frac{1}{\|\mathbf{w}\|}$. Nilai maksimum margin atau nilai margin (jarak) antara bidang pembatas (berdasarkan rumus jarak garis ke titik pusat) adalah

$$\frac{1-b-(-1-b)}{\|\mathbf{w}\|} = \frac{2}{\|\mathbf{w}\|} \quad (2.13)$$

Hyperplane yang optimal adalah $\max \frac{|2|}{\|\mathbf{w}\|}$ atau *equivalent* dengan $\min \frac{1}{2} \|\mathbf{w}\|^2$.

Dengan menggabungkan kedua kendala pada persamaan (2.11) maka dapat dipresentasikan dalam pertidaksamaan sebagai berikut:

$$y_i(\mathbf{X}_i^T \mathbf{w} + b) - 1 \geq 0 \quad , i = 1, 2, \dots, n \quad (2.14)$$

Secara matematis, formulasi permasalahan optimasi SVM untuk klasifikasi linier dalam *primal space* adalah:

$$\min \frac{1}{2} \|\mathbf{w}\|^2 \quad (2.15)$$

Dengan fungsi kendala $y_i(\mathbf{X}_i^T \mathbf{w} + b) \geq 1 \quad , i = 1, 2, \dots, n$.

Pada formulasi di atas, ingin meminimalkan fungsi tujuan $\frac{1}{2} \|\mathbf{w}\|^2$ atau sama saja dengan memaksimalkan $\|\mathbf{w}\|^2$ atau $\|\mathbf{w}\|$. Maksimal margin $\frac{2}{\|\mathbf{w}\|}$ dapat diperoleh dari meminimalkan $2 \|\mathbf{w}\|^2$ atau $\|\mathbf{w}\|$.

Secara umum, persoalan optimasi (2.15) ini akan lebih mudah diselesaikan jika diubah ke dalam formula *lagrange*. Dengan demikian permasalahan optimasi dengan kendala dapat dirumuskan menjadi:

$$L_{pri}(\mathbf{w}, b, \alpha) = \frac{1}{2} \|\mathbf{w}\|^2 - \sum_{i=1}^n \alpha_i [y_i(\mathbf{X}_i^T \mathbf{w} + b) - 1] \quad (2.16)$$

dengan kendala $\alpha_i \geq 0$ (nilai dari koefisien *lagrange*). Penaksir \mathbf{w} dan b dengan meminimumkan L_{pri} terhadap \mathbf{w} dan b kemudian disamadengankan $\frac{\partial L_{pri}(\mathbf{w}, b, \alpha)}{\partial \mathbf{w}} = 0$ dan $\frac{\partial L_{pri}(\mathbf{w}, b, \alpha)}{\partial b} = 0$, sehingga diperoleh persamaan (2.17)

$$\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i \text{ dan } \sum_{i=1}^n \alpha_i y_i \quad (2.17)$$

Vektor \mathbf{w} seringkali bernilai besar (tak terhingga), tetapi nilai α_i terhingga. Untuk itu, formula *lagrange* L_{pri} (*primal problem*) diubah ke dalam L_D (*dual problem*). Dengan mensubstitusikan persamaan (2.17) ke persamaan (2.16) diperoleh L_D yang ditunjukkan pada persamaan (2.18):

$$L_D = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \quad (2.18)$$

Jadi, persoalan pencarian bidang pemisah terbaik dapat dirumuskan pada persamaan (2.19).

$$\max_{\alpha} L_D = \max \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \quad (2.19)$$

nilai α_i dapat diperoleh, yang nantinya akan digunakan untuk mencari nilai \mathbf{w} .

Jika nilai $\alpha_i > 0$ atau sebuah titik data ke- i untuk setiap $y_i(\mathbf{X}_i^T \mathbf{w} + b) = 1$. Penyelesaian masalah *primal* dan *dual* pada persamaan (2.16) dan (2.18) memberikan solusi yang sama ketika masalah optimasi adalah *convex*. Setelah menyelesaikan *dual problem*, maka suatu pengamatan baru $\mathbf{x}_{(new)}$ dapat diklasifikasikan menggunakan ukuran klasifikasi sebagai berikut:

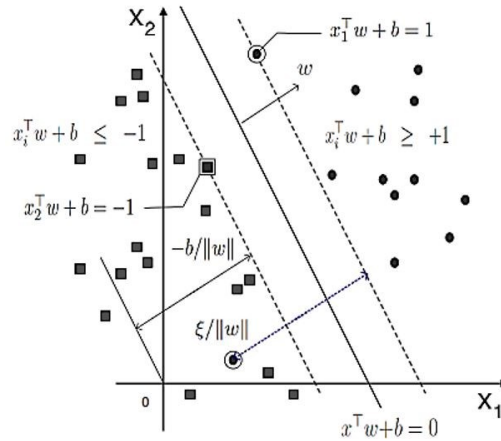
$$f(\mathbf{x}_{new}) = \text{sign}(\mathbf{X}_{new}^T \mathbf{w} + b) \quad (2.20)$$

Dimana $\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i$ dan $b = \frac{1}{n_{sv}} (\sum_{i=1}^{n_{sv}} - (\mathbf{X}_{new}^T \mathbf{w}))$ dengan \mathbf{x}_i adalah *support vector*, $\mathbf{x}_{(new)}$ adalah data yang diklasifikasikan, α_i adalah *lagrange multiplier* dan b adalah bias dan n_{sv} adalah jumlah *support vector*.

2.2.6.2 SVM Linier Non-separable

Chen [59] menyatakan pada kasus linier *nonseparable* yaitu mengklasifikasikan data linier yang tidak dapat dipisahkan maka kendala pada persamaan (2.11) harus diubah secara linier dengan penambahan variabel *slack* ξ_i yang menunjukkan pinalti terhadap ketelitian pemisahan yang memungkinkan suatu titik berada di dalam *margin error* ($0 \leq \xi_i \leq 1, \forall_i$) atau dinamakan misklasifikasi ($\xi_i > 1$), sehingga \mathbf{x}_i diklasifikasikan menjadi:

$$\begin{aligned} \mathbf{X}_i^T \mathbf{w} + b &\geq 1 - \xi_i \text{ untuk } y_i = 1 \text{ (untuk kelas +1)} \\ \mathbf{X}_i^T \mathbf{w} + b &\geq -(1 - \xi_i) \text{ untuk } y_i = -1 \text{ (untuk kelas -1)} \end{aligned} \quad (2.21)$$



Gambar 2.7 Bidang Pemisah Terbaik dengan Margin (d) Terbesar Linier *Non-Separable* [59]

Bidang pemisah terbaik dengan margin (d) terbesar pada linier *non-separable*, dapat diilustrasikan pada Gambar 2.7. Pencarian bidang pemisah terbaik dengan penambahan variabel ξ_i sering juga disebut dengan *soft margin hyperplane*. Formula pencarian bidang pemisah terbaik atau fungsi tujuan berubah menjadi:

$$\min_{\mathbf{w}, \xi} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i \quad (2.22)$$

Persamaan (2.21) dapat digabungkan ke dalam dua *constraint* dalam bentuk persamaan (2.23):

$$y_i(\mathbf{X}_i^T \mathbf{w} + b) \geq 1 - \xi_i \quad (2.23)$$

dengan $\xi_i \geq 0, C > 0$, dimana C adalah parameter yang menentukan besar biaya akibat kesalahan klasifikasi (*misclassification*) dari data *training* selama proses pembelajaran dan nilainya ditentukan oleh pengguna. Ketika nilai C besar, maka *margin* akan menjadi lebih kecil, yang mengindikasikan bahwa tingkat toleransi kesalahan akan menjadi lebih kecil ketika suatu kesalahan terjadi. Sebaliknya, ketika nilai C kecil, tingkat toleransi kesalahan akan menjadi lebih besar. Bentuk persamaan (2.22) memenuhi prinsip *Structural Risk Minimization* (SRM) dimana meminimumkan $\frac{1}{2} \|\mathbf{w}\|^2$ ekuivalen dengan meminimumkan dimensi VC (*Vapnik-Chervonenkis*). Nilai dari dimensi VC ini akan menentukan besarnya nilai kesalahan hipotesis pada data *testing* sedangkan meminimumkan $C \sum_{i=1}^n \xi_i$ ekuivalen dengan meminimumkan *error* pada data *training*. Fungsi *lagrange* untuk *primal problem* adalah:

$$L_{pri}(\mathbf{w}, b, \alpha) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i [y_i (\mathbf{X}_i^T \mathbf{w} + b) - 1 + \xi_i] - \sum_{i=1}^n \mu_i \xi_i \quad (2.24)$$

dimana $\alpha_i > 0$ dan $\mu_i \geq 0$ adalah *Lagrange Multiplier*. Kondisi KKT (*KarushKhun-Tucker*) untuk *primal problem* adalah:

$$\frac{\partial L_{pri}(\mathbf{w}, b, \alpha)}{\partial \mathbf{w}} = 0 \rightarrow \mathbf{w} - \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i = 0 \Leftrightarrow \mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i$$

$$\frac{\partial L_{pri}(\mathbf{w}, b, \alpha)}{\partial b} = 0 \rightarrow - \sum_{i=1}^n \alpha_i y_i = 0 \Leftrightarrow \sum_{i=1}^n \alpha_i y_i = 0$$

$$\frac{\partial L_{pri}(\mathbf{w}, b, \alpha)}{\partial \xi_i} = 0 \rightarrow C - \alpha_i - \mu_i = 0 \Leftrightarrow \alpha_i = C - \mu_i$$

Dengan kondisi untuk *Lagrange multipliers*:

$$\alpha_i \geq 0,$$

$$\mu_i \geq 0,$$

$$\alpha_i [y_i (\mathbf{X}_i^T \mathbf{w} + b) - 1 + \xi_i] = 0,$$

$$\mu_i \xi_i = 0$$

Dengan mensubstitusikan nilai $\mathbf{w} = \sum_{i=1}^n \alpha_i y_i \mathbf{x}_i$ ke dalam *primal problem* menjadi persamaan *dual problem* sebagai berikut:

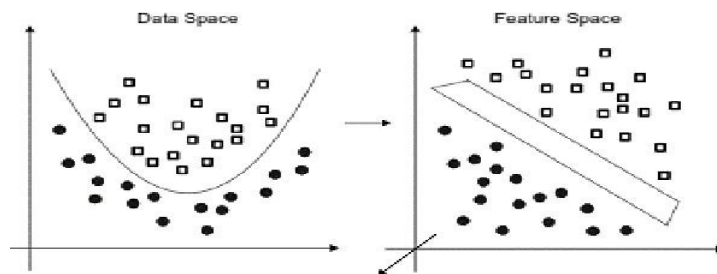
$$\max_{\alpha} L_D = \max \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \quad (2.25)$$

dengan $0 \leq \alpha_i \leq C$ dan $\sum_{i=1}^n \alpha_i y_i = 0$

Sampel \mathbf{x}_i untuk $\alpha_i > 0$ (*support vector*) yaitu titik yang berada di atas margin atau dalam margin ketika *soft margin* digunakan. *Support vector* sering menyebar dan level penyebarannya berada pada batas atas (*upper bound*) untuk *misclassification rate* [63].

2.2.6.3 SVM Non-linier Separable

Menurut Chen [59], pada kenyataan tidak semua data bersifat linier sehingga sulit untuk mencari bidang pemisah secara linier. Diberikan beberapa titik baru $x \in Y$ dan ingin memprediksi hubungan $y \in Y = \{-1,1\}$, maksudnya adalah memilih y dimana (x,y) hampir mirip ke *training* sampel. Akhirnya, memerlukan pengukuran kemiripan dalam X dan dalam $\{-1,1\}$. Permasalahan ini dapat diselesaikan dengan mentransformasikan data ke dalam dimensi ruang yang berdimensi lebih tinggi sehingga dapat dipisahkan secara linier pada *feature space* yang baru. SVM juga bekerja pada data nonlinier.



Gambar 2.8 Pemetaan Ruang Data Dua Dimensi (kiri) ke Dalam Ruang *Feature* Tiga Dimensi (kanan) [46]

$$\mathbf{R}^2 \leftrightarrow \mathbf{R}^3$$

Klasifikasi nonlinier yang ditunjukkan pada Gambar 2.8, suatu pemetaan data dengan struktur nonlinier melalui suatu fungsi : $\varphi: \mathbf{R}^p \rightarrow \mathbf{H}$ ke dalam ruang berdimensi tinggi \mathbf{H} dimana aturan klasifikasi bersifat linier. Perhatikan bahwa semua *vector training* \mathbf{x}_i terdapat dalam persamaan (2.25) sebagai *scalar product* dari bentuk $\mathbf{x}_i^T \mathbf{x}_j$. Pada SVM nonlinier, *scalar product* ditransformasikan ke $\varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}_j)$. Fungsi transformasi pada SVM adalah menggunakan “*Kernel*

Trick” [60]. Kegunaan *kernel trick* untuk menghitung *scalar product* melalui sebuah fungsi kernel. Proyeksi fungsi : $\varphi: \mathbf{R}^p \rightarrow \mathbf{H}$ memastikan bahwa *inner product* $\varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}_j)$ ditunjukkan oleh fungsi kernel

$$K(x_i, x_j) = \varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}_j) \quad (2.26)$$

Jika suatu fungsi kernel K pada persamaan (2.26), dapat digunakan tanpa perlu mengetahui fungsi transformasi φ secara eksplisit.

Diberikan sebuah kernel K dan data $x_1, x_2, \dots, x_n \in X$ maka matriks $K = (K(x_i, x_j))_{ij}$ berukuran $n \times n$ disebut *Gram matrix* untuk data x_1, x_2, \dots, x_n . Sebuah syarat cukup dan perlu untuk matriks simetri K , dengan $K_{ij} = K(x_i, x_j) = K(x_j, x_i) = K_{ji}$, untuk K definit positif disebut “*Mercer’s Theorem*” [61].

$$\sum_{i=1}^n \sum_{j=1}^n \varphi_i \varphi_j K(x_i, x_j) \geq 0 \quad (2.27)$$

Contoh sederhana pada sebuah *kernel trick* yang menunjukkan bahwa kernel dapat dihitung tanpa perhitungan fungsi *mapping* φ secara eksplisit adalah fungsi pemetaan:

$$\varphi(x_1, x_2) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)^T \quad (2.28)$$

Sehingga menjadi :

$$\mathbf{w}^T \varphi(x) = w_1 w_1^2 + \sqrt{2} w_2 x_1 x_2 + w_3 x_2^2$$

Dengan dimensi pada *feature space* adalah kuadrat, padahal dimensi asalnya adalah linier. Metode kernel menghindari pembelajaran secara eksplisit *mapping* data ke dalam *feature space* dimensi tinggi, seperti pada contoh berikut:

$$\begin{aligned} f(x) &= \mathbf{w}^T \mathbf{x} + b \\ &= \sum_{i=1}^n \alpha_i \mathbf{x}_i^T \mathbf{x} + b \\ &= \sum_{i=1}^n \alpha_i \varphi(\mathbf{x}_i)^T \varphi(\mathbf{x}) + b \text{ dalam } \textit{feature space} \\ &= \sum_{i=1}^n \alpha_i K(x_i, x) + b \end{aligned} \quad (2.29)$$

Hubungan kernel dengan fungsi *mapping* adalah:

$$\begin{aligned}
\varphi(\mathbf{x}_i)^T \varphi(x) &= (x_{i1}^2, \sqrt{2}x_{i1}x_{i2}, x_{i2}^2)(x_1^2, \sqrt{2}x_1x_2, x_2^2)^T \\
&= x_{i1}^2x_1^2 + 2x_{i1}x_{i2}x_1x_2 + x_{i2}^2x_2^2 \\
&= (\mathbf{x}_i^T \mathbf{x})^2 \\
&= K(\mathbf{x}_i, \mathbf{x})
\end{aligned} \tag{2.30}$$

Sedangkan, untuk memperoleh fungsi klasifikasi nonlinier dalam data *space*, bentuk secara umumnya diperoleh dari penerapan *kernel trick* ke persamaan (2.31):

$$L_D = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i \mathbf{x}_j) \tag{2.31}$$

yaitu memaksimumkan

$$L_D: \max_{\alpha} L_D = \max \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K(\mathbf{x}_i \mathbf{x}_j) \tag{2.32}$$

dengan, $\sum_{i=1}^n \alpha_i y_i = 0, 0 \leq \alpha_i \leq C; i = 1, 2, \dots, n$

Fungsi kernel yang umum digunakan pada metode SVM adalah:

1. Kernel Linier

$$K(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j \tag{2.33}$$

2. Kernel *Polynomial*

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma \mathbf{x}_i^T \mathbf{x}_j + r)^p, \gamma > 0 \tag{2.34}$$

3. Kernel *Radial Basis Function* (RBF)

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \tag{2.35}$$

4. Kernel *Sigmoid*

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\gamma \mathbf{x}_i^T \mathbf{x}_j + r) \tag{2.36}$$

Pemilihan fungsi kernel yang tepat merupakan hal yang sangat penting karena akan menentukan *feature space* dimana fungsi *classifier* akan dicari. Sepanjang fungsi kernelnya sesuai (cocok), SVM akan beroperasi secara benar meskipun tidak tahu pemetaan yang digunakan [62]. Menurut Scholkopf [59] [60],

fungsi kernel gaussian RBF memiliki kelebihan yaitu secara otomatis menentukan nilai, lokasi dari *center* serta nilai pembobot dan bisa mencakup nilai rentang tak terhingga. Gaussian RBF juga efektif menghindari *overfitting* dengan memilih nilai yang tepat untuk parameter C dan γ dan RBF baik digunakan ketika tidak ada pengetahuan terdahulu. Fungsi kernel yang direkomendasikan adalah fungsi kernel RBF karena dapat memetakan hubungan tidak linier, RBF lebih *robust* terhadap *outlier* karena fungsi kernel RBF berada antara selang $(-\infty, \infty)$ sedangkan fungsi kernel yang lain memiliki rentang antara (-1 sampai dengan 1) [66].

