

PENERAPAN DECISION TREE DAN RANDOM FOREST DALAM DETEKSI TINGKAT STRES MANUSIA BERDASARKAN KONDISI TIDUR

Sza Sza Amulya Larasati¹, Elok Nuraida Kusuma Dewi², Brahma Hanif Farhansyah³, Fitra
Abdurachman Bachtiar^{4*}, Fajar Pradana⁵

^{1,2,3,4,5}Universitas Brawijaya, Malang

Email: ¹szaszaal@student.ub.ac.id, ²eloknuraidakd@student.ub.ac.id, ³tifhanifhahn@student.ub.ac.id,
⁴fitra.bachtiar@ub.ac.id, ⁵fajar.p@ub.ac.id

*Penulis Korespondensi

(Naskah masuk: 06 November 2023, diterima untuk diterbitkan: 30 Oktober 2024)

Abstrak

Masalah kesehatan mental menjadi isu global yang sangat umum terjadi, termasuk perubahan suasana hati, perbedaan kepribadian, ketidakmampuan mengatasi masalah, serta mengisolasi diri dari keramaian. Berdasarkan data dari World Health Organization (WHO), gangguan kecemasan dan stres menjadi gangguan mental yang paling sering terjadi dari 970 juta kasus yang dilaporkan sepanjang tahun 2019. Stres telah banyak dikaitkan dengan tidur. Penelitian ini akan mengungkap hubungan kondisi tidur pada manusia dengan tingkat stres yang sedang diderita dengan 5 tingkatan: normal, stres ringan, stres sedang, stres tinggi, stres sangat tinggi. Data yang digunakan merupakan data kontinyu dengan 8 fitur: 'sr' (*snoring rate*), 'rr' (*respiration rate*), 't' (*body temperature*), 'lm' (*limb movement*), 'bo' (*blood oxygen*), 'rem' (*rapid eye movement*), 'sh' (*sleeping hours*), dan 'hr' (*heart rate*). Setiap fitur memiliki rentang nilai yang tidak sama, sehingga dilakukan normalisasi untuk menyeragamkan rentang tersebut. *Hyperparameter tuning* dilakukan dengan teknik k-fold cross validation dan model dirancang dengan algoritma klasifikasi Decision Tree serta Random Forest. Hasilnya, 5 fitur: tingkat mendengkur, laju respirasi, pergerakan anggota tubuh termasuk bola mata, serta detak jantung saat tidur berbanding lurus dengan tingkat stres. Semakin tinggi nilai kelima fitur tersebut mengindikasikan tingkat stres yang lebih tinggi. Sedangkan dengan 3 fitur lainnya: suhu tubuh, kadar oksigen, dan waktu tidur memberikan hasil sebaliknya. Dengan kata lain, ketiga nilai tersebut berbanding terbalik dengan tingkat stres yang diderita. Model Decision Tree memiliki akurasi 0,99 dan Random Forest memiliki akurasi 1,0. Hasil penelitian ini diharapkan dapat memberikan *insight* bagi peneliti lain pada bidang yang sama dan dapat menjadi acuan dalam mendeteksi stres yang sedang diderita.

Kata kunci: *Decision Tree, Deteksi, Random Forest, Stres, Tidur*

THE APPLICATION OF DECISION TREE AND RANDOM FOREST IN DETECTING HUMAN STRESS LEVELS BASED ON SLEEP CONDITIONS

Abstract

Stress is often associated with sleep. This research aims to uncover the relationship between human sleep conditions and the level of stress experienced, categorized into five levels: not stressed, very mildly stressed, mildly stressed, highly stressed, and very highly stressed. The data used consists of continuous data with eight features: 'snoring rate' (snoring rate), 'respiration rate' (respiration rate), 'body temperature' (body temperature), 'limb movement' (limb movement), 'blood oxygen' (blood oxygen), 'rapid eye movement' (rapid eye movement), 'sleep hours' (sleep hours), and 'heart rate' (heart rate). Each feature has a different value range, so normalization is performed to standardize these ranges. Hyperparameter tuning is done using k-fold cross-validation, and the model is designed using the Decision Tree and Random Forest classification algorithms. The results show that five features: snoring rate, respiration rate, limb movement including eye movement, and heart rate during sleep are directly proportional to the level of stress. Higher values for these five features indicate higher levels of stress. On the other hand, the other three features: body temperature, blood oxygen level, and sleep hours yield the opposite results. In other words, the values of these three features are inversely proportional to the level of stress experienced. The Decision Tree model has an accuracy of 0.99, and the Random Forest model has an accuracy of 1.0. The results of this research are expected to provide insights for other researchers in the same field and can serve as a reference for detecting ongoing stress

