

BAB I

PENDAHULUAN

1.1 Latar Belakang

Relaksasi psikofisiologis merupakan kondisi yang dapat diidentifikasi melalui perubahan parameter fisiologis, seperti detak jantung (*heart rate*/HR) dan tekanan darah (*blood pressure*/BP). Parameter tersebut banyak digunakan sebagai indikator objektif dalam menilai respons tubuh terhadap stres maupun relaksasi karena berhubungan langsung dengan aktivitas sistem saraf otonom [1]. Ketika individu berada dalam kondisi relaksasi, aktivitas sistem saraf simpatis cenderung menurun, yang ditandai dengan penurunan detak jantung dan tekanan darah [2].

Aromaterapi lavender (*Lavandula angustifolia*) merupakan salah satu intervensi non-farmakologis yang banyak digunakan untuk menurunkan kecemasan dan meningkatkan relaksasi. Beberapa penelitian menunjukkan bahwa inhalasi aromaterapi lavender memiliki efek signifikan terhadap penurunan tekanan darah dan detak jantung[3]. Penelitian sebelumnya melaporkan bahwa pemberian aromaterapi lavender pada pasien hipertensi mampu menurunkan tekanan darah sistolik dan diastolik secara signifikan [4]. Penelitian lain juga menunjukkan bahwa inhalasi lavender dapat menurunkan detak jantung serta menekan aktivitas sistem saraf simpatis, sehingga mendukung kondisi relaksasi fisiologis [5]. Oleh karena itu, pengukuran detak jantung dan tekanan darah menjadi variabel yang relevan dalam mengevaluasi tingkat relaksasi setelah pemberian aromaterapi [6].

Seiring berkembangnya teknologi, analisis respons fisiologis tidak hanya dilakukan secara deskriptif, tetapi juga melalui pendekatan komputasional berbasis machine learning. Metode terkini yang banyak diterapkan dalam analisis data kesehatan adalah teknik *clustering* dan *classification* untuk mengidentifikasi pola serta mengelompokkan kondisi fisiologis pasien[7]. Salah satu algoritma clustering yang umum digunakan adalah *K-Means Clustering*, yang terbukti efektif dalam mengelompokkan data kesehatan berdasarkan indikator fisiologis seperti tekanan darah dan detak jantung [8]. Dalam penelitian sebelumnya, *K-Means* digunakan untuk mengelompokkan pasien berdasarkan tingkat risiko kesehatan dengan memanfaatkan fitur tekanan darah sistolik dan diastolik [9]. Namun demikian, efektivitas *K-Means* sangat dipengaruhi oleh pemilihan pusat kluster awal (*initial centroids*) yang dapat mempengaruhi hasil pengelompokan [10].

Tantangan utama dalam penelitian sistem informasi kesehatan adalah sulitnya mendapatkan dataset klinis riil pasca-intervensi yang terbuka untuk publik. Pengumpulan data langsung pada subjek manusia memiliki kendala etika, biaya, dan waktu. Salah satu solusi untuk mengatasi kelangkaan data ini adalah melalui pendekatan simulasi komputasional berbasis parameter literatur medis (*Synthetic*

Data Generation), yang memungkinkan pengujian algoritma dilakukan secara valid tanpa melibatkan subjek fisik secara langsung [11], [12].

Selain clustering, klasifikasi tingkat kondisi fisiologis juga sering melibatkan pendekatan hybrid atau kombinasi algoritma. Beberapa penelitian menunjukkan bahwa penggabungan *K-Means* sebagai tahap pra-klasifikasi dengan algoritma klasifikasi seperti *Naive Bayes* mampu meningkatkan akurasi prediksi [13]. Algoritma *Naive Bayes* sendiri telah banyak digunakan dalam bidang kesehatan, misalnya pada prediksi risiko diabetes dan analisis data medis berbasis probabilistik karena kesederhanaan dan efisiensinya [14]. Dengan mengintegrasikan *K-Means* untuk pengelompokan awal data fisiologis dan *Naive Bayes* untuk tahap klasifikasi, sistem diharapkan mampu mengklasifikasikan tingkat relaksasi secara lebih akurat berdasarkan respons detak jantung dan tekanan darah [15].