2.2.7 K-Fold Cross Validation

K-folds Cross Validation adalah salah satu teknik untuk validasi yang sangat populer digunakan. Metode validasi dengan *k-folds* sangat cocok digunakan untuk kasus data yang jumlah sampelnya terbatas. Untuk melakukan proses klasifikasi tentunya data dibagi ke dalam *training* dan *testing*, dan ketika data yang digunakan untuk *training* sangat sedikit kemungkinan adalah data yang digunakan kurang representatif. Dalam *k-folds cross validation*, data (D) dibagi ke dalam *k-subsets* $D_1, \dots, D_2, \dots, D_k$ dengan jumlah yang sama. Data yang digunakan untuk *training* adalah *subsets* data *k-1* yang dikombinasikan secara bersama-sama dan kemudian diaplikasikan untuk sisa satu *subsets* data sebagai hasil *testing*. Proses ini diulangi sebanyak *k subsets* dan hasil akurasi klasifikasi yaitu hasil rata-rata dari setiap data *training* dan *testing*. *k-folds* yang biasa digunakan adalah 3, 5, 10 dan 20 [66].

2.2.8 Bagging (Bootstrap Aggregating)

Bagging (*Bootstrap Aggregating*) merupakan teknik yang diusulkan oleh Breiman yang dapat digunakan untuk mereduksi variansi estimator pada metode klasifikasi dan regresi [64]. Penggunaannya tidak dibatasi hanya untuk memperbaiki estimator. Teknik ini juga dapat memperbaiki stabilitas, meningkatkan akurasi, dan kekuatan prediktif. Jumlah replikasi minimum untuk tujuan klasifikasi yakni sebanyak 50 kali, serta optimum saat nilai akurasi tertinggi telah didapatkan [64]. Bagging merupakan salah satu prosedur intensif

perhitungan untuk memperbaiki estimator atau pengklasifikasi yang tidak stabil, khususnya masalah dimensi tinggi. Bagging merupakan salah satu bagian dari *Bootstrap*. Perbedaan pada Bagging dilakukan dengan membandingkan sampel secara berpasangan. Ide dasar dari Bagging adalah menggunakan *bootstrap resampling* untuk membangkitkan prediktor dengan banyak versi. Bagging merupakan salah satu prosedur intensif perhitungan untuk memperbaiki estimator atau pengklasifikasi yang tidak stabil, khususnya masalah data dimensi tinggi.

Menurut Breiman [64], Bagging merupakan implementasi sederhana dari pembangkitan replikasi quasi *learning sample*. Didefinisikan peluang dari kasus ke- n dari suatu *learning sample* adalah $p(n) = 1/N$. Kemudian ambil sampel sebanyak N kali dari distribusi $\{p(n)\}$, secara ekuivalen merupakan sampel dari T dengan pengembalian. Himpunan sampel dari T disampel kembali menjadi himpunan *learning sample* T' . T' lebih dikenal dengan istilah sampel *bootstrap* dari T .

Ada beberapa peneliti yang memberikan rekomendasi tentang banyaknya replikasi *bootstrap* sampling yang harus dibuat, diantaranya Sutton merekomendasikan untuk melakukan replikasi sebanyak 25 atau 50 kali [73]. Sedangkan Hastie *et al* [71] menyatakan bahwa peningkatan akurasi akan terjadi jika banyaknya replikasi ditingkatkan dari 50 ke 100 kali dan jika replikasinya ditingkatkan menjadi yang lebih dari 100 kali akan menghasilkan akurasi yang tidak lebih besar dibandingkan replikasi 100 kali.

Buhlmann dan Yu [65], secara singkat menyatakan algoritma Bagging.

1. Menyusun sampel $\mathcal{L}_i^* = (Y_i^*, \mathbf{X}_i^*), i = 1, \dots, n$, menurut distribusi empiris pada pasangan

$$\mathcal{L}_i = (Y_i, \mathbf{X}_i), i = 1, \dots, n. \quad (2.37)$$

2. Hitung estimator bootstrap $\theta_n^*(x)$ dengan prinsip *plug-in* yaitu:

$$\theta_n^*(x) = h_n(\mathcal{L}_1^*, \dots, \mathcal{L}_n^*) \quad (2.38)$$

dengan

$$\theta_n(x) = h_n(\mathcal{L}_1, \dots, \mathcal{L}_n)$$

3. Estimator Bagging adalah sebagai berikut:

$$\theta_{nB}(x) = E^*[\theta_{nB}^*(x)] \approx B^{-1} \sum_{b=1}^B \theta_{nB}^*(x) \quad (2.39)$$

Secara heuristik kinerja variansi estimator Bagging $\theta_{nB}(x)$ adalah sama dengan atau lebih kecil dibandingkan estimator orisinal $\theta_{nB}^*(x)$.

2.2.9 Metode Synthetic Minority Oversampling Technique (SMOTE)

Metode *Synthetic Minority Oversampling Technique* (SMOTE) telah diusulkan oleh Chawla *et al* [70] sebagai solusi untuk mengatasi *imbalanced class* menggunakan prinsip yang berbeda dengan metode oversampling yang telah pernah diusulkan sebelumnya. Pada metode *oversampling* memiliki cara kerja memperbanyak pengamatan secara *random*, sedangkan pada metode SMOTE dengan cara membangkitkan data buatan sehingga jumlah data kelas minor setara dengan kelas mayor. Data buatan tersebut dibangkitkan atas dasar posisi k-tetangga terdekat atau disebut juga dengan metode *k-nearest neighbor*. Algoritma SMOTE terdiri dari dua bagian utama, bagian pertama berisi perulangan untuk mencari k-tetangga terdekat dan bagian kedua untuk membuat data buatan dari kelas minor.

2.2.10 Evaluasi Performansi Metode Klasifikasi

Evaluasi performansi yang dilakukan pada penelitian ini memiliki tujuan untuk mengukur kinerja model klasifikasi. *Confusion matrix* digunakan untuk mengevaluasi seberapa baik metode yang digunakan pada proses klasifikasi dengan cara mengenali kelas yang diprediksi dan kelas aktual klasifikasi. TP (*True Positive*) dan TN (*True Negative*) menunjukkan bahwa hasil klasifikasi (data yang dikenali/diprediksi) adalah benar, sedangkan FP (*False Positive*) dan FN (*False Negative*) menunjukkan hasil klasifikasi (data yang dikenali/diprediksi) yang tidak benar, dan N adalah jumlah total klasifikasi yang dilakukan. *Confusion matrix* adalah sebuah tabel dengan ukuran $m \times m$, dimana $m \geq 2$. Nilai *confusion matrix* pada baris menunjukkan kelas aktual klasifikasi saat ini, sedangkan nilai pada kolom menunjukkan kelas prediksi hasil klasifikasi. Seperti yang diperlihatkan pada Tabel 2.2. Kondisi idealnya, nilai FP dan FN seharusnya bernilai 0 atau mendekati 0. Berikut ini adalah Tabel *Confusion Matrix* pada kasus klasifikasi *binary*:

Tabel 2.2 Tabel *Confusion Matrix*

AKTUAL	PREDIKSI	
	Positif	Negatif
Positif	TP	FN
Negatif	FP	TN

Keterangan:

TP : *True Positive* (jumlah prediksi benar pada kelas positif)

FP : *False Positive* (jumlah prediksi salah pada kelas positif)

FN : *False Negative* (jumlah prediksi salah pada kelas negatif)

TN : *True Negative* (jumlah prediksi benar pada kelas negatif)

Dari tabel *confusion matrix* tersebut dapat dihitung menjadi beberapa parameter penilaian. Parameter penilaian yang digunakan dalam penelitian ini diantaranya adalah:

- 1) Akurasi adalah parameter untuk menilai keseluruhan efektivitas algoritma yang diusulkan terhadap nilai prediksi yang benar dari label kelas. Nilai dari akurasi dinyatakan dengan persamaan berikut:

$$Akurasi = \frac{TN+TP}{TN+TP+FN+FP} \quad (2.40)$$

- 2) Presisi atau *Positive Predictive Value* (PPV) adalah berapa persen kelas positif yang diprediksi dengan benar dari keseluruhan kelas yang diprediksi positif.

$$Presisi = \frac{TP}{TP+FP} \quad (2.41)$$

- 3) Sensitivitas atau *Recall* atau *True Positive Rate* (TPR) adalah berapa persen kelas positif yang diprediksi dengan benar dari keseluruhan kelas yang sebenarnya adalah positif atau proporsi kelas positif yang diidentifikasi atau diprediksi dengan benar.

$$Sensitivity = \frac{TP}{TP+FN} \quad (2.42)$$

- 4) *Specificity* atau *True Negative Rate* (TNR) adalah berapa persen kelas negatif yang diprediksi dengan benar dari keseluruhan kelas yang

sebenarnya adalah negatif atau proporsi kelas negatif yang diidentifikasi atau diprediksi dengan benar.

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (2.43)$$

5) *Negative predictive values* (NPV) adalah berapa persen kelas negatif yang diprediksi dengan benar dari keseluruhan kelas yang diprediksi negatif.

$$NPV = \frac{TN}{TN+FN} \quad (2.44)$$

6) *False positive rate* (FPR) adalah proporsi kelas positif yang diidentifikasi atau diprediksi dengan salah.

$$FPR = \frac{FP}{FP+TN} \quad (2.45)$$

7) *False negative rate* (FNR) adalah proporsi kelas negatif yang diidentifikasi atau diprediksi dengan salah.

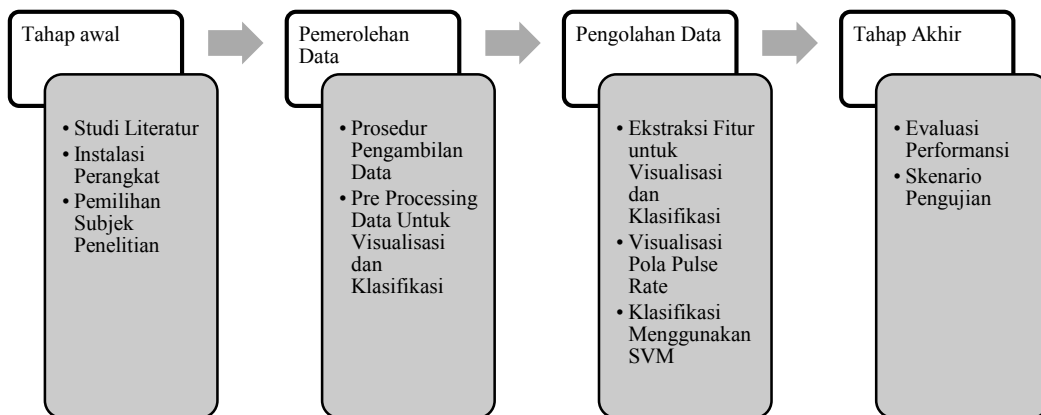
$$FNR = \frac{FN}{FN+TP} \quad (2.46)$$

Halaman ini sengaja dikosongkan

BAB 3

METODE PENELITIAN

Pada bab ini membahas prosedur pengambilan dan pengolahan data, serta bagaimana langkah-langkah dari metode yang dilakukan dari pengenalan pola sinyal *pulse rate* pada remaja menggunakan visualisasi dan prediksi emosi marah pada *gamer* menggunakan klasifikasi *Support Vektor Machine* (SVM). Tahapan penelitian dijelaskan pada Gambar 3.1 dibawah ini:



Gambar 3.1. Diagram alur Penelitian

3.1 Tahap Awal

Tahap awal adalah tahap untuk menentukan tema penelitian dari rumusan permasalahan, menentukan bagaimana penelitian akan dilakukan, menentukan alat yang diperlukan dan menentukan manfaat dari hasil penelitian. Hal ini bertujuan agar penelitian dapat dilakukan sesuai dengan yang diharapkan. Perihal yang dilakukan dalam tahap awal adalah sebagai berikut:

3.1.1 Studi Literatur

Studi literatur yang dilakukan dengan cara mencari sumber bacaan dari buku dan sumber daring yang berkaitan dengan tema penelitian, yaitu pengenalan

emosi menggunakan sensor fisiologi. Selanjutnya mencari tahu bagaimana pengolahan hasil sinyal tersebut agar dapat menjadi informasi baru.

3.1.2 Instalasi Perangkat

Perangkat keras dan lunak yang diperlukan untuk mendukung penelitian, sebelumnya dilakukan instalasi. Perangkat yang dipergunakan pada proses pengambilan dan pengolahan dataset sinyal *pulse rate* yaitu:

- 1) Sensor *pulse rate* dari Cooking Hack Libelium.
- 2) Modul *Printed Circuit Board E-Health sensor platform V2.0* dari Cooking Hack Libelium.
- 3) *Microcontroller* Arduino UNO Rev.3 berlisensi *open-source*.
- 4) Laptop.
- 5) Aplikasi *Coolterm* berlisensi *freeware* dari <http://freeware.the-meiers.org>.
- 6) Aplikasi Weka (*Waikato Environment for Knowledge Analysis*) versi 3.9.4 berlisensi *General Public License*.
- 7) Aplikasi Python versi 3.8 berlisensi *General Public License*.

Cooking Hacks adalah divisi perangkat keras berlisensi *open-source* dari Libelium (penyedia platform *Internet of Things*). Produk yang digunakan pada penelitian ini adalah adalah sensor *pulse rate* dan modul *Printed Circuit Board E-Health sensor platform V2.0*. Perangkat tersebut diinstallkan pada mikrokontroler *open-source* Arduino UNO Rev.3 yang akan di rekam oleh aplikasi *Coolterm*.

Hasil dari data *pulse rate*, masih berupa data mentah yang belum bisa memberikan informasi baru. Perangkat lunak untuk pengolahan data pada metode klasifikasi yang dilakukan menggunakan aplikasi Weka yang dikembangkan oleh *University of Waikato*. Selanjutnya pada tahap visualisasi pola sinyal *pulse rate*, dipergunakan bahasa pemrograman Python.

3.1.3 Pemilihan Subjek Penelitian

Pemilihan subjek penelitian yang diambil adalah remaja dengan rentang usia 20 hingga 25 tahun (10 laki-laki dan 2 perempuan) yang memiliki level permainan di atas pemula dan tidak memiliki riwayat penyakit kronis. Level

permainan yang dilakukan adalah level yang biasa subjek mainkan. Pada proses pengambilan data menggunakan sensor *pulse rate*. Ruang yang dipergunakan merupakan ruangan yang biasa subjek pergunakan untuk bermain *video game*. Gambar 3.2 merupakan salah satu contoh suasana pengambilan data bermain *video game*.



Gambar 3.2. Gambaran Suasana Subjek pada Saat Pengambilan Data

3.2 Pemerolehan Data

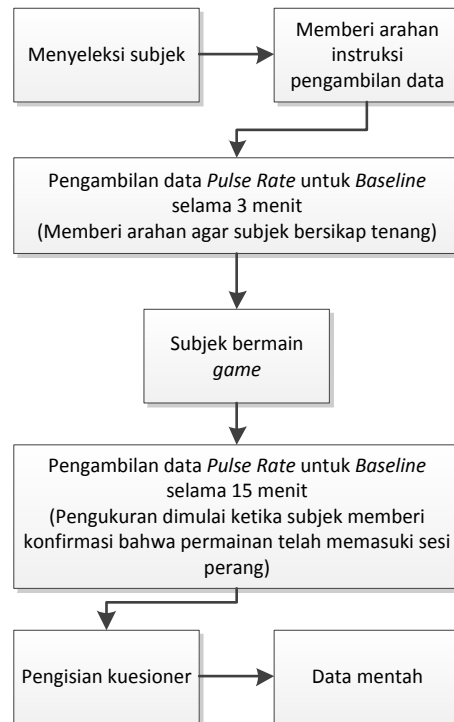
Hasil dari penentuan subjek penelitian, langkah selanjutnya yang dilakukan adalah melakukan prosedur pengambilan data. Pada tahap ini, terdapat penentuan stimulan yang akan dipergunakan hingga metode *pre processing* pada data *pulse rate* untuk dipergunakan pada visualisasi dan klasifikasi.

3.2.1 Prosedur Pengambilan Data

Pada prosedur pengambilan data, langkah pertama yang dilakukan adalah menentukan stimulan yang akan digunakan oleh subjek dalam pengambilan data *pulse rate*. Yaitu dengan cara memvalidasi *stimulation (video game)* menggunakan survey ke komunitas pengembang *video game* di kota Surabaya (GADaS-Game Developer arek Suroboyo).