Keywords: *Decision Tree, Detection, Random Forest, Sleep, Stress*

1. PENDAHULUAN

Isu kesehatan mental merupakan permasalahan global yang terjadi di seluruh dunia. Gangguan kesehatan mental melibatkan berbagai aspek, seperti perubahan suasana hati, perubahan kepribadian, kesulitan dalam mengatasi masalah, serta gemar mengisolasi diri dari keramaian (Ceja, et. al., 2018). Berdasarkan data dari WHO, gangguan kecemasan dan stres menjadi gangguan mental yang paling sering terjadi dari 970 juta kasus yang dilaporkan sepanjang tahun 2019. Stres dapat dikatakan sebagai respon fisiologis saat seseorang menyesuaikan diri dengan tekanan yang mendorong adanya perubahan psikologis dan emosional dalam menanggapi tekanan tersebut (Ahuja, & Banga, 2019). Gejala stres seperti merasa tidak berharga dan putus asa merupakan faktor risiko bunuh diri dan 55% diantaranya memiliki pikiran untuk bunuh diri (Frijanto, 2022). WHO juga menyatakan bahwa 800 ribu orang di seluruh dunia meninggal karena bunuh diri sepanjang tahun 2019 dan kerap dikaitkan dengan gangguan kejiwaan terutama stres. Fakta tersebut mengindikasikan bahwa stres menjadi masalah yang krusial dan harus diatasi dengan segera.

Tingkat stres seseorang dapat dideteksi dengan cara yang beragam. Beberapa penelitian sebelumnya telah menyoroti hubungan erat antara stres dan kondisi tidur. Tidur dapat didefinisikan sebagai keadaan ketika aktivitas sistem saraf cenderung berkurang dalam kondisi mata tertutup dan otot-otot pada keadaan santai. Berdasarkan penelitian (Veeramachaneni, et. al., 2019), durasi tidur yang konsisten dapat menjadi indikasi tingkat stres yang lebih rendah. Temuan serupa juga digagas oleh (Kartika, 2021) dan (Shanbhog & Medikonda, 2023) yang menyebutkan bahwa kualitas tidur memiliki korelasi positif dengan tingkat stres. Fakta-fakta ini menegaskan bahwa kondisi seseorang saat tidur dapat menjadi tolak ukur stres yang sedang diderita.

Analisis secara komprehensif pada kebiasaan tidur dapat memberikan informasi mengenai tingkat stres yang sedang diderita. Stres yang terdeteksi pada tahap awal dapat memberikan peluang kesembuhan yang lebih tinggi. Hasil deteksi ini dapat dijadikan sebagai acuan dalam penentuan penanganan lebih lanjut pada pengidap stres. Proses analisis kini banyak diimplementasikan dengan bantuan teknologi dan memberikan hasil yang cukup menjanjikan. Salah satu teknologi yang kini banyak dikembangkan adalah *machine learning*, yang memungkinkan mesin memperoleh *insight* data dan menentukan keputusan tanpa perintah eksplisit (Jordan & Mitchell, 2015). *Machine learning* dapat diimplementasikan dengan algoritma beragam pada dua pendekatan: *supervised learning* dan *unsupervised learning*.

Data "*Human Stress Detection in and through Sleep*" dipilih untuk melatih model deteksi stres

(Rachakonda, 2021). Data ini merupakan data kontinyu dengan 8 fitur yang merepresentasikan kondisi manusia saat tidur dengan 1 label kelas yang mengelompokkannya dalam 5 kategori stres. Beberapa fitur memiliki korelasi positif dengan label kelas, namun fitur lainnya memiliki korelasi negatif dengan label kelas. Korelasi ini membuktikan bahwa data bersifat non-linear, sehingga model akan dikembangkan dengan algoritma Decision Tree yang menggunakan serangkaian aturan untuk membuat prediksi berdasarkan fitur-fitur data dengan interpretabilitas tinggi. Sayangnya, model Decision Tree yang dibangun terlalu kompleks berpotensi tinggi untuk mengalami *overfitting* (Roshanski, et. al., 2023). Algoritma Random Forest juga diusulkan dalam pembuatan model deteksi stres yang kemudian akan dibandingkan kinerjanya dengan model Decision Tree. Kedua model ini sama-sama berbasis pohon keputusan, namun Random Forest akan membuat beberapa pohon keputusan menggunakan data bootstrap dan secara acak memilih subset variabel di setiap pohon keputusan (Yovita, 2021). Keduanya, baik Decision Tree dan Random Forest dapat mengidentifikasi garis non-linear pemisah antar kelas yang kompleks dalam data kontinyu yang digunakan.

Penelitian ini dikembangkan untuk mendeteksi tingkat stres manusia berdasarkan kondisi tidur. Model Decision Tree dikembangkan untuk mengenali tingkat stres tersebut pada data non-linear yang digunakan. Model Random Forest yang diusulkan akan dibandingkan dengan kinerja model Decision Tree untuk menangani kesalahan klasifikasi data. Model dengan kinerja paling baik akan menjadi luaran akhir penelitian ini yang diharapkan dapat mengenali stres berdasarkan pola kondisi tidurnya.

