

BAB 2

LANDASAN TEORI

2.1 Stres

2.1.1 Definisi Stres

Stres dapat didefinisikan sebagai hasil dari interaksi antara individu dengan lingkungannya, yang dipersepsikan sebagai beban atau tekanan yang melampaui kemampuan dirinya untuk mengatasinya [12]. Definisi ini menekankan bahwa stres bukan semata-mata akibat dari stresor (sumber tekanan) itu sendiri, melainkan juga bergantung pada bagaimana individu menilai dan merespons stresor tersebut.

2.1.2 Gejala dan Faktor Penyebab Stres

Stres pada mahasiswa dapat dikenali melalui empat aspek gejala, yaitu gejala fisik, gejala emosional, gejala kognitif, dan gejala interpersonal. Gejala fisik dapat berupa sakit kepala, gangguan tidur, atau kelelahan; gejala emosional dapat berupa mudah marah atau cemas; gejala kognitif dapat berupa kesulitan berkonsentrasi; sedangkan gejala interpersonal dapat berupa penarikan diri dari lingkungan sosial. Faktor penyebab stres dapat dibedakan menjadi faktor internal, seperti kemampuan dan kecerdasan individu, serta faktor eksternal, seperti tuntutan kampus, keluarga, dan lingkungan sekitar [1].

2.1.3 Tahapan Stres (*General Adaptation Syndrome*)

Selye [6] melalui model General Adaptation Syndrome (GAS) menjelaskan bahwa respons stres bersifat progresif, yaitu berkembang melalui fase alarm sebagai reaksi awal, fase resistensi ketika tubuh berusaha beradaptasi, hingga fase kelelahan (*exhaustion*) ketika energi adaptasi telah habis. Implikasi penting dari model ini adalah bahwa apabila stresor tidak segera ditangani, kondisi individu dapat berkembang menjadi lebih kronis seperti *burnout*. Model tersebut menjadi dasar konseptual mengapa penyaringan (*screening*) tingkat stres perlu dilakukan sedini mungkin, agar intervensi dapat diberikan sebelum kondisi memburuk.

Perlu ditegaskan bahwa penelitian ini tidak mengukur fase GAS secara langsung. Penanda fisiologis fase *alarm* seperti peningkatan detak jantung

dan pelepasan hormon kortisol berada di luar jangkauan instrumen kuesioner *self-report* yang digunakan. Sejalan dengan definisi operasional yang telah dijelaskan pada Subbab 1.1, istilah “stres awal” pada penelitian ini merujuk pada waktu dilakukannya deteksi, bukan pada fase stres menurut GAS. Penelitian ini mengklasifikasikan tingkat keparahan stres (*Mild, Moderate, Severe, dan Very Severe*) berdasarkan gejala yang dilaporkan sendiri oleh mahasiswa, sebagaimana diuraikan pada Subbab 2.1.2.

2.2 Stres Akademik pada Mahasiswa dalam Mengerjakan Skripsi

Stres akademik adalah bentuk stres yang muncul akibat tuntutan-tuntutan yang berkaitan dengan kegiatan akademik, termasuk dalam proses penyusunan skripsi. Mahasiswa tingkat akhir yang sedang mengerjakan skripsi rentan mengalami stres akademik karena dihadapkan pada berbagai tuntutan, seperti ketidakmampuan mengelola waktu, ekspektasi tinggi terhadap diri sendiri, serta ketidakpastian dalam proses bimbingan dengan dosen pembimbing [3]. Selain faktor tersebut, kemampuan mencari bantuan secara akademik (*academic help-seeking*) dan resiliensi akademik juga turut memengaruhi tingkat stres yang dialami mahasiswa dalam menyelesaikan skripsi [2]. Tingginya prevalensi stres pada mahasiswa Indonesia, yang tercatat lebih tinggi dibandingkan rata-rata Asia maupun dunia [1], serta data nasional yang menunjukkan satu dari tiga remaja Indonesia mengalami masalah kesehatan mental [5], menegaskan bahwa isu stres akademik pada mahasiswa, termasuk dalam konteks pengerjaan skripsi, merupakan permasalahan nyata yang perlu mendapat perhatian khusus.