Sebelum pengambilan data dilakukan penyampaian informasi perihal prosedur dan dilakukan persetujuan dengan memanfaatkan data untuk penelitian sebagaimana mestinya. Tahapan pengambilan data pada yang dilakukan yaitu: pemberian instruksi kepada subjek tentang bagaimana prosedur yang akan

dilakukan, kemudian pemasangan sensor *pulse rate* untuk pengambilan data pertama (*baseline*), pengambilan data selanjutnya ketika subjek bermain *video game*, dan terakhir pengisian kuesioner. Data *baseline* yang diambil untuk penelitian ini selama 3 menit, sedangkan data bermain *video game* yang diambil selama 15 menit. Berikut pada Gambar 3.3 diagram alir pada prosedur pengambilan data:

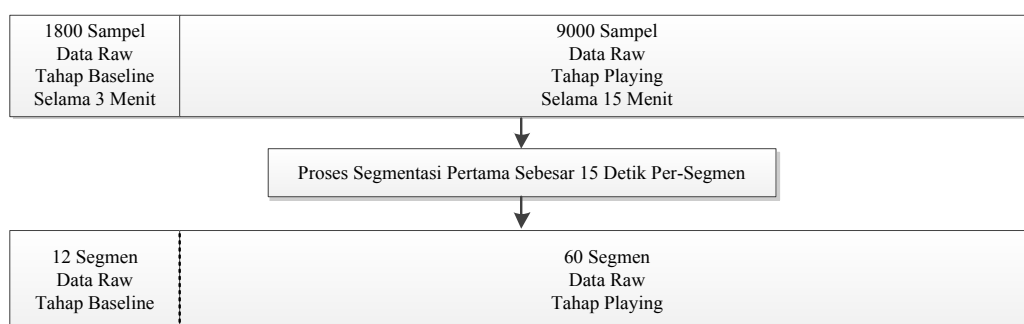


Gambar 3.3 Diagram alir pada Prosedur pengambilan data

3.2.2 Pre-Processing Data

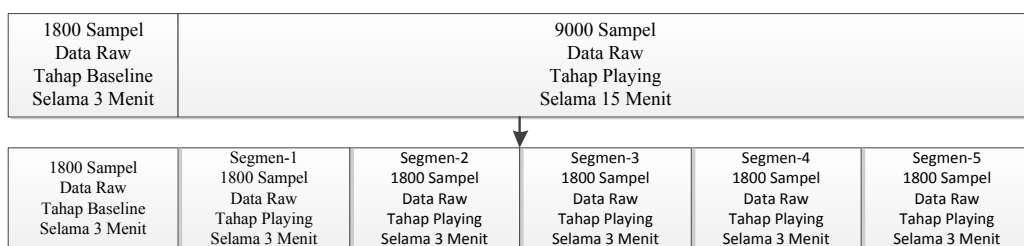
Pre-Processing adalah tahapan setelah mendapatkan data *raw* menjadi data yang bisa di komputasi. Pada tahapan ini terdapat proses *noise filtering*, yaitu menyaring data yang tidak diperlukan yang mengandung kesalahan atau terdapat data anomali. Proses segmentasi terdapat dua buah, pertama proses segmentasi untuk keperluan ekstraksi fitur pada proses klasifikasi dan kedua proses segmentasi untuk keperluan ekstraksi fitur pada proses visualisasi. Pada proses segmentasi untuk proses klasifikasi, *segmenting window size* dilakukan dengan cara membagi data berdasarkan *time domain* selama 15 detik untuk 1 ukuran segmen. Alasan proses ini dilakukan karena emosi adalah respons fisiologis

jangka pendek dan juga agar *window size* pada proses ekstraksi fitur berjumlah atau berukuran sama pada setiap segmen. Metode *segmenting window size* ini terinspirasi dari Wen *et al* [54] yang membagi sinyal fisiologi menjadi ukuran yang sama berdasarkan waktu (*window time-series*). Selanjutnya, dari proses segmentasi pertama ini akan didapatkan 12 segmen pada tahap *baseline* dan 60 segmen pada tahap *playing*. Segmen-segmen tersebut akan diperlakukan pada proses ekstraksi fitur terhadap 12 orang subjek, seperti yang diilustrasikan pada Gambar 3.4 dibawah:



Gambar 3.4 Blok Diagram Pada *Window Segmentation* untuk Proses Klasifikasi

Selanjutnya pada Pada proses segmentasi untuk proses visualisasi, *segmenting window size* dilakukan dengan cara membagi data berdasarkan *time domain* selama 3 menit untuk 1 ukuran segmen, seperti yang diilustrasikan pada Gambar 3.5 dibawah ini:



Gambar 3.5 Blok Diagram Pada *Window Segmentation* untuk Proses Visualisasi

3.3 Pengolahan Data

Meskipun terdapat dua buah metode yang diusulkan untuk pengenalan pola sinyal *pulse rate*, yaitu visualisasi dan klasifikasi, namun ekstraksi fitur untuk pengolahan data yang dilakukan terdapat kesamaan, yaitu sama-sama

mempergunakan fitur dari statistik sebagai *input* pengolahan data yang dilakukan pada penelitian ini.

3.3.1 Ekstraksi Fitur untuk Visualisasi dan Klasifikasi

Ekstraksi fitur merupakan ciri-ciri yang unik dari variabel atau atribut *input* yang bertujuan agar algoritma dapat melakukan pengenalan pola pada proses klasifikasi. Pada penelitian ini ekstraksi fitur yang digunakan adalah fitur berdasarkan domain waktu., yaitu pengolahan fitur yang dilakukan setelah data di bagi menjadi beberapa segmen berdasarkan domain waktu. Penelitian ini terinspirasi oleh penelitian sebelumnya yang menggunakan ekstraksi fitur dalam domain waktu [19]. Fitur yang diusulkan untuk visualisasi adalah *Difference Value Of Mean* (DV) and *Absolute Deviation Of Mean* (AD). Sedangkan fitur untuk klasifikasi adalah *Standard Deviation*, *Mean*, *Maximum*, *Minimum*, *Range*, *Variance*, *Skewnes* dan *Kurtosis*.

Pertama, sebelum melakukan perhitungan nilai *Difference Value Of Mean* (DV), dilakukan pengelompokan data didasarkan pada deret waktu sehingga setiap segmentasi yang diekstraksi menjadi rata-rata adalah seimbang. Persamaan *Difference Value Of Mean* yang digunakan sebagai berikut:

$$DV = \bar{x}_{PFI} - \bar{x}_B \quad (3.1)$$

Dimana DV = perbedaan nilai rata-rata, \bar{x}_{PFI} = nilai rata-rata pada tahap bermain *video game* per fase ke i , dan \bar{x}_B = nilai rata-rata pada tahap *baseline*. Seperti yang di ilustrasikan pada Gambar 3.5, dalam 15 menit bermain *video game*, dibagi kedalam 5 segmentasi, hal ini bertujuan agar pola yang terbentuk dapat dengan mudah dimengerti. Gambaran struktur data hasil ekstraksi fitur dari parameter *Difference Value Of Mean* yang akan digunakan sebagai *input* grafik visualisasi diilustrasikan pada Tabel 3.1 berikut:

Tabel 3.1 Struktur Data pada Fitur *Difference Value Of Mean* untuk Visualisasi

Subjek	Fase ke-1	Fase ke-2	...	Fase ke-5	Fitur
1	DV
2	DV
3	DV
...
10	DV
11	DV
12	DV

Proses ekstraksi fitur statistik selanjutnya adalah menghitung nilai *Absolute Deviation Of Mean* (AD) seperti dalam persamaan dibawah. Rangkaian waktu perubahan nilai AD digunakan untuk menentukan tren perubahan pola sinyal PR yang terkait dengan kemarahan. Nilai AD adalah nilai penyimpangan absolut dari nilai rata-rata dalam deret data tertentu [23]. AD lebih efisien dan lebih mudah dipahami daripada standar deviasi dalam situasi realistis dan untuk distribusi normal [24]. Nilai AD dihitung setiap 1 detik selama 3 menit untuk tahap *baseline* dan 15 menit untuk tahap bermain *video game*. Kemudian, terdapat 1.800 sampel sinyal *pulse rate* di tahap *baseline* dan 9.000 di tahap bermain *video game*. Selanjutnya, nilai AD yang diperoleh menjadi *input* data dari grafik visualisasi. Persamaan *Absolute Deviation Of Mean* yang digunakan sebagai berikut:

$$AD = \sum_{i=1}^N |x_i - \bar{x}| \quad (3.2)$$

Dimana AD = Penyimpangan absolut, $\sum_{i=1}^N$ = adalah jumlah nilai dimulai dari i ke N , x_i = Nilai x yang ke - i , dan \bar{x} = Nilai rata-rata. Gambaran struktur data hasil ekstraksi fitur dari parameter *Absolute Deviation Of Mean* yang akan digunakan sebagai *input* grafik visualisasi diilustrasikan pada Tabel 3.2 berikut:

Tabel 3.2 Struktur Data pada Fitur *Absolute Deviation Of Mean* untuk Visualisasi

Running		
Time	Nilai	Fitur
1	...	AD
2	...	AD
3	...	AD
...
10799	...	AD
10799	...	AD
10800	...	AD

Pada proses segmentasi berdasarkan *time domain* akan menghasilkan 6 segmen, dimana 1 segmen berdurasi 3 menit. Proses selanjutnya adalah ekstraksi fitur statistik di setiap segmen, hal ini bertujuan untuk mendapatkan ciri yang lebih spesifik dalam rentang waktu yang kecil. Fitur yang diusulkan adalah *Standard Deviation, Mean, Maximum, Minimum, Range, Variance, Skewnes dan Kurtosis*. Proses ekstraksi fitur seperti Tabel 3.3 menghasilkan 144 *instances* pada kelas *baseline* dan 720 *instances* pada kelas *playing*.

Tabel 3.3 Struktur Data pada Ekstraksi Fitur untuk Klasifikasi

Sampel	Fitur ke-1	Fitur ke-2	...	Fitur ke-7	Kelas
1	<i>Baseline</i>
2	<i>Baseline</i>
3	<i>Baseline</i>
...
862	<i>Playing</i>
863	<i>Playing</i>
864	<i>Playing</i>

Dari proses ekstraksi fitur yang akan dipergunakan untuk klasifikasi dan pengelompokan *class attribute*, maka diperoleh dataset yang *imbalanced class*.

Perbandingan *class attribute* yang didapatkan adalah 1 berbanding 5, seperti yang di tunjukan pada tabel 3.4 dibawah ini:

Tabel 3.4 Tabel Persentase *Class Attribute*

<i>Class Attribute</i>	Jumlah	Persentase
Kelas <i>Baseline</i>	144	16.67%
Kelas <i>Playing</i>	720	83.33%
Total	864	100%

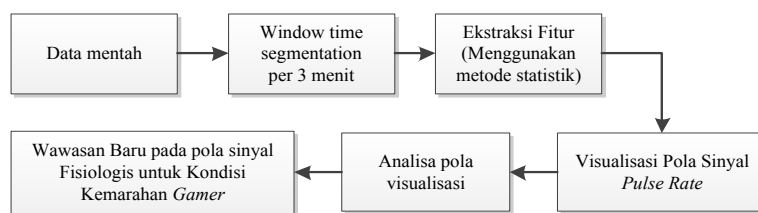
3.3.2 Normalisasi Fitur untuk Klasifikasi

Hasil ekstraksi fitur yang dihasilkan mempunyai nilai atribut dengan rentang yang besar. Oleh karena itu, agar persebaran data tidak terlalu jauh dan dominasi variabel yang berdekatan tidak terjadi, maka pada penelitian ini menggunakan normalisasi. SVM membutuhkan normalisasi, karena algoritma klasifikasi ini menggunakan margin atau jarak antara *hyperplane* dengan data terdekat pada masing-masing kelas. Penelitian ini menggunakan skema normaliasasi *Min-Max*, yaitu dengan skala -1 ke 1. Skema ini dipilih karena merujuk dari penelitian yang telah dilakukan oleh Weijun *et al* [68] yang menyimpulkan bahwa skema normaliasasi *Min-Max* tersebut pada algoritma klasifikasi SVM memiliki tingkat akurasi dan waktu komputasi terbaik jika dibandingkan dengan skema normalisasi lainnya seperti: *Zero-Mean normalization*, *Sigmoidal normalization*, *Softmax Normalization*, *Decimal Scaling*, dan *Max Normalization*.

3.3.3 Visualisasi Pola *Pulse Rate*

Nilai yang dipergunakan sebagai *input* untuk visualisasi data diperoleh dari ekstraksi fitur yang telah disebutkan pada penjelasan sebelumnya. Metode visualisasi yang dipergunakan adalah dengan mempresentasikan tren perubahan sinyal *pulse rate* berdasarkan seri waktu ketika tahap *baseline* dan bermain *video game*. Hal ini dilakukan untuk mengetahui pola maupun jangkauan nilai yang dipresentasikan. Pada parameter *Difference Value Of Mean* telah dipergunakan oleh Wibawa *et al* [22] untuk memvisualisasikan perbedaan nilai rata-rata pada

emosi dasar orang lanjut usia. Namun pada visualisasi yang diusulkan, menggunakan fase-fase sebagai pembagi waktu. Terdapat 5 fase yang akan divisualisasikan, fase pertama merupakan hasil perbandingan nilai rata-rata antara 3 menit pertama tahap bermain *video game* dengan 3 menit *baseline*, sedangkan fase ke-2 adalah perbandingan nilai rata-rata 3 menit ke-2 bermain *video game* dengan 3 menit ke-3 bermain *video game*. Hal ini dilakukan hingga mendapatkan 5 fase yang akan divisualisasikan bagi setiap subjek. Selanjutnya pada parameter *Absolute Deviation Of Mean* telah dipergunakan dalam berbagai penelitian sebelumnya untuk mempresentasikan pola perubahan data dalam seri waktu. Parameter ini telah berhasil mendeteksi keberadaan anomali yang kemungkinan adalah prekursor gempabumi pada penelitian yang dilakukan oleh Sulastri dkk [72].



Gambar 3.6 Diagram Alir Untuk Analisa Pola Menggunakan Visualisasi

3.3.4 Klasifikasi Menggunakan SVM

Algoritma klasifikasi yang digunakan adalah SVM. Aplikasi untuk proses klasifikasi menggunakan Weka 3.9.4 dan *Library SVM* yang digunakan adalah LibSVM (*Library Support Vector Machine*) yang dikembangkan oleh Chang *et al.*[49]. Terdapat permasalahan yang pada model klasifikasi SVM yang diusulkan untuk pengenalan emosi marah menggunakan sinyal *pulse rate*, yaitu permasalahan dalam menentukan *hyperparameter* yang optimal dan kelas atribut yang tidak seimbang (*imbalanced class*). *Imbalanced class* yaitu keadaan ketika terdapat salah satu kelas yang memiliki jumlah *instances* lebih banyak dari kelas lainnya. Sehingga, pada penelitian ini akan digunakan metode *Grid Search* untuk menentukan *hyperparameter* yang optimal dan menggunakan metode Bagging (*Bootstrap Aggregating*) untuk mengatasi permasalahan *imbalanced class*. Pada metode *Grid Search* bertujuan untuk memilih parameter *Cost* (C) dan *Gamma* (γ),

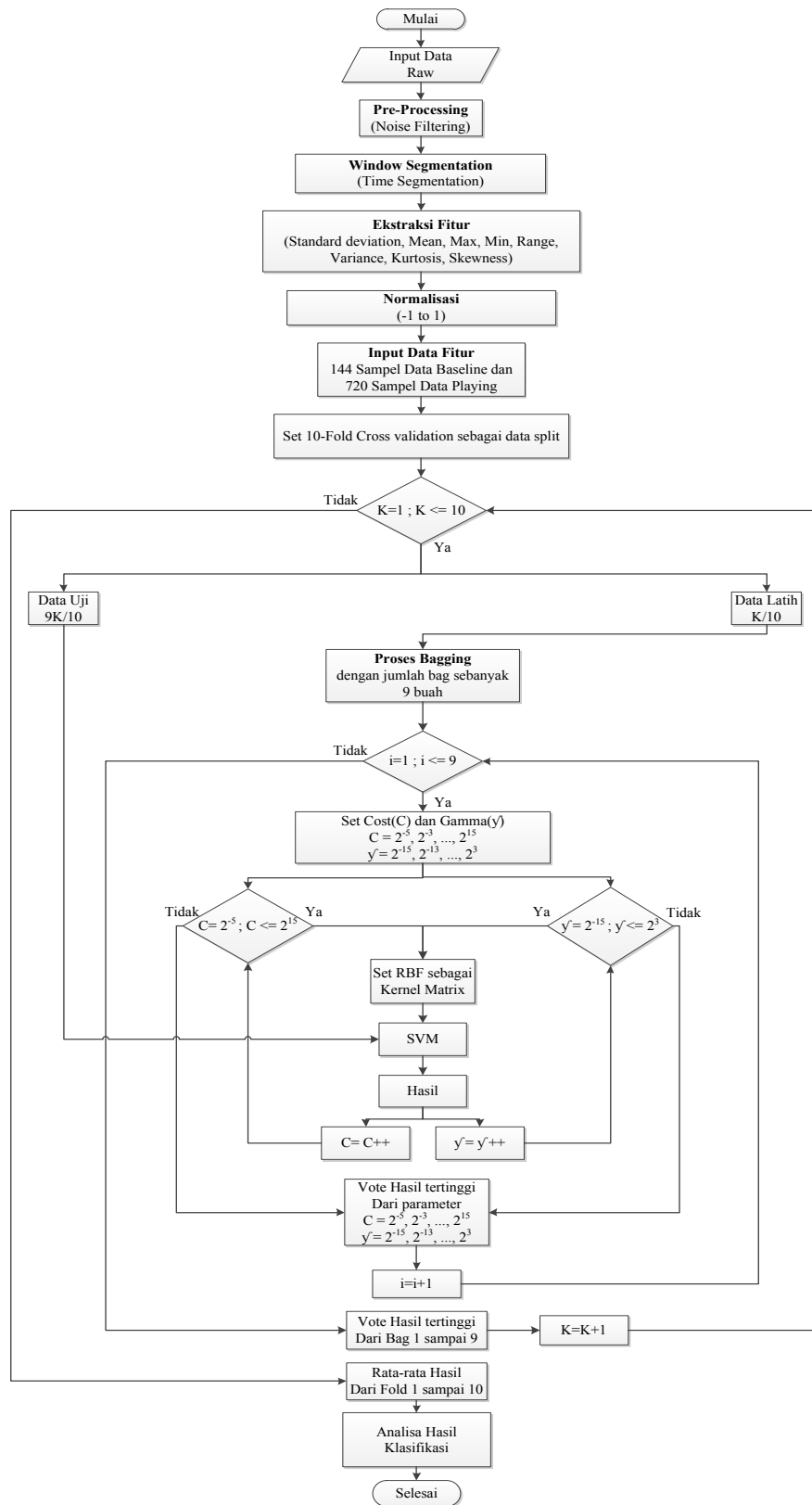
seperti yang dilakukan oleh Hsu *et al* [63], agar pada setiap iterasi digunakan penambahan parameter secara eksponensial dimana pada pencarian nilai parameter yang digunakan di setiap iterasi pada kernel RBF berada di *range* $C = 2^{-5}, 2^{-3}, \dots, 2^{15}$ dan $\gamma = 2^{-15}, 2^{-13}, \dots, 2^3$. Sedangkan metode Bagging ini memiliki cara kerja dengan membangun beberapa model algoritma yang sama (*multiple classifier*) dari sub-sampel yang berbeda dari data latih. Jumlah *Bag* yang dibuat sebanyak 9 buah, hal ini dilakukan karena pada proses *k-fold cross-validation* sebelum proses Bagging, *instances* yang dialokasikan sebagai data *training* sebanyak 90% dari keseluruhan *instances*, sehingga jumlah *instances* yang akan dibagi pada proses Bagging adalah 10% dari keseluruhan *instances* pada setiap *bag*, dan setiap *bag* akan memiliki jumlah yang sama dengan data *testing*. Skenario yang diusulkan adalah dengan cara: pada setiap *bag* yang dihasilkan oleh metode Bagging terdapat *classifier* SVM yang akan bekerja melakukan klasifikasi, selanjutnya pada setiap proses klasifikasi yang terjadi, metode *Grid Search* diimplementasikan untuk menentukan parameter *Cost* dan *Gamma* yang menghasilkan nilai prediksi yang terbaik. Kemudian, model validasi yang digunakan untuk menggeneralisasi data adalah *k-fold Cross-validation* dengan nilai $k=10$.