2. STUDI LITERATUR

2.1 Stres

Berdasarkan data dari WHO, gangguan kecemasan dan stres menjadi gangguan mental yang paling sering terjadi dari 970 juta kasus yang dilaporkan sepanjang tahun 2019. Stres dapat didefinisikan sebagai keadaan ketegangan emosional yang disebabkan oleh keadaan yang bersifat menuntut dan tidak dapat dihindari (Rachakonda, 2021). Stres dapat dipicu oleh berbagai faktor, seperti tekanan pekerjaan, konflik dalam hubungan *interpersonal*, masalah finansial, atau tragedi traumatis. Stres merupakan kondisi kompleks yang melibatkan aspek emosional, kognitif, dan biologis. Stres dapat dikategorikan menjadi beberapa tingkatan, stres dengan tingkat lebih dapat menyebabkan gangguan pada sistem tubuh manusia serta merangsang respon pertahanan sistem saraf pusat. Respon ini tergantung pada reaksi fisiologis

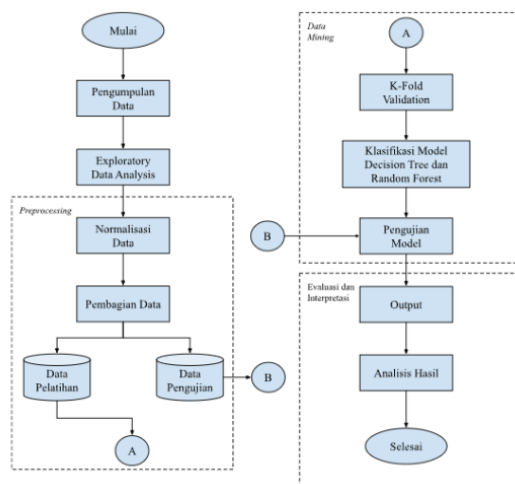
individu yang mencakup respon neuro-endokrin dan perilaku, serta melibatkan perubahan dalam aktivitas dan fungsi sumbu hipotalamus-hipofisis-adrenal (HPA).

Pengaruh stres pada setiap individu berbeda. Beberapa orang memiliki mekanisme penanganan stres yang efektif dan mampu mengatasi tuntutan yang ada, sementara orang lain mungkin mengalami kesulitan menangani tantangan yang mereka hadapi dan berpotensi mengalami stres dengan tingkat yang lebih tinggi. Pada tingkat yang moderat, stres dapat memberikan motivasi untuk menghadapi tantangan dan mencapai tujuan. Namun, stres yang berkepanjangan dapat memberikan dampak negatif pada kesehatan fisik dan mental.

2.2 Tidur

Tidur merupakan aktivitas yang sangat penting untuk menjaga keseimbangan tubuh manusia. Tidur dapat didefinisikan sebagai keadaan ketika aktivitas sistem saraf cenderung berkurang, kondisi mata tertutup, dan otot-otot dalam keadaan santai. Penelitian sebelumnya (Veeramachaneni, et. al., 2019), (Kartika, 2021), (Shanbhog & Medikonda, 2023) berhasil menunjukkan bahwa tidur memiliki korelasi positif dengan tingkat stres manusia. Kualitas tidur yang lebih baik dapat mengindikasikan tingkat stres yang rendah. Tidur yang terganggu telah dikaitkan dengan peningkatan kasus penyakit, sindrom kelelahan berat, insomnia psikofisiologis yang berlangsung lama, sistem kekebalan tubuh yang lemah, PTSD, dan risiko kecelakaan kerja yang lebih tinggi. (Rachakonda, et. al. 2020).

3. METODE PENELITIAN



Gambar 1. Metodologi penelitian

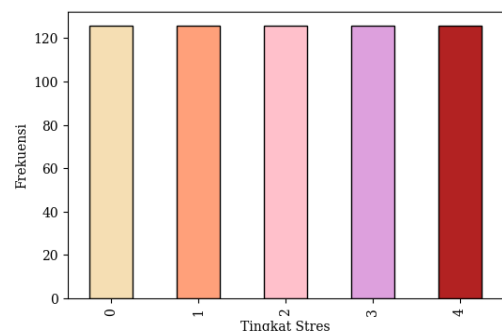
Gambar 1 menunjukkan metodologi penelitian yang akan dilakukan dalam penelitian ini. Penelitian diawali dengan proses pengumpulan data dilanjutkan dengan tahapan Exploratory Data Analysis untuk menemukan wawasan awal pada data. Preprocessing data dilakukan untuk memastikan konsistensi data,

melibatkan proses transformasi yakni normalisasi data serta pembagian data menjadi data pelatihan dan pengujian. Data hasil preprocessing ini akan digunakan dalam proses pelatihan dan pengujian model pada tahapan data mining. Hasil pengujian akan dievaluasi dengan beberapa metrik untuk mengetahui kinerja model dan membantu proses analisis data.

3.1 Pengumpulan Data

Penelitian ini menggunakan data sekunder yang didapatkan melalui penyedia *open data* yakni pada website *Kaggle*. Dataset “*Human Stress Detection in and through Sleep*” (Rachakonda, 2021) dipilih sebagai data untuk melatih dan menguji model Decision Tree dan Random Forest sebagai pendeteksi tingkat stres. Data ini berisi kumpulan nilai mengenai kondisi tidur seseorang.

Dataset “*Human Stress Detection in and through Sleep*” merupakan data numerik kontinyu sebanyak 630 baris yang mengukur delapan keadaan, yakni : ‘sr’ (*snoring rate*), ‘rr’ (*respiration rate*), ‘t’ (*body temperature*), ‘lm’ (*limb movement*), ‘bo’ (*blood oxygen*), ‘rem’ (*rapid eye movement*), ‘sh’ (*sleeping hours*), dan ‘hr’ (*heart rate*). Data ini memberikan informasi mengenai tingkat stres yang didefinisikan sebagai ‘sl’ (*stress level*) dengan lima kategori, yakni: 0 (tidak stres), 1 (stres sangat rendah), 2 (stres rendah), 3 (stres tinggi), 4 (stres sangat tinggi). Label kelas ‘sl’ tersebar rata di setiap kategori, ditunjukkan pada Gambar 2.