1.2 Rumusan Masalah

1. Bagaimana menganalisa simulasi data fisiologis sintetis yang realistis dengan tetap mempertahankan korelasi multivariat antara HR, HRV, SBP, dan DBP berdasarkan parameter statistik dari literatur terkait efek lavender?
2. Bagaimana penerapan algoritma *K-Means* dalam melakukan segmentasi pola perubahan fisiologis dan menginterpretasikan kluster yang terbentuk sebagai tingkat relaksasi rendah, sedang, dan tinggi?
3. Bagaimana kinerja algoritma *Gaussian Naive Bayes* dalam mengklasifikasikan tingkat relaksasi berdasarkan data hasil simulasi, serta bagaimana ketahanan model terhadap variasi data sintetis?

1.3 Batasan Masalah

1. Sumber Data
 - a. Data kondisi awal (*baseline*) menggunakan dataset sekunder *Psychological State Dataset* (2024) [16].
 - b. Data pasca-intervensi diperoleh melalui simulasi statistik menggunakan distribusi normal multivariat dengan parameter yang merujuk pada literatur.
 - c. Penelitian tidak melibatkan pengumpulan data klinis langsung maupun eksperimen pada subjek manusia.
2. Variabel Penelitian
 - a. Parameter fisiologis yang dianalisis terbatas pada HR, HRV, SBP, dan DBP.
 - b. Faktor psikologis, lingkungan, dan demografis tidak dimasukkan dalam analisis.
3. Metodologi Komputasional

- a. Algoritma yang digunakan dibatasi pada *K-Means* untuk clustering dan *Gaussian Naive Bayes* untuk klasifikasi.
 - b. Perbandingan performa dilakukan terbatas pada algoritma *Machine Learning* konvensional (*Random Forest, SVM, Decision Tree*) untuk mengukur efisiensi relatif model usulan.
 - c. Simulasi mengasumsikan distribusi normal multivariat dengan hubungan korelasi linier.
4. Lingkup Validasi
 - a. Validasi penelitian bersifat statistik dan metodologis, bukan klinis.
 - b. Hasil penelitian hanya berlaku pada konteks data sintetis dengan karakteristik tertentu.

1.4 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengembangkan model simulasi data fisiologis sintetis yang merepresentasikan perubahan parameter fisiologis pasca-efek lavender dengan mempertahankan korelasi multivariat yang realistis.
2. Menerapkan pendekatan *hybrid K-Means* dan *Naive Bayes* untuk segmentasi dan klasifikasi tingkat relaksasi berdasarkan data fisiologis hasil simulasi.
3. Mengevaluasi performa algoritma *Gaussian Naive Bayes* menggunakan metrik akurasi, presisi, *recall*, *F1-score*, serta validasi silang, sekaligus menganalisis kontribusi masing-masing fitur fisiologis.
4. Menyusun kerangka kerja simulasi dan evaluasi algoritma yang dapat direplikasi pada penelitian lain dengan keterbatasan data klinis.

1.5 Manfaat Penelitian

1. Manfaat Teoritis
 - a. Memberikan kontribusi metodologis dalam bidang *Health Informatics* terkait pemanfaatan simulasi data untuk mengatasi keterbatasan data klinis.
 - b. Menambah referensi penelitian mengenai penerapan metode *hybrid unsupervised* dan *supervised learning* pada data fisiologis sintetis.
 - c. Menyediakan kerangka validasi statistik simulasi data yang dapat digunakan pada penelitian intervensi non-farmakologis lainnya.
2. Manfaat Praktis
 - a. Menyediakan *framework* komputasional bagi peneliti sistem informasi kesehatan untuk evaluasi algoritma tanpa ketergantungan pada data klinis.
 - b. Memberikan gambaran penerapan algoritma *K-Means* dan *Naive Bayes* dalam klasifikasi tingkat relaksasi berbasis parameter fisiologis.
 - c. Menjadi bahan pembelajaran akademik terkait simulasi data, *clustering*, dan klasifikasi pada konteks kesehatan.