Langkah-langkah model klasifikasi SVM yang diusulkan untuk pengenalan emosi marah menggunakan sinyal *pulse rate* adalah sebagai berikut:

1. Melakukan *Pre-Processing*, yaitu proses *noise filtering*, yaitu menyaring data yang tidak diperlukan yang mengandung kesalahan.
2. Melakukan *window time segmentation*, yaitu membagi data raw berdasarkan *time domain* sebesar 15 detik per segmen. Maka didapatkan sebanyak 12 segmen pada kelas *baseline* dan 60 segmen pada kelas *playing*.
3. Melakukan ekstraksi fitur disetiap segmen data, yaitu fitur yang digunakan adalah *Standard Deviation, Mean, Maximum, Minimum, Range, Variance, Skewnes* dan *kurtosis*. Maka didapatkan sebanyak 144 *instances* pada kelas *baseline* dan 720 *instances* pada kelas *playing* di 12 orang subjek.
4. Melakukan normalisasi dengan nilai minimum -1 dan maksimum 1.

5. Menggunakan perlakuan *10-fold cross validation* sebagai data *split training* dan kernel matrix RBF pada algoritma SVM.
6. Menggunakan metode Bagging (*Bootstrap Aggregating*) dengan 9 *bag* pada analisa Bagging-SVM.
7. Menggunakan metode *Grid Search* untuk menentukan *hyperparameter* yang optimal di rentang nilai $C = 2^{-5}, 2^{-3}, \dots, 2^{15}$ dan $\gamma = 2^{-15}, 2^{-13}, \dots, 2^3$ pada prose klasifikasi SVM.
8. Menentukan parameter C dan γ dengan akurasi SVM terbaik pada analisa *Grid Search-SVM* dengan cara *voting*.
9. Melakukan *voting* terhadap hasil terbaik dari hasil klasifikasi SVM yang dihasilkan pada setiap *bag* metode Bagging.
10. Menentukan hasil terbaik dari setiap proses *fold cross validation* dengan menggunakan *averaging*.
11. Melakukan analisa di setiap hasil skenario percobaan.

Flowchart pada Gambar 3.7 berikut untuk mengilustrasikan aliran proses ususal model klasifikasi pada penelitian ini.



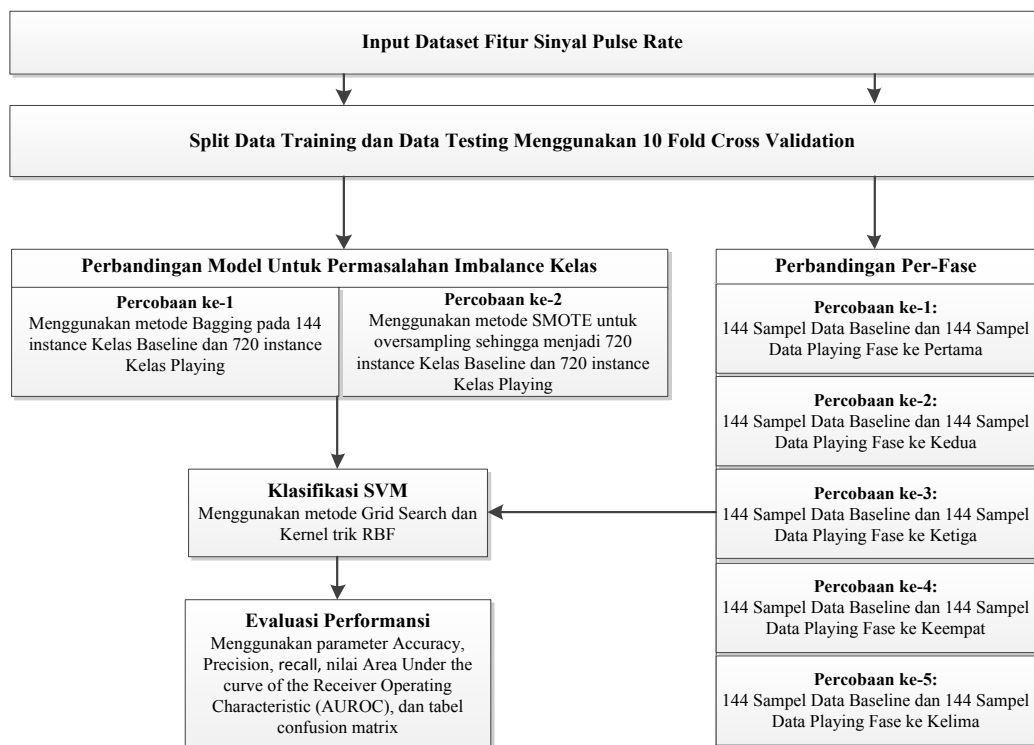
Gambar 3.7 Flowchart Usulan Model Bagging-SVM

3.4 Tahap Akhir

3.4.1 Skenario Pengujian

Penelitian ini mempunyai skenario seperti pada Gambar 3.7, dimana masing-masing skenario percobaan akan mendapatkan evaluasi sebagai kontrol terhadap pengujian selanjutnya. Skenario pertama, yaitu dengan melakukan percobaan perbandingan model untuk permasalahan *imbalanced class*. Model klasifikasi yang diusulkan akan dibandingkan dengan metode lain yang menggunakan pendekatan data untuk mengatasi permasalahan *imbalanced class*. Metode yang dipilih sebagai perbandingan yaitu SMOTE dengan *oversampling* pada sampel kelas minor. *Oversampling* yang dilakukan sebesar 40%, sesuai dengan persentase perbedaan kelas *baseline* dan *playing*. Metode SMOTE ini telah dilakukan oleh Didin dkk [67] untuk mengatasi *imbalanced class* pada data *network flow* menggunakan *Naive Bayes Classifier* dan mendapatkan akurasi sebesar 99,19%.

Skenario ke-2 dilakukan dengan cara membagi sampel data (*instances*) kelas *playing* menjadi 5 fase berdasarkan seri waktu. Fase pertama yaitu sampel data pada 3 menit pertama ketika bermain *video game*, fase ke-2 yaitu sampel data pada 3 menit dua ketika bermain *video game* dan seterusnya hingga fase ke lima. Hal ini dilakukan untuk menjawab pertanyaan, pada fase keberapa ukuran tingkat kebenaran (parameter akurasi) dari prediksi kelas atribut (emosi marah) yang diasumsikan bernilai rendah atau tinggi?. Selain itu juga untuk mengetahui persamaan pola yang didapatkan pada analisa menggunakan visualisasi parameter *Difference Value Of Mean (DV)* and *Absolute Deviation Of Mean (AD)*.



Gambar 3.8 Skenario Keseluruhan Pengujian

3.4.2 Evaluasi Performansi

Evaluasi performansi yang dilakukan pada perbandingan model untuk permasalahan *imbalanced class* dengan menggunakan parameter Akurasi, Presisi atau *positive predictive value* (PPV), Sensitivitas atau *recall* atau *true positive rate* (TPR), *Specificity* atau *true negative rate* (TNR), *Negative predictive values* (NPV), *False positive rate* (FPR), *False negative rate* (FNR), Tabel *Confusion Matrix*.

Tujuan dari tahap ini adalah untuk mengukur kinerja model klasifier yang diusulkan. Parameter *precision*, *recall* dipilih karena pada penelitian yang dilakukan oleh Saito *et al* [69] menyatakan bahwa parameter tersebut tidak terpengaruh terhadap penambahan sampel / *instances* yang klasifikasikan sebagai *True Negative*, sehingga hal tersebut lebih informatif dalam evaluasi model pada keadaan *imbalanced class* di proses klasifikasi biner jika dibandingkan dengan nilai *Area Under the ROC (Receiver Operating Characteristic)*.

BAB 4

HASIL DAN PEMBAHASAN

Pada bagian ini, akan diberikan uraian yang memuat pembahasan dari pola *pulse rate* pada remaja yang memainkan *video game* yang memiliki emosi marah sebagai emosi yang dominan. Selanjutnya penyelesaian tahapan-tahapan tujuan penelitian yang telah diberikan dalam metodologi penelitian, hingga diperoleh informasi dalam mendeteksi emosi marah dan mengetahui tren perubahan pola sinyal berdasarkan waktu dan menjawab pertanyaan: apakah algoritma klasifikasi SVM dapat digunakan untuk memprediksi emosi marah berbasis sinyal *pulse rate*.

4.1 Hasil Pengambilan Data

Pada sub-bab sebelumnya telah diterangkan bahwa proses validasi pemilihan *video game* sebagai stimulan dilakukan dengan cara survey ke komunitas pengembang *video game* di Kota Surabaya. Hasil survey didapatkan sebesar 77% memilih *video game* DOTA2 sebagai *video game* mempunyai emosi marah sebagai emosi yang dominan dan *video game* yang terlama dalam satu babak permainannya. Berikut pada Tabel 4.1 hasil survey yang didapatkan.

Tabel 4.1 Hasil Survey *video game*

No.	Nama <i>Game</i>	Persentase	Durasi per-babak
1	<i>Defense of the Ancients (DotA) 2</i>	77%	30-60 menit
2	<i>Pro Evolution Soccer (PES)</i>	18%	10-20 menit
3	<i>Game lainnya</i>	5%	-

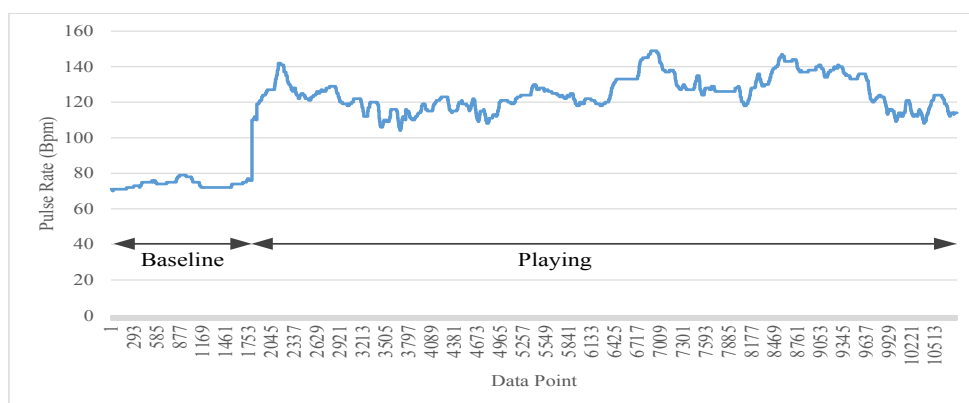
Jumlah subjek yang dilibatkan sebanyak 12 orang remaja dengan rata-rata usia 22 ± 1.8 tahun. Data yang didapatkan dari subjek, dibagi kedalam dua kelas, yaitu kelas *baseline* dan kelas bermain *video game (playing)*. Durasi kelas *baseline* selama 3 menit sedangkan kelas *playing* 15 menit, dimana dalam 1 detik

terdapat 10 sampel *pulse rate* yang terekam. Seluruh subjek telah dikonfirmasi menggunakan kuesioner bahwa tidak ada gangguan teknis dalam bermain *video game* seperti jaringan komputer yang tidak stabil dan lainnya. Dari pengisian kuesioner terdapat isian range 1 hingga 10 untuk pertanyaan seberapa emosi yang dirasakan setelah bermain *video game*, dan didapatkan rata-rata skor 6.8. Durasi bermain *video game* setiap subjek disajikan pada Tabel 4.2. Informasi awal yang dapat disimpulkan dari durasi bermain *video game* tersebut adalah tingginya intensitas emosi marah yang terjadi pada subjek yang bermain *video game*.

Tabel 4.2 Perincian Subjek Penelitian

Kode Subjek	Usia (Tahun)	Jenis Kelamin	Durasi Bermain Game dalam 1 Minggu (Jam)	Tingkat Emosi yang dirasakan Setelah Bermain Game (Skala 1-10)
R1	22	Pria	8	7
R2	25	Pria	10	8
R3	22	Pria	30	7
R4	25	Pria	14	9
R5	21	Pria	21	8
R6	25	Pria	7	8
R7	21	Pria	13	8
R8	23	Pria	30	9
R9	24	Wanita	14	8
R10	25	Wanita	12	8
R11	20	Pria	14	7
R12	22	Pria	4	9
Rata-rata	22 ± 1.8		14.75 ± 8.31	8

Data bermain *video game* yang diambil merupakan data yang sudah terkonfirmasi dengan subjek ketika permainan memasuki tahap yang diharapkan (emosi marah), sehingga jika terdapat kelebihan waktu maka akan di sesuaikan dengan waktu subjek mendapat stimulasi emosi marah. Data *pulse rate* yang tercatat sebanyak 600 sampel selama 1 menit. Hal ini memiliki arti bahwa pada *baseline* terdapat 1800 sampel dan pada bermain *video game* terdapat 9000 sampel. Gambar 4.1 dibawah merupakan tipikal data yang didapatkan dari salah satu subjek.



Gambar 4.1 Karakteristik Data Hasil Pengukuran pada Salah Satu Subjek

Sinyal fisiologi dalam hal ini *pulse rate* mempunyai pola fluktuasi dalam range tertentu, oleh karena itu perbedaan pola antar waktu dapat digali informasinya menggunakan grafik visualisasi dan prediksi menggunakan klasifikasi. Perubahan pola dapat dipengaruhi menggunakan stimulan, termasuk pengamatannya dikaitkan berdasarkan respon dan kondisi yang terjadi pada subjek. Pola yang terjadi pada Gambar 4.1 menjadi informasi yang dapat diamati secara langsung bagaimana fluktuasi yang terjadi terhadap stimulan bermain *video game*.

4.2 Hasil Pengolahan Data

Bagian ini akan menjelaskan hasil yang diperoleh dari metode yang diusulkan untuk menentukan pola sinyal *pulse rate* pada subjek yang memainkan *video game* DOTA2 sebagai *video game* yang mempunyai emosi marah sebagai emosi yang dominan. Dalam satu putaran permainan, memfilter menjadi selama

15 menit ketika para subjek mengkonfirmasi bahwa ini adalah waktu reaksi yang disebabkan oleh permainan. Dengan demikian, menurut Gambar 4.1, hasil yang diperoleh dibagi menjadi dua kelas, *baseline* dan *playing*. Telah menjelaskan dalam metode bagian metode penelitian, dengan mengusulkan 2 metode pengenalan emosi untuk menentukan pola sinyal *pulse rate* yang terjadi ketika subjek bermain *video game* dalam keadaan marah.