Gambar 2. Distribusi data pada label “sl”

3.2 Exploratory Data Analysis

EDA bertujuan untuk mengungkap wawasan awal dari data yang telah dikumpulkan. Analisis statistik dan berbagai teknik visualisasi data dilakukan pada tahapan ini untuk mengidentifikasi pola, anomali, dan tren yang mungkin tersembunyi dalam data tersebut. EDA juga dapat memberikan pemahaman yang mendalam terhadap struktur data pada data yang digunakan. Hal ini akan menjadikan EDA sebagai landasan yang sangat krusial dalam menentukan algoritma klasifikasi pada tahap selanjutnya.

3.3 Preprocessing

Tahap pemrosesan awal data bertujuan untuk mempersiapkan data pada proses pengolahan data di tahap selanjutnya. Tahap pemrosesan awal data melibatkan proses transformasi data dengan melakukan normalisasi karena hasil analisis statistik pada EDA memberikan informasi bahwa setiap fitur pada data memiliki rentang nilai yang tidak sama. Ketidaksamaan rentang ini dapat membuat fitur dengan rentang nilai yang lebih tinggi mendominasi perhitungan jarak karena memiliki bobot yang lebih tinggi sehingga akan dilakukan normalisasi data untuk menyamakan rentang data di setiap fitur. Teknik normalisasi yang dipilih adalah normalisasi Min-Max.

$$v' = \frac{v - \min_F}{\max_F - \min_F} (\text{new_max}_F - \text{new_min}_F) + \text{new_min}_F \quad (1)$$

Persamaan 1. Normalisasi Min-Max

Persamaan 1 menjelaskan proses normalisasi Min-Max dengan mengubah nilai data asli (v) dalam rentang $[0,1]$. Deviasi nilai v dari rentang data asli (\min_F hingga \max_F) disesuaikan dalam rentang baru dengan proporsi deviasi tersebut. Hasil normalisasi (v') dapat mempermudah proses perbandingan dan analisis data serta memastikan bahwa data-data dengan skala yang berbeda dapat digunakan secara seragam dalam berbagai konteks pengolahan data dan pembelajaran mesin.

Tahap *preprocessing* juga melibatkan proses pembagian data menjadi data pelatihan (80%) dan data pengujian (20%). Dari 630 data, 504 baris data akan digunakan dalam proses pelatihan dan 126 data lainnya digunakan sebagai data pengujian. Teknik sampling yang digunakan adalah Random Sampling, dengan memilih data pelatihan dan pengujian secara acak.

3.4 Data Mining

Algoritma yang digunakan dalam melatih model deteksi adalah Decision Tree dan akan dibandingkan kinerjanya dengan Random Forest. Beberapa parameter akan digunakan pada model Decision Tree dan untuk menentukan parameter yang paling baik akan dilakukan *hyperparameter tuning* dengan teknik Cross Validation dengan K-Fold.

Tahap *data mining* juga melibatkan proses pelatihan untuk menghasilkan model klasifikasi berdasarkan 504 data pelatihan. Model yang dilatih adalah model Decision Tree dan Random Forest. Pada tahap ini juga dilakukan pengujian model berdasarkan 126 data pengujian yang telah didefinisikan sebelumnya. Proses pengujian ini juga dilakukan pada kedua model yang nantinya kinerja keduanya akan dibandingkan

Decision Tree

Decision Tree merupakan salah satu algoritma pembelajaran mesin yang digunakan untuk membangun model prediktif atau klasifikasi berdasarkan pengambilan keputusan yang berstruktur

seperti pohon. Pohon keputusan terdiri dari simpul (*node*) yang mewakili keputusan, cabang (*branch*) yang mewakili hasil dari keputusan tersebut, dan daun (*leaf*) yang mewakili hasil klasifikasi akhir. Algoritma ini termasuk dalam *Supervised Learning* dengan data yang terus dibagi sesuai parameter input dan dilakukan secara rekursif hingga mencapai daun (Cruz, et. al. 2020). Decision Tree dapat diimplementasikan pada data kontinyu dan struktur pohon yang dihasilkan dapat mengklasifikasikan data lebih baik terutama pada data *non-linear separable* (Fratello & Tagliaferri, 2019).

Random Forest

Random Forest merupakan algoritma yang digunakan untuk mempertimbangkan banyak Decision Tree sehingga dapat didefinisikan sebagai *ensemble* dari algoritma Decision Tree. Random Forest akan mencoba menemukan fitur terbaik secara acak di antara semua fitur (Ahuja, R., & Banga, A., 2019).