2.3 Kuesioner sebagai Instrumen Pengumpulan Data

Kuesioner merupakan instrumen penelitian yang digunakan untuk mengukur suatu fenomena, baik fenomena alam maupun sosial, dengan cara memberikan serangkaian pertanyaan atau pernyataan tertulis kepada responden untuk dijawab [13]. Salah satu skala pengukuran yang umum digunakan dalam kuesioner psikologis maupun sosial adalah skala Likert, yaitu skala yang digunakan untuk mengukur sikap, pendapat, dan persepsi seseorang atau sekelompok orang terhadap suatu fenomena, dengan menjabarkan variabel yang diukur menjadi indikator-indikator yang kemudian disusun menjadi item-item pertanyaan [13].

Meskipun skala Likert umum digunakan dalam kuesioner psikologis,

instrumen pada penelitian ini disusun dalam format pilihan ganda (multiple choice) dengan opsi jawaban bersifat nominal. Format ini dipilih atas pertimbangan bahwa butir-butir hasil wawancara dengan psikolog berorientasi pada jenis pengalaman yang dominan dirasakan mahasiswa (misalnya jenis kendala utama, jenis strategi coping) dan bukan pada derajat persetujuan terhadap suatu pernyataan, sehingga lebih tepat direpresentasikan sebagai kategori daripada sebagai skala bertingkat.

Namun demikian, pilihan format ini membawa dua konsekuensi metodologis yang perlu diakui secara terbuka. Pertama, opsi jawaban nominal tidak memiliki urutan keparahan intrinsik, sehingga tidak dapat langsung dijumlahkan menjadi skor total sebagaimana lazim dilakukan pada skala Likert. Kedua, uji validitas item-total maupun uji reliabilitas konsistensi internal (Cronbach's alpha) tidak dapat diterapkan pada data nominal. Kedua keterbatasan ini menjadi catatan penting bagi penelitian ini dan dibahas lebih lanjut pada Subbab 4.8.1 serta Bab 5.

Agar data yang dihasilkan oleh kuesioner dapat dipertanggungjawabkan secara ilmiah, instrumen kuesioner umumnya perlu memenuhi dua syarat utama, yaitu validitas dan reliabilitas. Validitas menunjukkan derajat ketepatan antara data yang sesungguhnya terjadi pada objek penelitian dengan data yang berhasil dikumpulkan oleh instrumen, sedangkan reliabilitas menunjukkan derajat konsistensi dan stabilitas suatu instrumen apabila digunakan secara berulang pada objek yang sama [13]. Dalam penelitian ini, kuesioner yang digunakan disusun berdasarkan butir-butir pertanyaan yang diperoleh dari psikolog di lingkungan Universitas Multimedia Nusantara (UMN), sebagai alternatif dari instrumen baku umum seperti DASS-21 maupun *Student-Life Stress Inventory* yang banyak digunakan pada penelitian sejenis. Mengingat format jawaban yang bersifat nominal, pemenuhan validitas pada penelitian ini ditempuh melalui validitas isi (*content validity*), yaitu dengan menyusun butir pertanyaan langsung bersama psikolog yang berkompeten di bidangnya, dan bukan melalui uji validitas maupun reliabilitas statistik yang mensyaratkan data berskala. Keterbatasan pada aspek pengujian statistik ini diakui secara terbuka dan menjadi salah satu saran penyempurnaan pada Bab 5.

2.4 Machine Learning

Machine learning adalah cabang ilmu kecerdasan buatan yang memungkinkan sistem untuk belajar dari data dan membuat prediksi atau keputusan tanpa diprogram secara eksplisit untuk setiap aturan [14]. Salah

satu kategori utama *machine learning* adalah *supervised learning*, yaitu proses pembelajaran yang menggunakan data yang sudah diberi label (*labeled data*) agar model dapat belajar memetakan hubungan antara fitur input dan label keluaran. Klasifikasi merupakan salah satu tugas dalam *supervised learning* yang bertujuan untuk mengelompokkan data ke dalam kategori atau kelas tertentu berdasarkan fitur-fiturnya, misalnya mengklasifikasikan tingkat stres seseorang ke dalam kategori normal, ringan, atau sedang.