4.2.1 Hasil Ekstraksi Fitur untuk Visualisasi

Fitur pertama yang diusulkan sebagai *input* grafik visualisasi adalah *Difference Value Of Mean*. Pada Fitur ini, seperti yang telah dijelaskan pada sub-bab sebelumnya, terdapat 5 fase yang didapatkan dari hasil mencari nilai perbedaan dari nilai rata-rata per 3 menit. Fitur dari parameter *Difference Value Of Mean* yang diperoleh disajikan pada Tabel 4.3

Tabel 4.3 Hasil Ekstraksi Fitur *Difference Value Of Mean* untuk Visualisasi

Subjek	Fase ke-1	Fase ke-2	...	Fase ke-5	Fitur
1	20	28	...	12	DV
2	8	12	...	12	DV
3	52	44	...	52	DV
...
10	0	4	...	4	DV
11	12	8	...	12	DV
12	4	8	...	12	DV

Fitur berikutnya untuk *input* grafik visualisasi adalah *Absolute Deviation Of Mean*. Fitur ini diperoleh dari jumlah nilai *Absolute Deviation Of Mean* keseluruhan subjek yang diikutsertakan dalam penelitian. . Fitur dari parameter *Absolute Deviation Of Mean* yang diperoleh disajikan pada Tabel 4.4.

Tabel 4.4 Hasil Ekstraksi Fitur *Absolute Deviation Of Mean* untuk Visualisasi

Running Time	Nilai	Fitur
1	0.854409722	AD
2	0.854409722	AD
3	0.854409722	AD
...
10799	1.457458333	AD
10799	1.457458333	AD
10800	1.457458333	AD

4.2.2 Hasil Ekstraksi Fitur untuk Klasifikasi

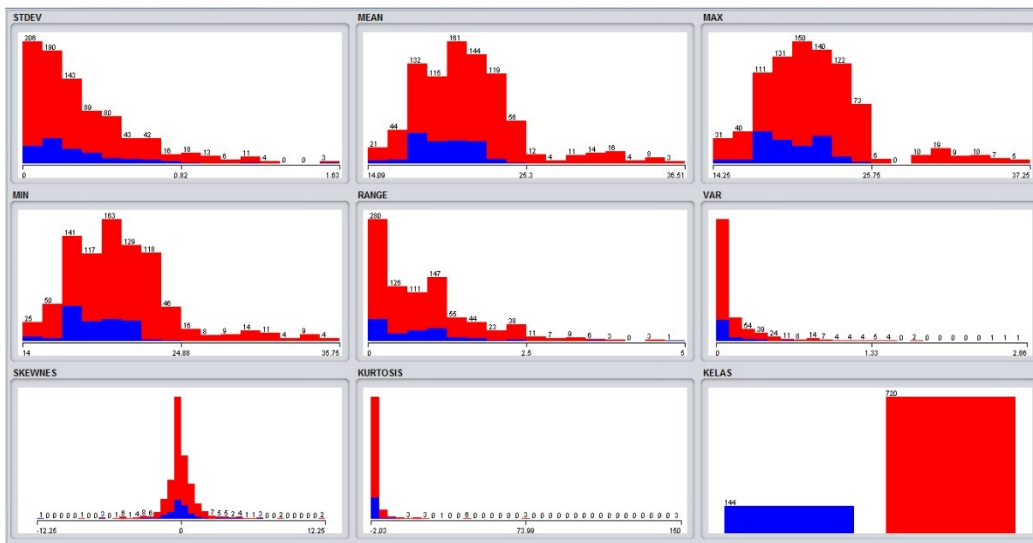
Setelah fitur-fitur untuk visualisasi, terdapat juga fitur untuk klasifikasi. Fitur ini sebagai *instances* yang akan digunakan oleh algoritma klasifikasi untuk memprediksi emosi marah yang terjadi saat bermain *video game*. Prediksi yang dihasilkan akan menjadi informasi seberapa jauh kebenaran prediksi dari model klasifikasi yang diusulkan oleh penelitian ini. Fitur-fiturnya adalah sebagai berikut: *Standard Deviation, Mean, Maximum, Range, Minimum, Variance, Skewnes dan Kurtosis*. Nilai dari hasil ekstraksi fitur-fitur tersebut, disajikan pada Tabel 4.5 dibawah ini:

Tabel 4.5 Hasil Ekstraksi Fitur untuk Klasifikasi SVM

No	STDEV	MEAN	MAX	MIN	RANGE	VAR	SKEWNES	KURTOSIS	KELAS
1	2.3822294	71.378333	77	69	1.25	5.6750167	0.710048	-0.6463552	<i>Baseline</i>
2	3.6187947	64.373333	72	59	0.5	13.095675	0.8683343	-0.5598985	<i>Baseline</i>
3	2.1926131	73.895	79	70	0.75	4.8155392	0.6122288	-0.3045834	<i>Baseline</i>
4	4.9044905	78.635	90	72	1	24.054027	0.70239	-0.5118614	<i>Baseline</i>
5	1.8073338	70.055556	74	66	0.75	3.2676178	0.0601024	-0.5350975	<i>Baseline</i>
...
434	0.3494203	19.861667	20.25	19.25	1	0.1220945	-0.6138602	-0.8883479	<i>Playing</i>
435	0.0491539	19.24	19.25	19	0.25	0.0024161	-4.7424118	20.767212	<i>Playing</i>
436	0.2815197	19.755	20	19.25	0.75	0.0792534	-0.8685864	-0.6664786	<i>Playing</i>
437	0.2279664	18.923333	19.25	18.75	0.5	0.0519687	0.6486891	-1.4902607	<i>Playing</i>
...
862	0	23.75	23.75	23.75	0	0	0	0	<i>Playing</i>
863	0.0351173	23.995	24	23.75	0.25	0.0012332	-6.9266022	46.599015	<i>Playing</i>
864	0	24	24	24	0	0	0	0	<i>Playing</i>

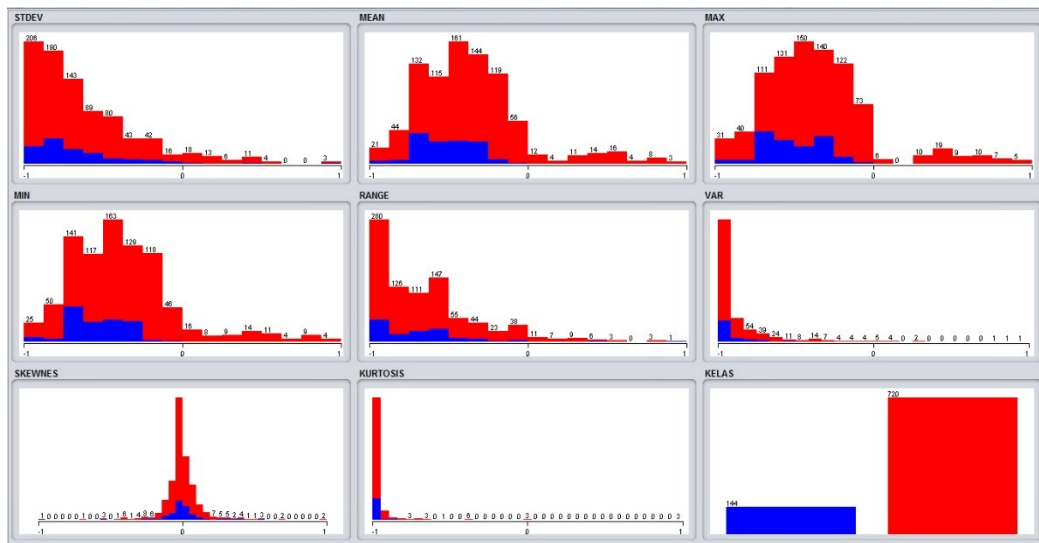
4.2.3 Hasil Normalisasi Fitur Visualisasi

Hasil fitur untuk proses klasifikasi yang didapatkan memiliki rentang nilai yang besar, oleh karena itu pada model klasifikasi yang diusulkan ini menggunakan normalisasi. Penelitian ini menggunakan skema normalisasi *Min-Max*, yaitu dengan skala -1 ke 1. Berikut adalah grafik-grafik perbedaan fitur-fitur yang belum dinormalisasi dan sesudah dinormalisasi.



Gambar 4.2 Grafik Persebaran Data Pada Fitur-Fitur Sebelum Normalisasi

Penerapan visualisasi kedalam grafik membantu memudahkan memahami persebaran data dan pola di setiap fitur-fitur. Dari grafik persebaran data pada fitur-fitur sebelum dan sesudah normalisasi dapat disimpulkan bahwa pola yang terbentuk tidak berubah.



Gambar 4.3 Grafik Persebaran Data Pada Fitur-Fitur Sesudah Normalisasi

4.3 Hasil Visualisasi Data *Pulse Rate*

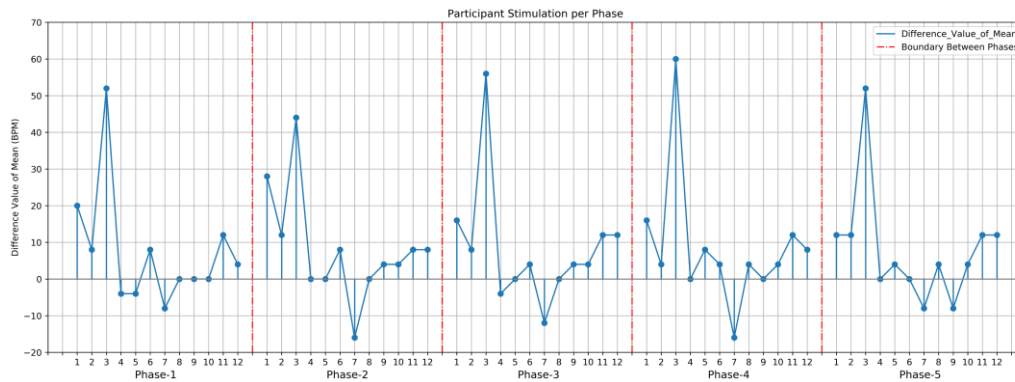
Bagian ini menjelaskan hasil yang diperoleh dari metode yang diusulkan untuk menentukan pola sinyal PR untuk subjek yang memainkan DOTA2 sebagai permainan yang mempunyai emosi marah sebagai emosi yang dominan. Telah dijelaskan di bagian metode analisis data dengan mengusulkan 2 metode pengukuran untuk menentukan pola sinyal PR yang terjadi ketika subjek bermain *video game* dalam kemarahan.

4.3.1 Hasil Visualisasi Pada Parameter *Difference Value Of Mean (DV)*

Terdapat 5 fase pada pola visualisasi yang terbentuk sebagai hasil dari metode DV. Hasil ini menggambarkan perubahan pola *pulse rate* dari setiap subjek setiap 3 menit selama bermain *video game* selama 15 menit. Dari Gambar 4.4, dapat dilihat perubahan yang signifikan pada pola *pulse rate* di fase ke-3 dan ke-4 dan ke-5. Ini menunjukkan bahwa emosi kemarahan telah diinduksi. Sedangkan pada fase ke-3 dan ke-5, merupakan fase yang terdapat paling banyak subjek yang mengalami perubahan yang signifikan, dengan masing-masing 5 subjek.

Sementara itu, pada fase pertama, adalah fase dimana terdapatnya subjek yang paling sedikit mengalami perubahan yang signifikan. Perubahan signifikan

yang ditentukan adalah perubahan nilai parameter lebih dari 10 bpm. Perubahan yang terjadi termasuk kenaikan dan penurunan nilai *pulse rate*. Dari 12 subjek, 7 dari subjek menunjukkan perubahan *pulse rate* lebih dari 10 bpm dibandingkan dengan kondisi awal dan kondisi setiap fase bermain *video game*.



Gambar 4.4 Representasi Grafis *Difference Value Of Mean (DV)* Per Fase

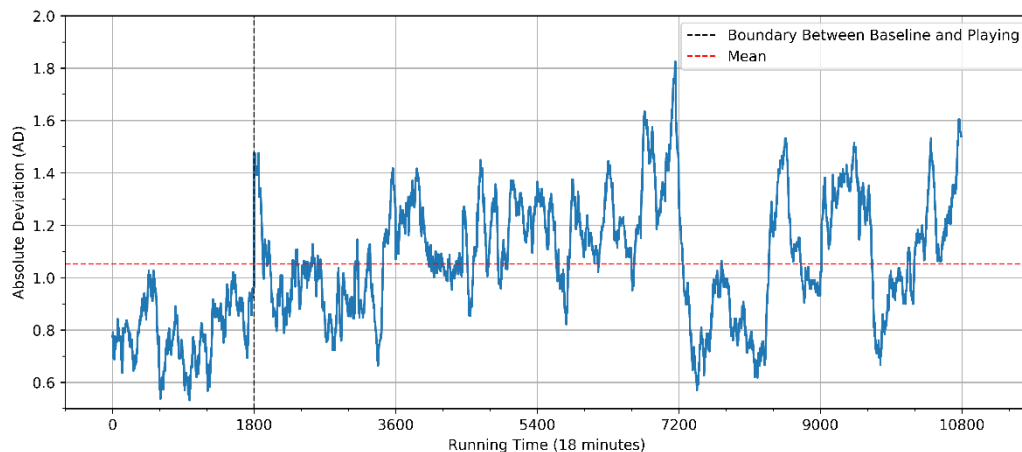
Tabel 4.6 menunjukkan nilai minimum dan maksimum dari masing-masing subjek. Terdapat 8 subjek yang memiliki *pulse rate* di atas 100 bpm saat bermain *video game*. Namun, ada 1 subjek yang memiliki perbedaan yang sangat signifikan, yaitu subjek ke-3 dengan nilai rata-rata *pulse rate* sebesar 73 bpm pada fase *baseline* dan 125 bpm pada fase bermain. Hal yang sama terjadi pada subjek pertama, keenam, kesepuluh dan kesebelas. Peningkatan *pulse rate* terjadi di atas 100bpm. Ini dapat menjadi perhatian lebih lanjut untuk dijelaskan secara medis karena memiliki potensi untuk menjadi takikardia [55]. Dengan hanya melihat nilai-nilai ini, kita dapat menjelaskan bahwa sulit untuk melihat apakah kemarahan subjek dipicu oleh permainan. Namun, dengan menggunakan parameter *Difference Value Of Mean* dapat menjelaskan bagaimana setiap fase berubah dalam bentuk grafik visualisasi.

Tabel 4.6 Hasil Pengambilan Data *Pulse Rate* Per Subjek

Kode Subjek	<i>Baseline (bpm)</i>			<i>Playing (bpm)</i>		
	<i>Mean</i>	<i>Max</i>	<i>Min</i>	<i>Mean</i>	<i>Max</i>	<i>Min</i>
R1	71.38	77	69	90.69	107	70
R2	64.37	72	59	73.35	85	66
R3	73.90	79	70	125.16	149	104
R4	78.64	90	72	77.50	90	65
R5	70.06	74	66	73.31	85	63
R6	86.64	94	81	92.74	104	79
R7	75.49	99	69	63.97	73	56
R8	79.57	86	74	81.66	93	71
R9	81.98	88	75	79.65	93	63
R10	87.24	96	80	91.06	105	83
R11	85.61	95	72	94.56	103	83
R12	84.22	93	75	88.57	97	82

4.3.2 Hasil Visualisasi Pada Parameter *Absolute Deviation Of Mean (AD)*

Di bagian ini, parameter yang digunakan sebagai input untuk grafik visualisasi adalah *Absolute Deviation Of Mean (AD)*. Nilai AD dihitung dari semua subjek menjadi satu deret nilai. Sebelumnya percobaan beberapa parameter dilakukan agar visualisasi yang terbentuk dapat menghadirkan kondisi yang realistis. Namun, merujuk pada studi sebelumnya, pemilihan parameter ini dapat menggambarkan bagaimana pola terjadi antara *baseline* dan bermain *video game*. Pada garis batas antara garis *baseline* dan bermain *video game*, inisiasi emosi marah telah terjadi sebagaimana diilustrasikan dalam grafik visualisasi pada Gambar 4.5.



Gambar 4.5 Representasi Grafis *Absolute Deviation Of Mean (AD)* dari Semua Subjek

Peningkatan di atas rata-rata dalam fase bermain terjadi dalam tiga menit ke-2 (3600 hingga 5400 poin), ke-3 (5400 hingga 7200 poin), dan tiga menit ke-5 (9000 hingga 10800 poin). Sementara itu, puncak kenaikan berada di tiga menit ke-3 (5400 hingga 7200 poin). Hasil ini juga linier dengan hasil yang diperoleh saat menggunakan parameter *Difference Value Of Mean*. Pada parameter *Difference Value Of Mean*, sebaran nilai *baseline* berada di bawah nilai rata-rata. Kemudian, skenario untuk membuat subjek tetap tenang ketika mengambil data *pulse rate* di *baseline* adalah seperti yang diharapkan. Selain itu, pengamatan yang dapat dilakukan dari hasil grafik ini adalah distribusi data *outlier* yang berada di luar band deviasi rata-rata. *Outlier* bisa tersebar di atas jangkauan nilai rata-rata ataupun dibawahnya. *Outlier* dapat diperoleh dari persamaan:

$$[\bar{x} - AD, \bar{x} + AD] \quad (4.1)$$

Hasil persamaan di atas diperoleh dari fase pertama setelah *baseline* (poin ke- 1801 hingga 3600), terdapat 379 *outlier*. Di fase ke-2 setelah *baseline* (poin ke- 3601 hingga 5400) terdapat 669 *outlier*, di fase ke-3 (poin ke- 5401 hingga 7200), terdapat 632 poin, di fase ke-4 (poin ke- 7201 hingga 9000) terdapat 849 poin *outlier*, dan di kisaran ke-5 (poin ke- 7201 hingga 9000) terdapat 999 poin *outlier*. Ini memberikan informasi bahwa sebagian besar penyimpangan rata-rata terjadi pada tahap ke-5. *Outlier* pada parameter *Difference Value Of Mean*

menunjukkan bahwa kenaikan ataupun penurunan *pulse rate* yang terjadi sangat tinggi pada fase-fase tersebut.