Random Forest akan membuat beberapa pohon keputusan menggunakan data bootstrap dan secara acak memilih subset variabel di setiap pohon keputusan (Yovita, 2021). Metode Random Forest memungkinkan setiap pohon keputusan dibangun dengan sampel acak dari data pelatihan. Selain itu, pada setiap simpul dalam pohon keputusan hanya dipilih sebagian atribut secara acak saja untuk memecah data. Variasi ini membuat Random Forest dapat mengurangi *overfitting* dan meningkatkan generalisasi dalam prediksi atau klasifikasi. Hasil akhir dari Random Forest didapatkan dengan menggabungkan prediksi atau klasifikasi dari setiap pohon keputusan dalam *ensemble*. Dalam kasus klasifikasi, prediksi akhir dapat ditentukan berdasarkan mayoritas suara dari pohon-pohon keputusan, sedangkan dalam kasus regresi diambil berdasarkan rata-rata prediksi dari pohon-pohon tersebut.

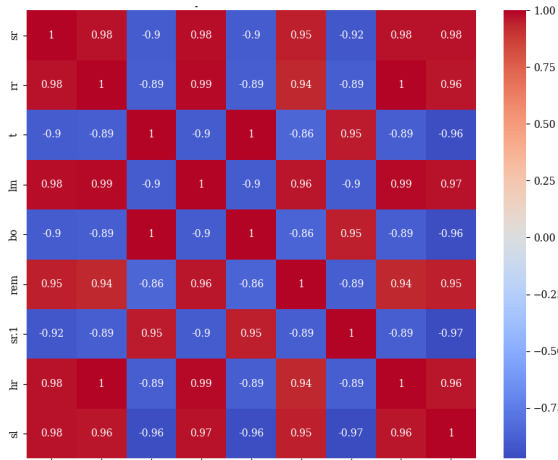
3.5 Interpretasi dan Evaluasi Model

Interpretasi dan evaluasi model merupakan tahap akhir yang ditujukan untuk mengevaluasi kinerja model. Keseluruhan model dievaluasi menggunakan *classification report* yang memberikan informasi mengenai presisi, *recall*, *f1-score*, *support*, dan akurasi model. Selain itu, *confusion matrix* juga digunakan sebagai evaluator model yang memberikan informasi mengenai seberapa banyak kelas yang diklasifikasikan dengan benar dan salah pada set data pengujian.

4. HASIL DAN PEMBAHASAN

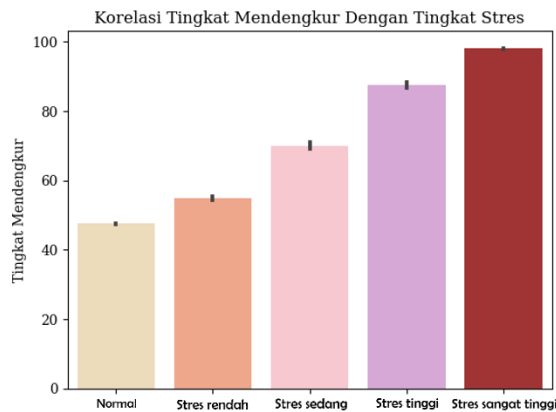
Proses pelatihan melibatkan 504 baris data pelatihan dengan nilai yang berbeda. Perhitungan korelasi antar fitur memberikan hasil bahwa lima fitur memiliki korelasi positif dengan label kelas, namun tiga fitur lainnya memiliki korelasi negatif. Korelasi

ini ditunjukkan pada Gambar 3 dalam bentuk visualisasi *heatmap*.



Gambar 3. *Heatmap* korelasi antar fitur

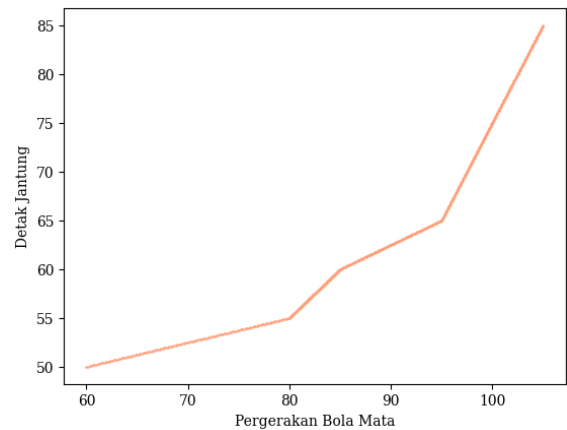
Berdasarkan Gambar 4 didapatkan informasi bahwa tingkat stres berbanding lurus dengan tingkat mendengkur, laju respirasi, pergerakan anggota tubuh termasuk bola mata, serta detak jantung saat tidur. Peningkatan nilai tersebut mengacu pada tingkat stres yang lebih tinggi. Lain halnya dengan suhu tubuh, kadar oksigen, dan waktu tidur yang berbanding terbalik, peningkatan ketiganya mengindikasikan tingkat stres yang lebih rendah.