2.4.1 Decision Tree

Decision Tree atau pohon keputusan adalah model prediktif berbentuk struktur pohon, yang terdiri atas simpul akar (*root node*), simpul internal yang merepresentasikan suatu atribut atau fitur, cabang yang merepresentasikan hasil pengujian atribut tersebut, serta simpul daun (*leaf node*) yang merepresentasikan label kelas hasil akhir [15]. Proses pembentukan *decision tree* dilakukan dengan memilih atribut terbaik untuk memisahkan data menjadi subset yang lebih homogen pada setiap simpul, proses ini disebut *splitting*. Pemilihan atribut terbaik umumnya didasarkan pada kriteria *Information Gain* atau *Gini Index*. *Gini Index* mengukur tingkat ketidakmurnian (*impurity*) suatu himpunan data, yang dapat dirumuskan sebagai berikut:

$$Gini(D) = 1 - \sum_{i=1}^n p_i^2 \quad (2.1)$$

dengan p_i adalah proporsi data pada kelas i dalam himpunan data D , dan n adalah jumlah kelas. Semakin rendah nilai *Gini Index* suatu atribut, semakin baik atribut tersebut dalam memisahkan data ke dalam kelas-kelas yang homogen [15]. Kelemahan utama *decision tree* tunggal adalah kecenderungannya untuk mengalami *overfitting*, yaitu kondisi ketika model terlalu menyesuaikan diri dengan data latih sehingga performanya menurun pada data baru.

Secara algoritmik, proses pembentukan *decision tree* dapat dirangkum sebagai prosedur rekursif sebagaimana disajikan pada Algoritma 2.1.

```

1 FUNCTION BuildTree(D) :
2     IF seluruh data pada D berada pada kelas yang sama
3     OR D kosong
4     OR kedalaman maksimum tercapai THEN
5         RETURN simpul daun dengan label kelas mayoritas pada D
6

```

```

7   terbaik <- None
8   giniTerbaik <- infinity
9
10  FOR setiap atribut A pada D:
11      hitung Gini(D, A)
12      IF Gini(D, A) < giniTerbaik THEN
13          terbaik <- A
14          giniTerbaik <- Gini(D, A)
15
16  bagi D menjadi D1, D2, ..., Dk berdasarkan nilai atribut
    terbaik
17
18  FOR i = 1 TO k:
19      Ti <- BuildTree(Di)
20
21  RETURN simpul internal dengan atribut terbaik
    dan cabang T1, T2, ..., Tk
22

```

Kode 2.1: Pseudocode pembentukan Decision Tree (CART)

2.4.2 Random Forest

Random Forest adalah algoritma *ensemble learning* yang diperkenalkan oleh Breiman [11], yang dibangun dengan menggabungkan keluaran dari sejumlah besar *decision tree* untuk menghasilkan satu prediksi akhir yang lebih akurat dan stabil. Setiap *decision tree* dalam *Random Forest* dibangun dari subset data yang diambil secara acak dengan pengembalian (*bootstrap sampling*) dari data latih, serta hanya menggunakan subset fitur yang dipilih secara acak pada setiap proses *splitting*. Teknik ini disebut *bootstrap aggregating* atau *bagging*, yang bertujuan untuk mengurangi variansi model dan mencegah *overfitting* yang umum terjadi pada *decision tree* tunggal [11].

Secara umum, cara kerja algoritma *Random Forest* dalam tugas klasifikasi dapat dijabarkan sebagai berikut:

1. Mengambil N sampel data secara acak dengan pengembalian dari data latih.
2. Membangun satu *decision tree* untuk setiap sampel yang telah diambil, dengan pemilihan fitur secara acak pada setiap simpul.
3. Melakukan proses ini secara berulang hingga terbentuk sejumlah T pohon keputusan.

4. Untuk data uji baru, setiap pohon dalam hutan memberikan satu prediksi kelas.
5. Hasil akhir klasifikasi ditentukan melalui mekanisme *voting* mayoritas (modus) dari seluruh prediksi pohon yang terbentuk.

Selain menghasilkan prediksi, *Random Forest* juga dapat memberikan ukuran *feature importance*, yaitu kontribusi relatif setiap fitur terhadap penurunan *impurity* (Persamaan 2.1) secara rata-rata di seluruh pohon dalam hutan. Informasi ini bermanfaat untuk mengidentifikasi indikator kuesioner yang paling berpengaruh terhadap hasil klasifikasi stres awal mahasiswa.