4.4 Hasil Klasifikasi Menggunakan SVM

Hasil klasifikasi yang akan dijabarkan menurut usulan model dan skenario percobaan yang diterangkan di bab 3. Pertama, usulan model klasifikasi pada penelitian ini adalah menerapkan metode Bagging untuk permasalahan *imbalanced class*, dan menerapkan metode *grid search* pada setiap proses klasifikasi SVM untuk menentukan *hyperparameter* yang optimal. Selanjutnya, menggunakan metode *k-fold cross validation* sebagai metode *split data training* dan *testing*.

Kedua, sebagai pembanding model yang diusulkan, dilakukannya percobaan metode SMOTE yang berbasis pendekatan data untuk mengatasi permasalahan *imbalanced class*. Cara kerja metode ini adalah dengan melakukan *oversampling* pada kelas minor hingga jumlah kelas minor sama dengan jumlah kelas mayor.

Ketiga, percobaan dilakukan dengan cara membagi kelas *playing* ke dalam 5 segmen berdasarkan waktu. Ukuran pembagian kelas *playing* sama dengan ukuran kelas *baseline*. Dimana ukuran kelas *baseline* adalah 3 menit dengan jumlah *class attribute* sama dengan 1800, maka ukuran setiap segmen pada kelas *playing* adalah 3 menit dengan jumlah *class attribute* sama dengan 1800, karena ukuran waktu pada kelas *playing* adalah 15 menit, maka kelas *playing* dapat dibagi menjadi 5 segmen. Percobaan ini memiliki tujuan untuk mendeteksi, pada fase keberapa performansi dari prediksi kelas atribut (emosi marah) yang diasumsikan bernilai rendah atau tinggi?. Selain itu juga untuk mengetahui persamaan pola yang didapatkan pada analisa menggunakan visualisasi parameter statistik *Difference Value Of Mean (DV)* and *Absolute Deviation Of Mean (AD)*.

4.4.1 Hasil Klasifikasi Algoritma SVM Menggunakan Metode Bagging

Pada metode ini, menggunakan *10-fold cross-validation* untuk membagi data *training* dan *testing* algoritma klasifikasi SVM. Data *training* yang didapatkan dari metode *cross-validation* akan dibagi kembali menjadi 9 bag

menggunakan metode Bagging. Pada proses Bagging akan dibuat *bootstrap* dengan cara *random sampling* secara penggantian, kemudian setiap *bootstrap* yang dibuat akan digunakan untuk melatih algoritma klasifikasi SVM.

Pada klasifikasi menggunakan algoritma SVM, metode yang akan digunakan untuk membantu proses klasifikasi data adalah metode *grid search*. Pada metode ini, pencarian parameter C dan γ untuk model yang optimal menggunakan prinsip *grid search*. Berdasarkan rekomendasi dari Hsu *et al* [63] pencarian parameter C dan γ di rentang nilai $C = 2^{-5}, 2^{-3}, \dots, 2^{15}$ dan $\gamma = 2^{-15}, 2^{-13}, \dots, 2^3$. Selanjutnya, matrix kernel *Radial Basis Function* (RBF) akan digunakan untuk membantu dalam mengatasi masalah *feature space*. Perulangan proses ini dilakukan hingga bag ke-9. Kemudian, penentuan parameter yang optimal menggunakan *majority vote* terhadap hasil klasifikasi yang didapatkan. Hasil yang didapatkan akan disimpan dan akan dilakukan kembali hingga split ke-10 dari split data yang yang dihasilkan oleh metode *10-fold cross-validation*. Algoritma SVM menggunakan metode Bagging yang digunakan adalah sebagai berikut:

Algoritma 1. SVM-Bagging

Input : *Instances* Kelas *Baseline* ; *Instances* Kelas *Playing* ; Jumlah *Bag* ; Parameter C (*cost*) ; Parameter γ (*gamma*).

Output : Akurasi ; Presisi atau *positive predictive value* (PPV) ; Sensitivitas atau *recall* atau *true positive rate* (TPR) ; *Specificity* atau *true negative rate* (TNR) ; *Negative predictive values* (NPV) ; *False positive rate* (FPR) ; *False negative rate* (FNR) ; Tabel *Confusion Matrix*.

Begin :

1. Menggunakan perlakuan *10-fold cross validation* sebagai metode *split* data *training* dan *testing*.
2. Menggunakan metode Bagging (*Bootstrap Aggregating*) dengan 9 *bag*.
3. Set kernel matriks RBF menggunakan persamaan:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right), \gamma > 0$$

4. Set parameter C dan γ di rentang nilai $C = 2^{-5}, 2^{-3}, \dots, 2^{15}$ dan $\gamma = 2^{-15}, 2^{-13}, \dots, 2^3$

5. Menggunakan klasifikasi SVM pada setiap *bag* dengan persamaan:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i$$

$$\text{subject to } y_i (w^T \varphi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, l$$

6. Menentukan hasil klasifikasi SVM terbaik yang dihasilkan pada setiap *bag* dengan cara *majority voting* dengan persamaan:

$$f(x) = \text{sign} \left(\sum_{i=1}^T \text{sign}(f_i(x)) \right)$$

7. Menentukan hasil klasifikasi SVM terbaik yang dihasilkan pada setiap *split* atau lipatan di proses *k fold cross validation* dengan cara *averaging*.

Hasil yang didapatkan pada pengujian metode Bagging (*Bootstrap Aggregating*) dengan menggunakan 144 *instances* pada kelas *baseline* dan 720 *instances* pada kelas *playing*, akurasi memiliki skor sebesar 82.98%, *specificity* atau *true negative rate* (TNR) memiliki skor sebesar 97.78% dan *negative predictive values* (NPV) memiliki skor 84.31%.

Pada penelitian ini, klasifikasi dilakukan untuk melihat bagaimana performansi *classifier* yang dihasilkan dapat menunjukkan tingkat prediksi kelas negatif (kelas *playing*) yang mempresentasikan emosi marah. Hal ini dilakukan dengan cara mendeteksi karakteristik data yang dikelompokkan ke kelas *playing* berbeda dengan karakteristik data yang dikelompokkan ke kelas *baseline*.

Parameter ukur akurasi memiliki interpretasi ukuran identifikasi atau prediksi yang benar dari label kelas *baseline* (tidak marah) dan *playing* (marah), *specificity* atau *true negative rate* (TNR) memiliki interpretasi berapa persen karakteristik data dari keadaan marah yang diprediksi dengan benar dari keseluruhan karakteristik data yang sebenarnya adalah kondisi marah atau proporsi karakteristik data dari keadaan marah yang diidentifikasi atau diprediksi dengan benar. Sedangkan *negative predictive values* (NPV) memiliki interpretasi

berapa persen karakteristik data yang benar-benar mempresentasikan kondisi marah dari keseluruhan karakteristik data yang diprediksi sebagai kondisi marah.

Parameter lain yang dapat diukur selain TNR dan NPV adalah presisi atau *positive predictive value* (PPV), sensitivitas atau *recall* atau *true positive rate* (TPR), *false positive rate* (FPR), *false negative rate* (FNR). Evaluasi performansi yang dihasilkan selain dari parameter tersebut, secara komprehensif juga dapat menggunakan tabel *confusion matrix* untuk mengetahui bagian setiap *instances* yang dikelompokkan oleh *classifier*.

Tabel 4.7 *Confusion Matrix* dari SVM Menggunakan Metode Bagging

AKTUAL	PREDIKSI	
	<i>Baseline</i>	<i>Playing</i>
<i>Baseline</i>	13	131
<i>Playing</i>	18	702

Persentase karakteristik data yang benar-benar mempresentasikan kondisi tidak marah dari keseluruhan karakteristik data yang diprediksi sebagai kondisi tidak marah atau PPV memiliki skor 44.83%. Sedangkan presentase karakteristik data dari keadaan tidak marah yang diidentifikasi atau diprediksi dengan benar atau TPR memiliki skor 9.02%.

Parameter evaluasi untuk prediksi yang salah yang digunakan adalah *False positive rate* (FPR) dan *False negative rate* (FNR). Persentase karakteristik data yang tidak termasuk ke prediksi kondisi marah dari keseluruhan karakteristik data yang dilabeli sebagai kondisi marah atau FPR memiliki skor 2.22%. Sedangkan FNR adalah persentase karakteristik data yang tidak termasuk ke prediksi kondisi tidak marah dari keseluruhan karakteristik data yang dilabeli sebagai kondisi tidak marah memiliki skor 90.97%. Keseluruhan hasil evaluasi dan persamaan yang digunakan ditampilkan pada Pada Tabel 4.8

Tabel 4.8 Hasil Klasifikasi Algoritma SVM Menggunakan Metode Bagging

Parameter penilaian	Persamaan	Hasil
Akurasi	$\frac{TN + TP}{TN + TP + FN + FP}$	82.98%
Presisi atau <i>positive predictive value</i> (PPV)	$\frac{TP}{TP + FP}$	44.83%
Sensitivitas atau <i>recall</i> atau <i>true positive rate</i> (TPR)	$\frac{TP}{TP + FN}$	9.02%
<i>Specificity</i> atau <i>true negative rate</i> (TNR)	$\frac{TN}{TN + FP}$	97.78%
<i>Negative predictive values</i> (NPV)	$\frac{TN}{TN + FN}$	84.31%
<i>False positive rate</i> (FPR)	$\frac{FP}{FP + TN}$	2.22%
<i>False negative rate</i> (FNR)	$\frac{FN}{FN + TP}$	90.97%

Bedasarkan dari hasil yang didapatkan pada pengujian metode Bagging, klasifikasi pada kondisi *imbalanced class* memiliki performansi yang tetap baik. Walaupun parameter ukur kelas positif (PPV dan TPR) memiliki skor dalam kategori rendah, yang mana hal tersebut dikarenakan proses *learning* terhadap kelas *baseline* tidak sebanyak kelas *playing*, namun pada proses Bagging dipembagian setiap *bag* yang digunakan untuk proses klasifikasi telah diseimbangkan jumlah *instances* antara data *training* dan data *testing*.

4.4.2 Hasil Klasifikasi Algoritma SVM Menggunakan Metode SMOTE

Metode ini dilakukan sebagai perbandingan performansi model klasifikasi yang di usulkan. Selain itu metode ini untuk mengetahui bagaimana performansi prediksi kelas *baseline* dan kelas *playing* dalam keadaan seimbang.

Pada metode ini, menggunakan *10-fold cross-validation* untuk membagi data *training* dan *testing*. Data training yang didapatkan dari metode *cross-validation* akan menjadi data *training* pada proses klasifikasi. Sedangkan data *testing*, akan menjadi data validasi pada proses klasifikasi. Perbedaan metode SMOTE dengan baging terletak pada proses ini, dimana pada metode SMOTE, data *training* yang dihasilkan dari proses *10-fold cross-validation* akan langsung menjadi data *training* yang diklsifikasikan oleh *classifier*.

Pada klasifikasi menggunakan algoritma SVM, metode yang akan digunakan untuk membantu proses klasifikasi data adalah metode *grid search*. Pada metode ini, pencarian parameter C dan γ untuk model yang optimal menggunakan prinsip *grid search*. Berdasarkan rekomendasi dari Hsu *et al* [63] pencarian paramter C dan γ di rentang nilai $C = 2^{-5}, 2^{-3}, \dots, 2^{15}$ dan $\gamma = 2^{-15}, 2^{-13}, \dots, 2^3$. Selanjutnya, matrix kernel *Radial Basis Function* (RBF) akan digunakan untuk membantu dalam mengatasi masalah *feature space*. Perulangan proses ini dilakukan hingga K ke-10. Kemudian, hasil yang didapatkan akan disimpan dan akan dilakukan kembali hingga split ke-10 dari split data yang yang dihasilkan oleh metode *10-fold cross-validation*. Algoritma SVM menggunakan metode SMOTE yang digunakan adalah sebagai berikut:

Algoritma 1. SVM-SMOTE

Input : *Instances* Kelas *Baseline* ; *Instances* Kelas *Playing* ; Parameter C (*cost*) ; Parameter γ (*gamma*).

Output : Akurasi ; Presisi atau *positive predictive value* (PPV) ; Sensitivitas atau *recall* atau *true positive rate* (TPR) ; *Specificity* atau *true negative rate* (TNR) ; *Negative predictive values* (NPV) ; *False positive rate* (FPR) ; *False negative rate* (FNR) ; Tabel *Confusion Matrix*.

Begin :

1. Melakukan *oversampling* 40% pada kelas *baseline* menggunakan metode SMOTE.
2. Menggunakan perlakuan *10-fold cross validation* sebagai metode *split* data *training* dan *testing*.
3. Set kernel matriks RBF menggunakan persamaan:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2\right), \gamma > 0$$

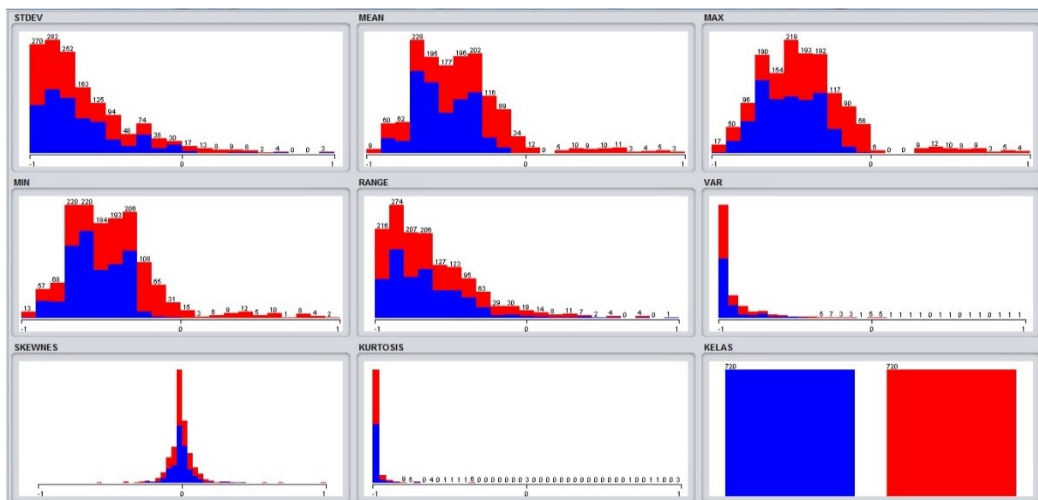
4. Set parameter C dan γ di rentang nilai $C = 2^{-5}, 2^{-3}, \dots, 2^{15}$ dan $\gamma = 2^{-15}, 2^{-13}, \dots, 2^3$
5. Menggunakan klasifikasi SVM dari hasil *split* dengan persamaan:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i$$

$$\text{subject to } y_i (w^T \varphi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, l$$

6. Menentukan hasil klasifikasi SVM terbaik yang dihasilkan pada setiap *split* atau lipatan di proses *k fold cross validation* dengan cara *averaging*.

Proses *oversampling* yang dilakukan menghasilkan pola fitur yang berbeda dengan pola sebelum *oversampling*, Gambar 4.6 menunjukkan pola fitur yang terbentuk dari proses *oversampling*.



Gambar 4.6 Grafik Persebaran Data Pada Fitur-Fitur Metode SMOTE

Evaluasi performansi yang dihasilkan untuk dapat dijabarkan menjadi parameter evaluasi lain, secara komprehensif menggunakan tabel *confusion matrix*. Tabel 4.9 menampilkan setiap *instances* yang dikelompokan oleh *classifier*.

Tabel 4.9 *Confusion Matrix* dari SVM Menggunakan Metode SMOTE

AKTUAL	PREDIKSI	
	<i>Baseline</i>	<i>Playing</i>
<i>Baseline</i>	443	277
<i>Playing</i>	65	655

Hasil yang didapatkan, skor TNR yang menunjukkan prediksi *instances* kelas *playing* terhadap *instances* yang di beri label *playing*, memiliki skor yang lebih rendah dari pada skor TNR metode Bagging. Kemudian, pada skor NPV metode SMOTE memiliki skor yang lebih rendah dari pada skor TNR metode Bagging. Pada skor TPR metode SMOTE memiliki skor yang lebih tinggi dari pada skor TPR metode Bagging. Pada skor PPV metode SMOTE memiliki skor yang lebih tinggi dari pada skor PPV metode Bagging. Pada skor FPR metode SMOTE memiliki skor yang lebih tinggi dari pada skor FPR metode Bagging. Pada skor FNR metode SMOTE memiliki skor yang lebih rendah dari pada skor FNR metode Bagging.

Kesimpulan yang dapat ditarik adalah penambahan *instances* pada kelas minoritas (*baseline*) memiliki perubahan prediksi yang signifikan pada parameter PPV, TPR, dan FNR. Hal ini disebabkan oleh proses *learning* pada kelas *baseline* untuk menentukan prediksi kelas *baseline* atau tidak termasuk kelas *baseline* menjadi lebih banyak.