Gambar 4. Korelasi tingkat mendengkur dengan tingkat stres

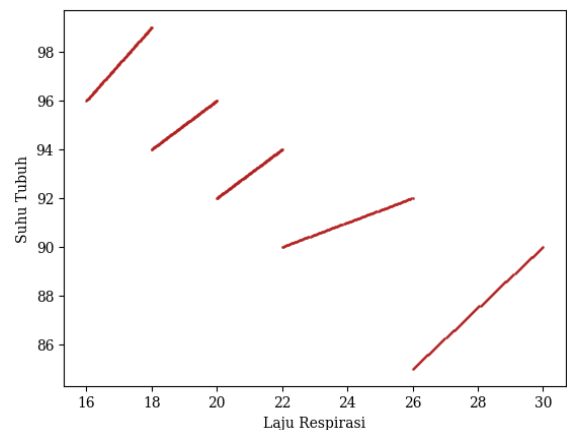
Delapan fitur yang digunakan merupakan data numerik, sedangkan tingkat stres merupakan data kategorikal. Apabila fitur ini direpresentasikan secara visual dapat mengungkap tren dari pengidap stres. Gambar 4 menunjukkan korelasi antara tingkat mendengkur seseorang ketika tidur dan tingkatan stres yang sedang dialami. Hubungan yang dihasilkan memvalidasi fakta bahwa tingkat mendengkur ini berbanding lurus dengan stres. Seseorang yang tidak terindikasi stres memiliki tingkat mendengkur yang rendah dan sebaliknya, seseorang dengan tingkat stres tinggi cenderung memiliki nilai tingkat mendengkur yang juga tinggi. *Insight* serupa juga dihasilkan pada fitur laju respirasi, pergerakan anggota tubuh

termasuk bola mata, serta detak jantung saat tidur jika dikaitkan dengan tingkat stres.



Gambar 5. Korelasi detak jantung dengan pergerakan bola mata saat tidur

Fitur yang berbanding lurus dengan tingkat stres juga dibandingkan satu sama lain. Detak jantung dan pergerakan bola mata memiliki korelasi yang sama dengan tingkat stres yakni berkorelasi positif, ketika kedua fitur ini dibandingkan memberikan informasi bahwa keduanya juga berbanding lurus. Korelasi ini ditunjukkan pada Gambar 5. Fakta ini juga divalidasi dengan *heatmap* pada Gambar 3 yang menyatakan bahwa detak jantung (*hr*) dan pergerakan bola mata (*rem*) memiliki koefisien korelasi sebesar 0,94. Fitur lainnya termasuk tingkat mendengkur, laju respirasi, pergerakan anggota tubuh ketika dibandingkan satu sama lain juga akan memberikan korelasi positif, artinya peningkatan salah satu fitur membuat peningkatan pada fitur lainnya.



Gambar 6. Korelasi suhu tubuh dan laju respirasi saat tidur

Lain halnya ketika fitur yang berkorelasi positif dibandingkan dengan fitur yang berkorelasi negatif dengan tingkat stres, hubungan yang dihasilkan dari keduanya adalah berbanding terbalik. Sebagai contoh, suhu tubuh memiliki korelasi negatif dengan tingkat stres dan laju respirasi memiliki korelasi positif dengan stres level. Korelasi yang dihasilkan dari suhu tubuh dan laju respirasi nyatanya berbanding terbalik satu sama lain, dimana

peningkatan suhu tubuh membuat laju respirasi lebih rendah. Hal ini diilustrasikan pada Gambar 6.

Korelasi-korelasi yang dihasilkan membuktikan bahwa data bersifat non-linear. *Decision Tree* dan *Random Forest* dipilih dalam implementasi model karena dapat menangani hubungan non-linear dan interaksi antara fitur dengan baik. Keduanya dapat mengidentifikasi garis non-linear pemisah antar kelas yang kompleks dalam data “*Human Stress Detection in and through Sleep*” dan cocok untuk data kontinyu. Pemilihan *Decision Tree* dan *Random Forest* juga dipilih dengan pertimbangan daya komputasi yang dibutuhkan cukup rendah.

Tabel 1. Hasil *cross validation* pada *hyperparameter tuning model*

| Parameter | Rata – rata Cross Validation | Standar Deviasi Cross Validation |
|------------------|---------------------------------|-------------------------------------|
| DT gini-none | 0,982139 | 0,013141 |
| DT gini-log2 | 0,984119 | 0,010113 |
| DT entropy-none | 0,986119 | 0,010089 |
| DT entropy-log 2 | 0,992079 | 0,009701 |
| Random Forest | 0,990099 | 0,012524 |

Model Decision Tree dibuat dengan parameter *criterion* dan *max_features* melalui proses *hyperparameter tuning* pada teknik Cross Validation. *Criterion* digunakan dalam penentuan pemisahan cabang, yakni dengan “gini” dan “entropi”, sedangkan *max_features* berperan dalam pengontrolan banyak fitur yang diambil dalam memilih pemisah pada setiap node, yakni “none” dan “log2”. Tabel 1 menunjukkan bahwa rata-rata Cross Validation Decision Tree paling baik ada pada parameter entropy-log2 dengan nilai 0,992079. Standar deviasi dari parameter tersebut memiliki nilai yang cukup kecil, yakni 0,009701. Nilai tersebut memberikan arti bahwa distribusi data pada model Decision Tree entropy-log2 tersebar dekat nilai rata-rata. Berdasarkan hal tersebut, model ini dinilai cukup stabil karena dapat membantu mengurangi risiko *overfitting*, terutama jika model cenderung mempelajari perubahan-perubahan kecil yang mungkin hanya mewakili *noise* dalam data.

Proses pengujian model *Decision Tree* dilakukan dengan data pengujian yang telah didefinisikan sebelumnya, yakni sejumlah 126 data. Akurasi digunakan sebagai metrik evaluasi untuk mengukur kinerja model dan memberikan nilai sebesar 0,99 yang ditunjukkan pada Tabel 2.