2.5 Evaluasi Model Klasifikasi

Performa model klasifikasi perlu dievaluasi menggunakan metrik tertentu agar dapat diketahui sejauh mana model mampu melakukan prediksi dengan benar. Salah satu alat yang umum digunakan untuk mengevaluasi performa model klasifikasi adalah *confusion matrix*, yaitu tabel yang menggambarkan jumlah prediksi benar dan salah yang dihasilkan oleh model, yang terdiri atas *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) [16]. Berdasarkan *confusion matrix* tersebut, beberapa metrik evaluasi dapat dihitung, yaitu:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \quad (2.2)$$

$$Presisi = \frac{TP}{TP + FP} \quad (2.3)$$

$$Recall = \frac{TP}{TP + FN} \quad (2.4)$$

$$F1-score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \quad (2.5)$$

Akurasi mengukur proporsi prediksi yang benar dari seluruh prediksi yang dilakukan oleh model. Presisi mengukur proporsi prediksi positif yang benar-benar tepat, sedangkan *recall* mengukur proporsi data positif sebenarnya yang berhasil dikenali oleh model. F1-score digunakan sebagai metrik gabungan antara presisi dan *recall*, yang sangat bermanfaat ketika jumlah data pada setiap kelas tidak seimbang [16].

2.6 Penelitian Terdahulu

Beberapa penelitian terdahulu yang relevan dengan topik deteksi stres menggunakan *machine learning* dirangkum pada Tabel 2.1.

Tabel 2.1. Ringkasan penelitian terdahulu yang relevan

Peneliti (Tahun)	Metode	Objek/Data	Hasil
Oktaviani, Rosmawarni, & Muslim [10]	Random Forest & SMOTE	Kuesioner 14 pertanyaan pada mahasiswa tingkat akhir yang mengerjakan skripsi	SMOTE Random Forest unggul pada data tidak seimbang
Larasati, Dewi, Farhansyah, Bachtiar, & Pradana [9]	<i>Decision Tree</i> & Random Forest	Data kondisi tidur untuk deteksi tingkat stres	Random Forest akurasi 1,00; <i>Decision Tree</i> 0,99
Septiani [17]	Random Forest	Kuesioner diagnosa tingkat depresi mahasiswa saat pandemi Covid-19	Random Forest mampu mengklasifikasikan tingkat depresi dengan baik
Zahrah & Muljono [8]	Perceptron, GBTC, Naive Bayes, Logistic Regression, AdaBoost	Deteksi stres pelajar berbasis data terstruktur	Perceptron terbaik dengan akurasi 97,27%
Susanti [18]	Naive Bayes (hybrid)	Klasifikasi tingkat stres mahasiswa Teknik Informatika saat perkuliahan	Model hybrid Naive Bayes mampu mengklasifikasikan tingkat stres mahasiswa
Rahayu, Fitria, Septhya, Rahmadden, & Efrizoni [7]	<i>Decision Tree</i> , Random Forest, Naive Bayes, KNN	Data teks Twitter untuk deteksi depresi dan kecemasan	Model klasifikasi teks mampu mengenali pola depresi/kecemasan

Berdasarkan tabel tersebut, dapat dilihat bahwa algoritma Random Forest secara konsisten menunjukkan performa yang baik dalam berbagai tugas klasifikasi stres maupun depresi, baik yang menggunakan data kuesioner, data fisiologis (kondisi tidur), maupun data teks. Penelitian yang paling mendekati topik penelitian ini adalah penelitian Oktaviani, Rosmawarni, dan Muslim [10], yang juga menggunakan data kuesioner pada mahasiswa yang sedang mengerjakan skripsi. Namun, penelitian tersebut menggunakan instrumen kuesioner umum dan belum secara spesifik dirancang untuk keperluan penyaringan dini pada populasi mahasiswa di institusi tertentu, serta belum menyoroti populasi mahasiswa pada institusi tertentu dengan instrumen yang disusun bersama psikolog di lingkungan kampus tersebut. Kesenjangan inilah yang menjadi posisi kontribusi dari penelitian ini, yaitu membangun model klasifikasi berbasis Random Forest untuk mendeteksi stres awal pada mahasiswa UMN, menggunakan kuesioner yang diperoleh langsung dari psikolog UMN.