Tabel 4.10 Hasil Klasifikasi Algoritma SVM Menggunakan Metode SMOTE

Parameter penilaian	Persamaan	Hasil
Akurasi	$\frac{TN + TP}{TN + TP + FN + FP}$	76.25%
Presisi atau <i>positive predictive value</i> (PPV)	$\frac{TP}{TP + FP}$	87.20%
Sensitivitas atau <i>recall</i> atau <i>true positive rate</i> (TPR)	$\frac{TP}{TP + FN}$	61.53%
<i>Specificity</i> atau <i>true negative rate</i> (TNR)	$\frac{TN}{TN + FP}$	90.97%
<i>Negative predictive values</i> (NPV)	$\frac{TN}{TN + FN}$	70.28%
<i>False positive rate</i> (FPR)	$\frac{FP}{FP + TN}$	9.03%
<i>False negative rate</i> (FNR)	$\frac{FN}{FN + TP}$	38.47%

4.4.3 Hasil Klasifikasi Algoritma SVM dengan *Class Attribute Per Fase*

Percobaan selanjutnya dengan membagi sampel data (*instances*) menjadi 5 bagian berdasarkan seri waktu. Fase pertama yaitu sampel data pada 3 menit pertama ketika bermain *video game*, fase ke-2 yaitu sampel data pada 3 menit dua ketika bermain *video game* dan seterusnya hingga fase ke lima. Hal ini dilakukan agar masing-masing kelas mempunyai jumlah *instances* yang seimbang. Sehingga masing-masing fase memiliki 144 *instances* pada kelas *baseline* dan kelas *playing*.

Proses klasifikasi pada percobaan ini, menggunakan *10-fold cross-validation* untuk membagi data *training* dan *testing*. Data *training* yang didapatkan dari metode *cross-validation* akan menjadi data *training* pada proses klasifikasi. Sedangkan data *testing*, akan menjadi data validasi pada proses klasifikasi.

Pada klasifikasi menggunakan algoritma SVM, metode yang akan digunakan untuk membantu proses klasifikasi data adalah metode *grid search*. Pada metode ini, pencarian parameter C dan γ untuk model yang optimal menggunakan prinsip *grid search*. Berdasarkan rekomendasi dari Hsu *et al* [63] pencarian parameter C dan γ di rentang nilai $C = 2^{-5}, 2^{-3}, \dots, 2^{15}$ dan $\gamma = 2^{-15}, 2^{-13}, \dots, 2^3$. Selanjutnya, matrix kernel *Radial Basis Function* (RBF) akan digunakan untuk membantu dalam mengatasi masalah *feature space*. Perulangan proses ini dilakukan hingga K ke-10. Kemudian, hasil yang didapatkan akan disimpan dan akan dilakukan kembali hingga split ke-10 dari split data yang yang dihasilkan oleh metode *10-fold cross-validation*. Algoritma SVM dengan mencoba membagi *class attribute* per fase adalah sebagai berikut:

Algoritma 1. SVM-Class Attribute Per Fase

Input : *Instances* Kelas *Baseline* ; *Instances* Kelas *Playing* ; Parameter C (*cost*) ; Parameter γ (*gamma*).

Output : Akurasi ; Presisi atau *positive predictive value* (PPV) ; Sensitivitas atau *recall* atau *true positive rate* (TPR) ; *Specificity* atau *true negative rate* (TNR) ; *Negative predictive values* (NPV) ; *False positive rate* (FPR) ; *False negative rate* (FNR) ; Tabel *Confusion Matrix*.

Begin :

1. Menggunakan perlakuan *10-fold cross validation* sebagai metode *split* data *training* dan *testing*.
2. Set kernel matriks RBF menggunakan persamaan:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\gamma\|x_i - x_j\|^2\right), \gamma > 0$$

3. Set parameter C dan γ di rentang nilai $C = 2^{-5}, 2^{-3}, \dots, 2^{15}$ dan $\gamma = 2^{-15}, 2^{-13}, \dots, 2^3$
4. Menggunakan klasifikasi SVM dari hasil *split* dengan persamaan:

$$\min_{w,b,\xi} \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i$$

subject to $y_i (w^T \varphi(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, l$

5. Menentukan hasil klasifikasi SVM terbaik yang dihasilkan pada setiap *split* atau lipatan di proses *k fold cross validation* dengan cara *averaging*.
6. Mengulangi dari langkah pertama untuk fase berikutnya.

Evaluasi performansi yang dihasilkan untuk dapat dijabarkan menjadi parameter evaluasi lain, secara komprehensif menggunakan tabel *confusion matrix*. Tabel 4.11 menampilkan setiap *instances* yang dikelompokan oleh *classifier* di setiap Fase yang di klasifikasi.

Tabel 4.11 *Confusion Matrix* dari SVM dengan *Class Attribute* Per Fase

FASE	AKTUAL	PREDIKSI	
		<i>Baseline</i>	<i>Playing</i>
Fase-1	<i>Baseline</i>	29	115
	<i>Playing</i>	20	124
Fase-2	<i>Baseline</i>	26	118
	<i>Playing</i>	17	127
Fase-3	<i>Baseline</i>	137	7
	<i>Playing</i>	107	37
Fase-4	<i>Baseline</i>	18	126
	<i>Playing</i>	12	132
Fase-5	<i>Baseline</i>	141	3
	<i>Playing</i>	102	42

Hasil yang didapatkan, ditampilkan pada Tabel 4.11 Pada fase ket-3 & fase ke-5, TPR memiliki skor yang lebih tinggi dari pada skor TNR, sedangkan

pada fase ke-1, 2 dan 4 skor TNR yang lebih tinggi dibandingkan dengan nilai TPR.

Apabila merujuk pada hasil visualisasi menggunakan parameter statistik *Difference Value Of Mean (DV)*. Dimana, pada visualisasi DV tersebut di fase ke-3 dan ke-5 terdapat subjek yang paling banyak memiliki perubahan yang signifikan dibanding fase lainnya, dimana perubahan signifikan yang dimaksud adalah perbedaan nilai *mean* yang di atas atau sama dengan 10bpm, sehingga perubahan yang tidak signifikan pada visualisasi parameter DV terdapat di fase ke-1, 2 dan 4, Maka perubahan signifikan yang terjadi pada hasil visualisasi tersebut merupakan prediksi karakteristik data dari kelas *baseline* secara dominan yang berhasil terdeteksi.

Selanjutnya, parameter PPV dan NPV mempunyai pola yang berbanding terbalik dengan TPR dan TNR. Pada fase ke-1, ke-2 dan ke-4, skor PPV mempunyai skor yang lebih tinggi dari fase lainnya, sedangkan pada parameter NPV mempunyai skor yang lebih tinggi pada fase ke-3 dan ke-5.

Kemudian, parameter FPR dan FNR mempunyai pola yang sama dengan TPR dan TNR. Yaitu pada fase ke-3 dan ke-5 parameter FPR memiliki skor yang tinggi dari pada fase lainnya. Sedangkan FNR pada fase ke-1, ke-2 dan ke-4.

Dengan kata lain dapat diinterpretasikan bahwa prediksi karakteristik data yang bernilai benar berdasarkan kelas yang telah diberi label, pada kelas positif (tidak marah) terjadi pada fase ke-3 dan ke-5, sedangkan prediksi karakteristik data yang bernilai benar berdasarkan kelas yang telah diberi label pada kelas negatif (kondisi marah) terjadi pada fase ke-1, ke-2, dan ke-4.

Selanjutnya, prediksi karakteristik data yang bernilai benar berdasarkan dari hasil prediksi klasifikasi, pada kelas positif (tidak marah) terjadi pada fase ke-1, ke-2, dan ke-4, sedangkan prediksi karakteristik data yang bernilai benar berdasarkan dari hasil prediksi klasifikasi pada kelas negatif (kondisi marah) terjadi pada fase ke-3 dan ke-5.

Sehingga kesimpulan yang dapat ditarik adalah, secara dominan emosi marah terjadi pada fase ke-1, ke-2, dan ke-4, dan tidak merasa emosi marah terjadi pada fase ke-3 dan ke-5. Hasil yang diperoleh disajikan pada Tabel 4.12 dibawah ini:

Tabel 4.12 Hasil Klasifikasi Algoritma SVM dengan *Class Attribute* Per Fase

Parameter penilaian	Hasil				
	Fase-1	Fase-2	Fase-3	Fase-4	Fase-5
Akurasi	53.13%	53.13%	60.42%	52.08%	63.54%
Presisi atau <i>positive predictive value</i> (PPV)	59.18%	60.47%	56.15%	60.00%	58.02%
Sensitivitas atau <i>recall</i> atau <i>true positive rate</i> (TPR)	20.14%	18.06%	95.14%	12.50%	97.92%
<i>Specificity</i> atau <i>true negative rate</i> (TNR)	86.11%	88.19%	25.69%	91.67%	29.17%
<i>Negative predictive values</i> (NPV)	51.88%	51.84%	84.09%	51.16%	93.33%
<i>False positive rate</i> (FPR)	13.89%	11.81%	74.31%	8.33%	70.83%
<i>False negative rate</i> (FNR)	79.86%	81.94%	4.86%	87.50%	2.08%

4.4.4 Analisa Hasil Klasifikasi Algoritma SVM

Parameter evaluasi untuk menentukan perbandingan performansi metode Bagging dan SMOTE adalah *weighted average* dari parameter PPV dengan NPV, TPR dengan TNR, dan FPR dengan FNR. Hasil yang didapatkan performansi usulan model klasifikasi menggunakan Bagging lebih baik dari pada SMOTE, meskipun pada metode SMOTE jumlah *instances* di kelas *baseline* telah ditingkatkan menjadi seimbang terhadap kelas *playing* dan skor prediksi PPV dan TPR menjadi meningkat dan FNR menjadi menurun.

Weighted average adalah rerata yang dihitung dengan melibatkan nilai bobot di setiap variabel. Hal ini dilakukan karena setiap parameter yang akan dihitung reratanya mempunyai bobot yang berbeda.

Bedasarkan dari hasil yang didapatkan pada pengujian metode Bagging dan SMOTE, klasifikasi pada kondisi kelas yang tidak seimbang (*imbalanced class*) memiliki performansi yang tidak buruk. Namun metode Bagging memiliki

performansi yang lebih maksimal jika dibandingkan dengan metode SMOTE. Berikut pada Tabel 4.13 hasil *weighted average* terhadap masing-masing parameter yang dihasilkan oleh metode Bagging dan metode SMOTE.

Tabel 4.13 Hasil *Weighted Average* pada Parameter Metode Bagging dan SMOTE

Parameter penilaian	Persamaan	Hasil	
		Metode Bagging	Metode SMOTE
<i>Weighted Averages</i> dari skor PPV dan NPV	$\frac{(PPV * (TP + FN)) + (NPV * (FP + TN))}{TP + FN + FP + TN}$	77.7%	78.7%
<i>Weighted Averages</i> dari skor TPR dan TNR	$\frac{(TPR * (TP + FN)) + (TNR * (FP + TN))}{TP + FN + FP + TN}$	83%	76.3%
<i>Weighted Averages</i> dari skor FPR dan FNR	$\frac{(FPR * (TP + FN)) + (FNR * (FP + TN))}{TP + FN + FP + TN}$	76.2%	23.8%

BAB 5

KESIMPULAN

5.1 Kesimpulan

Bedasarkan hasil dan analisa pengolahan data, maka dapat dipaparkan kesimpulan bahwa :

1. Pola yang ditemukan pada proses menggunakan ekstraksi parameter statistik *Difference Value Of Mean* (DV) dapat menunjukkan bahwa induksi emosi kemarahan yang terjadi pada setiap subjek berada pada fase ke-3, ke-4 dan ke-5. Sedangkan pada fase ke-3 dan ke-5, merupakan fase yang terdapat paling banyak subjek yang mengalami perubahan yang signifikan (lebih dari 10bpm), dengan masing-masing 5 subjek. Selanjutnya, parameter *Absolute Deviation Of Mean* (AD) dapat menunjukkan peningkatan di atas rata-rata dalam fase bermain terjadi dalam tiga menit ke-2 (3600 hingga 5400 poin), ke-3 (5400 hingga 7200 poin), dan tiga menit ke-5 (9000 hingga 10800 poin). Sementara itu, puncak kenaikan berada di tiga menit ke-3 (5400 hingga 7200 poin).
2. Prediksi yang dihasilkan oleh proses klasifikasi yang diusulkan menggunakan metode Bagging memiliki performansi yang lebih baik dari pada metode SMOTE. Metode Bagging menghasilkan skor akurasi sebesar 82.98%, *Weighted Averages* dari skor PPV dan NPV sebesar 77.7%, *Weighted Averages* dari skor TPR dan TNR sebesar 83%, dan *Weighted Averages* dari skor FPR dan FNR sebesar 76.2%
3. Sedangkan prediksi yang dihasilkan oleh proses klasifikasi menggunakan metode SMOTE mempunyai performansi dengan skor akurasi sebesar 76.25%, *Weighted Averages* dari skor PPV dan NPV sebesar 78.7%, *Weighted Averages* dari skor TPR dan TNR sebesar 76.3%, dan *Weighted Averages* dari skor FPR dan FNR sebesar 23.8%
4. Deteksi emosi marah per fase menggunakan klasifikasi SVM, parameter TPR menghasilkan performansi tertinggi pada fase ke-3 & fase ke-5, yaitu 95.14% dan 97.92%. Sedangkan parameter TNR menghasilkan

performansi tertinggi pada fase ke-1, ke-2 dan ke-2, yaitu 86.11%, 88.19% dan 91.67%.

5. Apabila merujuk pada hasil visualisasi menggunakan parameter statistik *Difference Value Of Mean (DV)*. Dimana, pada visualisasi DV tersebut di fase ke-3 dan ke-5 terdapat subjek yang paling banyak memiliki perubahan yang signifikan dibanding fase lainnya, dimana perubahan signifikan yang dimaksud adalah perbedaan nilai mean yang di atas atau sama dengan 10bpm, sehingga perubahan yang tidak signifikan pada visualisasi parameter DV terdapat di fase ke 1,2 dan 4, Maka perubahan signifikan yang terjadi pada hasil visualisasi tersebut merupakan prediksi karakteristik data dari kelas *baseline* secara dominan yang berhasil terdeteksi.
6. Analisa atau pengenalan pola emosi lebih akurat menggunakan metode klasifikasi dibanding dengan menggunakan visualisasi, karena pada klasifikasi terdapat proses learning untuk menentukan pengelompokan berdasarkan karakteristik data dari sampel yang sebelumnya telah diberi label.

5.2 Saran

Penelitian lebih lanjut dapat dilakukan dengan cara :

1. Penelitian ini memiliki data kelas yang tidak seimbang (*imbalanced class*), oleh sebab itu pengenalan pola yang dilakukan memiliki proses yang lebih lama. Penelitian selanjutnya yang dapat dilakukan, dengan cara menyeimbangkan panjang data berdasarkan waktu antara fase *pre-stimulation* dan fase *in-stimulation* maupun *post-stimulation*.
2. Melakukan validasi dan menambahkan risalah pada setiap langkah pengambilan data. Kemudian melakukan konfirmasi menggunakan kuesioner tentang bagaimana kondisi dan apa yang dirasakan oleh responden dengan cara pendekatan lainnya. Hal ini dilakukan dengan tujuan agar bisa menjadi referensi pendukung untuk menentukan label kelas dalam mendeteksi emosi. Karena ini adalah *study* yang

membutuhkan pengulangan dan penelitian lebih banyak varian agar bisa menjadi kesimpulan yang bisa menjadi informasi yang akurat.

3. Mengusulkan metode lain untuk penyelesaian permasalahan kelas yang tidak seimbang, dengan menggunakan solusi pendekatan data *undersampling* maupun gabungan dari *undersampling* dan *oversampling*, selain itu dapat juga dengan menggunakan solusi pendekatan algoritma *boosting* dan *stacking* maupun gabungan dari solusi pendekatan data dengan solusi pendekatan algoritma.

DAFTAR PUSTAKA

- [1] Kemenkes RI, "Tantangan e-Kesehatan di Indonesia" Jakarta, 2016.
- [2] B. Bonchev, "Adaptation in Affective Video Games: A Literature Review," *Cybernetics And Information Technologies*, vol. 16, no. 3, 2016.
- [3] R. Doumbouya, M. S Benlamine, A. Dufresne, and C. Frasson, "Game scenes evaluation and player's dominant emotion prediction," *Intelligent Tutoring Systems*, pp. 54–65, 2018.
- [4] S. D. Pressman and S. Cohen, "Does positive affect influence health?," *Psychol. Bull.*, vol. 131, no. 6, pp. 925–971, 2005.
- [5] S. D. Kreibig, "Autonomic nervous system activity in emotion: A review," *Biol. Psychol.*, vol. 84, no. 3, pp. 394–421, 2010.
- [6] A. D. Wibawa, M. H. Purnomo, A. Marzuki, and L. D. Rumpa, "Physiological Pattern of Human State Emotion Based on ECG and Pulse Sensor," *Journal of Theoretical and Applied Information Technology*, vol.93. No.1, 2016.
- [7] R.K. Dishman, Y. Nakamura, M. E. Garcia, R. W. Thompson, A. L. Dunn, and S. N. Blair, "Heart Rate Variability, Trait Anxiety, and Perceived Stress Among Physically Fit Men and Women," *International Journal Psychophysiol.*, vol.37 pp. 121–133, 2000.
- [8] B. H. Friedman, and J. F. Thayer, "Autonomic balance revisited: panic anxiety and heart rate variability," *J Psychosom Res.*, vol.44, pp. 131–151, 1998.
- [9] M. Pagani, F. Lombardi, S. Guzzetti, O. Rimoldi, R. Furlen, P. Pizzinelli, G. Sandrone, G. Malfatto, S. Dell, E. Piccaluga, M. Turiel, G. Baselli, S. Cerutti, and A. Malliani, "Power Spectral Analysis Of Heart Rate and Arterial Pressure Variabilities as A Marker Of Sympathovagal Interaction In Man and Conscious Dog," *Circ Res.*, vol. 59, pp. 171–192, 1986.
- [10] G. Chanel, C. Rebetez, M. Betrancourt, and T. Pun, "Boredom, Engagement And Anxiety As Indicators For Adaptation To Difficulty In Games," *Proceedings of the 12th International Conference on Entertainment and Media in the Ubiquitous Era*, pp. 13–17, 2008.
- [11] M. Ambinder, "Biofeedback In Gameplay: How Valve Measures Physiology To Enhance Gaming Experience," *Presented at the Game Developer's Conference*, 2011.
- [12] R.L. Mandryk and M.S. Atkins, "A Fuzzy Physiological Approach for Continuously Modeling Emotion during Interaction with Play Environments," *Int. l J. Human-Computer Studies*, vol. 65, pp. 329-347, 2007.