Tabel 2. Hasil Pengujian Model *Decision Tree*

| Label | Presisi | Recall | F1-score | Support |
|----------|---------|--------|----------|---------|
| 0 | 1,00 | 1,00 | 1,00 | 25 |
| 1 | 1,00 | 1,00 | 1,00 | 25 |
| 2 | 1,00 | 0,96 | 0,98 | 27 |
| 3 | 0,96 | 1,00 | 0,98 | 22 |
| 4 | 1,00 | 1,00 | 1,00 | 27 |
| Accuracy | | | 0,99 | 126 |

Tabel 2 menunjukkan adanya data yang salah diklasifikasikan yakni pada data pada kelas 2 dan 3. Berdasarkan Tabel 3, model gagal mengklasifikasikan satu data pada kelas 2 menjadi kelas 3. Akibatnya, kinerja model menurun dan

menghasilkan nilai *presisi*, *recall*, dan *F1-score* yang lebih rendah pada Tabel 2.

Tabel 3. *Confusion matrix model decision tree*

| | | Predicted Label | | | | |
|------------|---|-----------------|----|----|----|----|
| | | 0 | 1 | 2 | 3 | 4 |
| True Label | 0 | 25 | 0 | 0 | 0 | 0 |
| | 1 | 0 | 25 | 0 | 0 | 0 |
| | 2 | 0 | 0 | 26 | 1 | 0 |
| | 3 | 0 | 0 | 0 | 22 | 0 |
| | 4 | 0 | 0 | 0 | 0 | 27 |

Model Random Forest diimplementasikan untuk mengatasi misklasifikasi yang terjadi khususnya sensitivitas terhadap variasi kecil dalam data. Random Forest akan membangun banyak pohon keputusan independen secara paralel dimana setiap pohon dibangun dengan subset acak dari data pelatihan. Hasil prediksi dari setiap pohon akan digabungkan dengan cara voting atau rata-rata untuk regresi dalam menghasilkan prediksi akhir.

Tabel 4 menunjukkan evaluasi kinerja model *Random Forest* dengan akurasi model 1,0. Hal ini membuktikan bahwa model Random Forest dapat mengatasi kesalahan klasifikasi pada label 2 dan 3 dan berhasil mendeteksi keseluruhan data dengan benar.

Tabel 4. Hasil pengujian model *random forest*

| Label | Presisi | Recall | F1-score | Support |
|----------|---------|--------|----------|---------|
| 0 | 1,00 | 1,00 | 1,00 | 25 |
| 1 | 1,00 | 1,00 | 1,00 | 25 |
| 2 | 1,00 | 1,00 | 1,00 | 27 |
| 3 | 1,00 | 1,00 | 1,00 | 22 |
| 4 | 1,00 | 1,00 | 1,00 | 27 |
| Accuracy | | | 1,00 | 126 |

5. KESIMPULAN DAN SARAN

Stres merupakan masalah kesehatan mental yang perlu disorot karena menjadi penyumbang kematian dengan angka yang cukup besar. Deteksi stres berdasarkan kondisi tidur merupakan pendekatan yang sangat menjanjikan karena tidur merupakan kegiatan yang pasti dilakukan oleh seseorang sehingga data yang dihasilkan menimbulkan bias yang minimum. Model deteksi tingkat stres ini dibangun dengan algoritma Decision Tree dan menghasilkan akurasi 0,99. Angka ini menunjukkan bahwa model tidak mampu memprediksi semua data dengan baik sehingga masih ada data yang salah diklasifikasikan. Model lainnya kemudian dikembangkan dengan algoritma berbasis *tree* lainnya yakni Random Forest dan secara efektif dapat menangani masalah misklasifikasi yang terjadi. Model Random Forest ini memberikan akurasi sebesar 1,0.

Penelitian ini berhasil membuktikan bahwa kondisi dan kebiasaan ketika tidur memiliki hubungan dengan tingkat stres yang sedang dialami. Tingkat stres berbanding lurus dengan tingkat mendengkur, laju respirasi, pergerakan anggota tubuh termasuk bola mata, serta detak jantung saat tidur. Peningkatan nilai tersebut sejalan dengan semakin

tingginya tingkat stres yang sedang diderita. Di sisi lain suhu tubuh, kadar oksigen, dan waktu tidur memberikan hasil sebaliknya, peningkatan ketiga nilai ini berbanding terbalik dengan tingkat stres yang sedang dialami sehingga mengindikasikan tingkat stres yang lebih rendah. Pengetahuan ini diharapkan dapat menjadi insight bagi peneliti lain terutama bagi peneliti dengan bidang sejenis sehingga dapat menghasilkan informasi lain yang lebih berharga.

Data yang digunakan dalam penelitian ini merupakan data kontinyu yang dapat merepresentasikan *state* selama seseorang tidur dan model sama-sama dikembangkan dengan basis *tree*. Kinerja model dengan akurasi 1,0 tidak menjamin keandalan model jika diterapkan pada data yang lain khususnya data diskrit. Dalam pengembangan penelitian selanjutnya, model dapat dikembangkan dengan algoritma lain yang lebih kompleks khususnya yang dapat mengolah data diskrit.

DAFTAR PUSTAKA

- AHUJA, R. and BANGA, A. 2019. Mental Stress Detection in University Students using Machine Learning Algorithms. *Procedia Computer Science*, [online] 152, pp.349–353. doi:<<https://doi.org/10.1016/j.procs.2019.05.007>> [Diakses 28 Maret 2023]
- CRUZ, A.P., PRADEEP, A., SIVASANKAR, K.R. dan KRISHNAVENI, K.S. 2020. A Decision Tree Optimised SVM Model for Stress Detection using Biosignals. 2020 International Conference on Communication and Signal Processing (ICCSP). [online] Tersedia di : <<https://doi.org/10.1109/iccsp48568.2020.9182043>> [Diakses 28 Maret 2023]
- FRATELLO, M. AND TAGLIAFERRI, R. (2019). *Encyclopedia of Bioinformatics and Computational Biology*. [online] Google Books. Tersedia di: https://books.google.co.id/books?hl=en&lr=&id=rs51DwAAQBAJ&oi=fnd&pg=PA374&dq=decision+tree+random+forest+for+non+linear+separable+data&ots=q_1Z4hSoK-&sig=GmIvzATm-Vxq5dtKAukeogX7UjU&redir_esc=y#v=onepage&q=decision%20tree%20random%20forest%20for%20non%20linear%20separable%20data&f=false [Diakses 26 September 2023].
- FRIJANTO, A. (2022). Depresi dan Bunuh Diri. [online] Kemkes.go.id. Tersedia di : https://yankes.kemkes.go.id/view_artikel/1450/depresi-dan-bunuh-diri#:~:text=Kondisi%20memprihatinkan%20ini%20juga%20didukung,juta%20pertahun%20di%20seluruh%20dunia. [Diakses 3 Agustus 2023].
- CEJA, G., E., RIEGLER, M., NORDGREEN, T., JAKOBSEN, P., OEDEGAARD, K., J., TORRESEN, J. (2018) ‘Mental Health Monitoring with Multimodal Sensing and machine learning: A survey’, *Pervasive and Mobile Computing*, 51, pp. 1–26. doi:10.1016/j.pmcj.2018.09.003.
- JORDAN, M. I., MITCHELL, T. M. 2015. Machine learning: Trends, perspectives, and prospects. [online] Tersedia di : <<https://www.science.org/doi/10.1126/science.aaa8415>> [Diakses 3 Agustus 2023]
- KAMUS BESAR BAHASA INDONESIA (KBBI) Online. Dalam Kamus Besar Bahasa Indonesia. [online] Tersedia di : <<https://kbbi.kemdikbud.go.id/entri/stres>> [Diakses 1 Mei 2023]
- KARTIKA, A.A. (2021) Hubungan tingkat stres Dengan Kualitas Tidur Pada mahasiswa/Mahasiswi fakultas Kedokteran Pada saat pandemi covid-19, *Majalah Kedokteran Andalas*. Tersedia di: <http://jurnalmka.fk.unand.ac.id/index.php/art/article/view/873> [Diakses 22 Agustus 2023]
- LAAVANYA RACHAKONDA, BAPATLA, A.K., MOHANTY, S.P. AND KOUIGIANOS, E. (2021). SaYoPillow: Blockchain-Integrated Privacy-Assured IoMT Framework for Stress Management Considering Sleeping Habits. *IEEE Transactions on Consumer Electronics*, [online] 67(1), pp.20–29. doi:<https://doi.org/10.1109/tce.2020.3043683>.
- RACHAKONDA, L. 2021. Human Stress Detection in and through Sleep. [online] Kaggle.com. Tersedia di: <https://www.kaggle.com/datasets/laavanya/human-stress-detection-in-and-through-sleep?select=SaYoPillow.csv> [Diakses 13 April 2023].
- ROSHANSKI, I., KALECH, M., ROKACH, L. 2023. Automatic Feature Engineering for Learning Compact Decision Trees. [online] Tersedia di : <<https://doi.org/10.1016/j.eswa.2023.120470>> [Diakses 3 Agustus 2023]
- SHANBHOG M, S., and MEDIKONDA, J. 2023. A clinical and technical methodological review on stress detection and sleep quality prediction in an academic environment. [online] Tersedia di : <<https://www.sciencedirect.com/science/article/pii/S0169260723001864>> [Diakses 3 Agustus 2023]
- UB.AC.ID. 2023. Tampilan Implementasi Metode Decision Tree untuk Sistem Pendeteksi Stres berdasarkan Detak Jantung dan Kelenjar Keringat. [online] Tersedia di : <<https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/12126/5523>> [Diakses 25 Maret 2023]
- VEERAMACHANENI, K. et al. (2019) ‘Intraindividual variability in sleep and perceived stress in young adults’, *Sleep Health*,

5(6), pp. 572–579.
doi:10.1016/j.sleh.2019.07.009.

YOVITA (2021). Algoritma Machine Learning yang Harus Kamu Pelajari di Tahun 2021. [online] Dqlab.id. Tersedia di :
<https://dqlab.id/algoritma-machine-learning-yang-perlu-dipelajari> [Diakses 3 Agustus 2023].