- [13] E. Zamani, M. Chashmi, and N. Hedayati, "Effect of Addiction to Computer Games on Physical and Mental Health of Female and Male Students of Guidance School in City of Isfahan," *Addiction and Health*, Vol 1, pp.98-104, 2009.
- [14] World Health Organization, "Gaming disorder," *Health statistics and information systems*, 2018. [Online]. Available: <https://icd.who.int/browse11/11m/en#/http://id.who.int/icd/entity/1448597234>.
- [15] Kemenkes RI, "Bermain Game Online: Mengisi Waktu Luang, Bersenang-senang atau Ketergantungan," Jakarta, 2018.
- [16] Kemenkes RI, "Ketahui Tiga Jenis Terapi untuk Atasi Adiksi," Jakarta, 2018.
- [17] L. Zamani, *Kecanduan Game Online, Puluhan Pelajar Diobati di Rumah Sakit Jiwa Solo*, Jakarta: Kompas, 2019.
- [18] S. Rida, *Rumah Sakit Jiwa Tangani 81 Pasien Kecanduan Game Online Rata-rata Anak Dibawah Umur*, Jambi : Tribun News, 2020.
- [19] D.E Kurniawan, "Pengaruh Intensitas Bermain Game Online Terhadap Perilaku Prokrastinasi Akademik pada Mahasiswa Bimbingan dan Konseling Universitas PGRI Yogyakarta". *Jurnal Konseling GUSJIGANG*, vol. 3, no. 1, 2017.
- [20] N. Badariah, *Hubungan Perilaku Agresi dan Kecanduan Game Online pada Mahasiswa*, Thesis: Fakultas Kedokteran Universitas Syiah Kuala, Bandar Lampung, 2018.
- [21] Rokom, *Inilah Dampak Kecanduan Game Online*, Kemkes: Sahabat Negeriku, 2018.
- [22] H. Masya dan D. A. Candra, *Faktor-Faktor Yang Mempengaruhi Perilaku Gangguan Kecanduan Game Online Pada Peserta Didik Kelas X Di Madrasah Aliyah Al Furqon Prabumulih Tahun Pelajaran 2015/2016*, Thesis: Fakultas Tarbiyah dan Keguruan, IAIN Raden Intan, 2016.
- [23] M. A. Suplig, "Pengaruh Kecanduan Game Online Siswa SMA Kelas X Terhadap Kecerdasan Sosial Sekolah Kristen Swasta Di Makassar," *Jurnal Jaffray*, Vol. 15, No. 2, Oktober 2017.
- [24] K. Harahap and I. Beydha, *Game Online dan Prestasi Belajar (Studi Korelasional Pengaruh Game Online Terhadap Prestasi Belajar Siswa Kelas VIII SMP Nurul Hasanah Kelurahan Padang Bulan Medan)*, Thesis: Universitas Sumatera Utara, Medan, 2013.
- [25] Y. R. Dwi dan J. T. Purnomo, "Hubungan Kecanduan Game Online Terhadap Penyesuaian Sosial Pada Remaja," *Jurnal Humaniora Yayasan Bina Darma*, vol. IV, no. 1, hal. 027-044, Januari-Juni 2017.

- [26] S. Mihara and S. Higuchi, "Cross-sectional and longitudinal epidemiological studies of Internet gaming disorder: A systematic review of the literature," *Psychiatry and Clinical Neurosciences.*, no. 71, pp. 425–444, 2017.
- [27] L. Wartberg, L. Kriston, M. Kramer, A. Schwedler, T.M Lincoln, R. Kammerl, "Internet gaming disorder in early adolescence: Associations with parental and adolescent mental health." *European Psychiatry.*, vol. 43, pp.14-18 2017.
- [28] S. T. T'ng, K. H. Ho, D. E. Sim, C. H. Yu, P. Y. Wong, "The mediating effect of Internet gaming disorder's symptoms on loneliness and aggression among undergraduate students and working adults in Malaysia." *PsyCh Journal.*, pp.320, 2019.
- [29] H. Song, F. Fang, F. K. Arnberg, D. Mataix-Cols, L. F. de la Cruz, C. Almqvist, K. Fall, P. Lichtenstein, G. Thorgeirsson, U. A. Valdimarsdóttir, "Stress Related Disorders and Risk of Cardiovascular Disease: Population Based, Sibling Controlled Cohort Study." *BMJ.*, vol.365, pp. 11255, 2019.
- [30] C. Vlachakis, k. Dragoumani, S. Raftopoulou, M. Mantaïou, L. Papageorgiou, S. C. Tsaniras, V. Megalooikonomou, D. Vlachakis, "Human Emotions on the Onset of Cardiovascular and Small Vessel Related Diseases." *in vivo.*, vol.32, pp. 859-870, 2018.
- [31] Z. K. Nekouei, H. Neshatdoost, A. Yousefy, M. Sadeghi, G. Manshaee, "Psychological Factors and Coronary Heart Disease," *ARYA Atheroscler.*, vol.9, pp.102-111, 2013.
- [32] A. D. Wibawa, M. H. Purnomo, A. Marzuki, L. D. Rumpa, " Physiological Pattern of Human State Emotion Based on ECG and Pulse Sensor," *Journal of Theoretical and Applied Information Technology*, Vol.93. No.1, 2016.
- [33] R.K. Dishman, Y. Nakamura, M. E. Garcia, R. W. Thompson, A. L. Dunn, S. N. Blair, "Heart Rate Variability, Trait Anxiety, and Perceived Stress Among Physically Fit Men and Women," *International Journal Psychophysiol*, vol.37 pp. 121–133, 2020.
- [34] B. H. Friedman, J. F. Thayer, "Autonomic balance revisited: panic anxiety and heart rate variability," *J Psychosom Res*, vol.44, pp. 131–151, 1998.
- [35] M. Pagani, F. Lombardi, S. Guzzetti, O. Rimoldi, R. Furlen, P. Pizzinelli, G. Sandrone, G. Malfatto, S. Dell, E. Piccaluga, M. Turiel, G. Baselli, S. Cerutti, A. Malliani, "Power Spectral Analysis Of Heart Rate and Arterial Pressure Variabilities as A Marker Of Sympathovagal Interaction In Man and Conscious Dog," *Circ Res*, vol. 59, pp. 171–192, 1986.

- [36] L. Hakim, E. S. Pane, A. D. Wibawa, M. H. Purnomo, "Emotion Recognition in Elderly Based on SpO2 and Pulse Rate Signals Using Support Vector Machine." *IEEE ICIS.*, pp.474-479, 2018
- [37] A. Marzuki, L. D. Rumpa, A. D. Wibawa, and M. H. Purnomo, "Classification of Human State Emotion from Physiological Signal Pattern Using Pulse Sensor Based On Learning Vector Quantization," *Int. Semin. Intell. Technol. Its Appl.*, pp. 129–133, 2016.
- [38] Z. A. Khoirunnisaa, *Seleksi Kanal Pada Electroencephalograph (EEG) Menggunakan Metode Correlation Feature Selection (CFS) untuk Identifikasi Cybersickness*. Masters Thesis: Institut Teknologi Sepuluh November, 2018.
- [39] Newzoo, 2020 Global Games Market Report April Update, 2020. [Online]. Available: <https://newzoo.com/key-numbers/>
- [40] T. H. Apperley, "Genre and game studies: Toward a critical approach to video game genres," *Sage Publications*, vol. 37, no. 1, pp. 6-23, 2006.
- [41] The Entertainment Software Association (ESA), Essential Facts About the Video Game Industry in July 2020, [Online]. Available: https://www.thesa.com/wp-content/uploads/2020/07/Final-Edited-2020-ESA_Essential_facts.pdf
- [42] S. Sundari, *Kesehatan Mental Dalam Kehidupan*, Jakarta : Rineka Cipta, 2005.
- [43] K. R. Scherer, R. Banse, and H. G. Wallbott, "Emotion Inferences from Vocal Expression Correlate Across Languages and Cultures," *Journal of Cross-culturalpsychology*, vol.32, pp.76–92, 2002
- [44] A. Andersen, Peter, Guerrero, K. Laura, *Handbook of Communication and Emotion: Research, Theory, Applications and Contexts*, California: Academia Press, 1997.
- [45] J. Duffy, *Managing Anger and Aggression: Practical Guidance for Schools*, South Eastern Education and Library Board: Psychology/ Behavior Support Section, 2012.
- [46] R. L. Cautin, J. C. Overholser, and P. Goetz, "Assesment of Mode of Anger Expression In Adolescent Psychiatric Inpatients," *Proquest Sociology*, vol.36(141), pp. 163-170, 2001.
- [47] S. D. Kreibig, "Autonomic nervous system activity in emotion: A review," *Biol. Psychol.*, vol. 84, no. 3, pp. 394–421, 2010.
- [48] L. Sherwood, *Human Physiology: From Cells to System, Seventh, vol. 28, no. 3*. Canada: Yolanda Cassio, 2010.
- [49] R. Bott, Guyton and Hall, *Textbook of Medical Physiology 13ed, no. 1*. 2014.

- [50] Cooking hacks, “e-Health Sensor Platform V2.0 for Arduino and Raspberry Pi [Biometric / Medical Applications],” 2016. [Online]. Available: <https://www.cooking-hacks.com/documentation/tutorials/ehealthbiometric-sensor-platform-arduino-raspberry-pi-medical/>.
- [51] S. Walsh and E. King, *Pulse Diagnosis: A Clinical Guide*, Sidney: Churchill Livingstone Elsevier, 2007.
- [52] A. Jubran, “Pulse Oximetry,” *Crit. Care*, vol. 19, no. 1, pp. 1–7, 2015
- [53] K. M. Sensing, “How to Read SpO2.” Konica Minolta, 2006.
- [54] W. Wen, G. Liu, N. Cheng, J. Wei, P. Shangguan, and W. Huang, “Emotion Recognition Based on Multi-Variant Correlation of Physiological Signals,” *IEEE Trans. Affect. Comput.*, vol. 5, no. 2, pp. 126–139, 2014.
- [55] R. Gopinathannair and B. Olshansky, "Management of tachycardia," F1000Prime Reports, 2015.
- [56] C. Bonnaire, D. Baptista, “Internet gaming disorder in male and female young adults: The role of alexithymia, depression, anxiety and gaming type,” *Psychiatry Research*, vol. 272, pp. 521-530, 2019.
- [57] M. Kantardzic, *Data Mining: Concepts, Models, Methods, and Algorithms*, 2011.
- [58] C. Burges, *A Tutorial on Support Vector Machine for Pattern Recognition*. 955- 974, 1998.
- [59] Chen, P.-H., C.-J. L., & Scholkopf, B, *A Tutorial on ν -Support Vector Machines*, 2005.
- [60] B. Scholkopf and A. J. Smola, *Learning with Kernel: Support Vector Machines, Regularization, Optimization and Beyond*, Cambridge: MIT Press, 2002.
- [61] J. Mercer, “Functions of Positive and Negative Type and Their Connection with The Theory of Integral Equations,” *Philosophical Transactions of the Royal Society of London*, pp. 3-23, 1909
- [62] B. Santosa, *Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis*. Yogyakarta: Graha Ilmu, 2007.
- [63] C.W. Hsu, C. C. Chang, and C. J. Lin, *A Practical Guide to Support Vector Classification*, Thesis: National Taiwan University, 2003.
- [64] L. Breiman, *Bagging Prediktor*, New York: Department of statistics University of California, 1994.
- [65] P. Buhlmann, and B. Yu, “Analyzing Bagging,” *The Annals of Statistics*, vol. 30, no. 4, pp. 927-961, 2002.
- [66] A. Nugroho, B. Arief, and H. Dwi, “Support Vector Machines: Teori dan Aplikasinya Dalam Bioinformatika,” *Proceeding of Indonesian Scientific Meeting in Central Japan*, 2003.

- [67] F. D. Nizarul, *Deteksi Botnet Menggunakan Naive Bayes Classifier dengan SMOTE dan Metode BFS*. Masters Thesis: Institut Teknologi Sepuluh November, 2017.
- [68] W. Li, Z. Liu, "A method of SVM with Normalization in Intrusion Detection." *Procedia Environmental Sciences*, Vol. 11, Part A, Pages 256-262, 2011.
- [69] T. Saito, M. Rehmsmeier, "The Precision-Recall Plot Is More Informative than the ROC Plot When Evaluating Binary Classifiers on Imbalanced Datasets." *PLOS ONE*, Pages 1-22, 2015.
- [70] Chawla, N. W. & Bowyer, K. W., "SMOTE: Synthetic Minority Oversampling Technique. *Journal of Artificial Intelligence Research*", Volume 16, p. 321–357, 2002.
- [71] T. Hastie, R. Tibshirani, J. Friedman, "The Elements of Statistical Learning: Data Mining, Inference, and Prediction." *Springer Series In Statistics*, 2009.
- [72] Sulastri, S., Rohadi, S., Sunardi, B., Vita, A. N., & Prayogo, A. S, "Aplikasi Metode Mean Dan Median Absolute Deviation Pada Data Elektromagnet Sebagai Prekursor Gempa Bumi Di Pelabuhan Ratu." *Prosiding Seminar Nasional Fisika*, 2017.
- [73] C. D. Sutton, "Classification and Regression Trees, Bagging, and Boosting," *Handbook of Statistics*, vol. 24, no. 04, pp. 303–329, 2005.

LAMPIRAN

Lampiran I : Kuesioner

Pernyataan dan Kuesioner Pengambilan Data Kesehatan (Persentase Oksigen Dalam Darah dan Denyut Nadi)

1. Nama:
2. Usia:
3. Jenis Kelamin:
4. Berat Badan dan Tinggi Badan:
5. Dalam 1 hari berapa jam anda tidur?
6. Dalam 1 hari berapa kali anda makan (Makan Nasi dan Lauk)?
7. Apakah anda perokok, jika iya, berapa batang rokok yang anda hisap dalam 1 hari?
8. Dalam 1 hari apakah anda ada minum kopi, jika iya, berapa gelas kopi yang anda minum dalam 1 hari?
9. Bagaimana kondisi kesehatan anda saat ini, apakah anda batuk, flu, atau pusing atau lainnya:
10. Dalam 1 minggu berapa kali anda olahraga:
11. Dalam 1 minggu berapa kali anda bermain Game Dota:
12. Berapa jam dalam 1 kali anda bermain Game Dota:
13. Sudah berapa lama anda bermain Game Dota?
14. Selama anda bermain Dota, Emosi apa yang paling dominan yang anda rasakan?

Senang

Marah

Takut

Lainnya _____

Dalam Skala 0 sampai 9, Seberapa besar emosi yang anda rasakan?

Senang

0	1	2	3	4	5	6	7	8	9
Tidak				Sedang					Sangat

Marah

0	1	2	3	4	5	6	7	8	9
Tidak				Sedang					Sangat

Takut

0	1	2	3	4	5	6	7	8	9
Tidak				Sedang					Sangat

Lainnya _____

0	1	2	3	4	5	6	7	8	9
Tidak				Sedang					Sangat

15. Ketika anda bermain Game Dota tadi, hasil apa yang anda dapatkan:

Menang dengan mudah

Menang dengan Sulit

Dikalahkan dengan mudah

Kalah, namun nyaris menang

Dengan ini saya menyatakan bersedia untuk tidak menuntut apapun atas data yang saya serahkan ke **Pihak Pertama (Peneliti)**. Untuk diketahui Pihak Pertama akan menggunakan data yang diperoleh dari **Pihak Kedua (Responden)** dengan seperlunya dalam lingkup penelitian saja. Menyangkut perihal privasi data Pihak Kedua akan dijaga oleh Pihak Pertama.

Surabaya, September 2017

Pihak Pertama (Peneliti)

Pihak Kedua (Responden)

BIOGRAFI PENULIS



Dhanu Kurnia Utama, Magister Teknik Elektro Bidang Keahlian Telematika, Fakultas Teknologi Elektro, Institut Teknologi Sepuluh Nopember Surabaya lahir di Jambak, 23 April 1989 merupakan anak pertama dari Bapak Sumbar A Restianto Bin Rasyidjan dan Ibu Jumiati Binti Taslam. Penulis mempunyai satu adik kandung bernama Arga Dwi Priswara serta Istri yang bernama Norma Yunidya Wardhani dan Anak Falah Hamizan Dhandya Taqwa. Selama hidupnya, penulis menempuh pendidikan di SDN 27 Bandarejo (1994-2000), SMPN 1 Pasaman (2000-2003), SMAN 1 Bukittinggi (2003-2006). Penulis melanjutkan pendidikan Diploma III di Politeknik Negeri Padang Jurusan Teknik Komputer, Lalu Melanjutkan Diploma IV di Politeknik Elektronika Negeri Surabaya Jurusan Teknik Informatika. Penulis dapat dihubungi dengan kontak berikut:

Email : dhanu.kurnia.utama@gmail.com