**Analisis Perbandingan Algoritma Klasifikasi Random Forest dan AdaBoost dalam Mendeteksi Tingkat Stress**

**Abstrak**

**Bab I Pendahuluan**

**1.1. Latar Belakang**

**Kerangkanya:**

**Jelaskan indoensia kemudian sangkutkan dengan stress**

**Stress di Indonesia**

**Klasifikasi stress konvensional**

**Adanya AI sebagai teknologi mutahir yang digunakan di dunia kesehatan**

**Algoritma Klasifikasi Random Forest dan AdaBoost**

**Menjawab permasalahan mengenai tingkat stress**

**Menjawab SDGs 3**

**Dibandingkan dengan negara lain**

**Bab II Tinjauan Pustaka**

**Bab III Metode Penelitian**

**Bab IV Implementasi**

**Bab V Hasil Dan Pembahasan/Hasil Evaluasi Aplikasi Yang Telah Dikembangkan**

**Bab VI Penutup**

**Daftar Referensi**

**BAB I PENDAHULUAN**

**1.1. Latar Belakang**

Stres adalah kondisi tekanan mental yang dirasakan oleh individu saat menghadapi masalah dari lingkungan dan kesejahteraan sosial, yang dapat menyebabkan berbagai penyakit (Bhargava and Trivedi, 2018). Stres merupakan respons terhadap stres yang terkait dengan pekerjaan secara kronis dan merupakan upaya untuk beradaptasi atau melindungi diri dari stres tersebut. Tingkatan stres dapat diukur menggunakan berbagai alat, salah satunya adalah skala Depression Anxiety Stress Scale (DASS) dan Perceived Stress Scale (PSS) yang juga digolongkan menjadi tangkat rendah sedang daan berat (Hutabarat, Elita and Utomo, 2022). Stres dapat disebabkan oleh faktor lingkungan, faktor psikologis, faktor biologis, dan faktor social (Bhargava and Trivedi, 2018).

Dalam konteks pemuda, stres menjadi sangat penting karena ini adalah periode perubahan signifikan dalam hidup mereka. Mereka diharapkan menjadi elit dalam masyarakat, dan oleh karena itu, mereka harus meningkatkan kemampuan manajemen stres mereka untuk menjalani kehidupan yang sehat setelah memasuki Masyarakat (Williams and Williams, 2020). Stres utamanya berasal dari ujian akademis, hubungan interpersonal, masalah hubungan, perubahan hidup, dan eksplorasi karir. Stres seperti itu biasanya dapat menyebabkan masalah psikologis, fisik, dan perilaku (Lupien et al., 2009).

Peristiwa yang memicu stres disebut stresor, dan mereka mencakup berbagai situasi — mulai dari bahaya fisik nyata hingga menyajikan presentasi kelas atau menghadapi mata kuliah paling sulit selama satu semester. Beberapa stresor utama di antara mahasiswa adalah akademis, keuangan, hubungan, karier, dan manajemen waktu (Ismail Nur Maulid, Titi Saparina L and Noviati, 2022).

Di tempat kerja, burnout terjadi ketika para profesional menggunakan strategi penanganan yang tidak efektif untuk mencoba melindungi diri dari stres terkait pekerjaan. Dimensi 'overload' (beban kerja berlebih), 'lack of development' (kurangnya perkembangan), dan 'neglect' (pengabaian), yang termasuk dalam tipe 'frenetic' (sibuk berlebihan), 'under-challenged' (tidak cukup tantangan), dan 'worn-out' (letih), secara singkat mendefinisikan burnout (Montero-Marin et al., 2014).

Di era komunikasi digital, video konferensi juga telah diidentifikasi sebagai sumber stres baru, yang disebut sebagai Zoom fatigue. Deplesi sumber daya fisiologis dan kognitif ini sebagai konsekuensi dari penggunaan yang berkepanjangan dan tidak sesuai dengan alat video konferensi (Riedl, 2022).

Memahami sumber stres dan bagaimana individu dapat mengatasi stres sangat penting. Setelah mengidentifikasi penyebabnya, lebih banyak penekanan dapat diberikan pada tahap perkembangan anak menjadi remaja. Mereka seharusnya dibesarkan dalam lingkungan yang positif, dan lebih banyak penekanan harus diberikan pada kegiatan di luar ruangan dan menciptakan lingkungan belajar yang ramah dengan meminimalkan dampak negatif stresor (Putri and Azalia, 2022).

Kecerdasan buatan (AI) telah menjadi teknologi mutakhir yang digunakan dalam berbagai sektor, termasuk kesehatan. AI telah digunakan dalam berbagai aplikasi kesehatan, mulai dari diagnosis penyakit hingga pengembangan obat dan terapi baru (Hayat, Sutarno and Erwin, 2017).

Salah satu contoh penggunaan AI dalam dunia kesehatan adalah dalam sistem pakar untuk diagnosis penyakit. Misalnya, metode certainty factor telah digunakan dalam sistem pakar untuk mendiagnosis penyakit jerawat. Sistem ini dapat mendeteksi penyakit jerawat dengan akurasi 100%, memudahkan pengguna untuk berkonsultasi tanpa harus bertemu dengan dokter spesialis (Puteri and Bhakti, 2019).

Random Forest adalah algoritma klasifikasi yang termasuk dalam kategori ensemble learning. Algoritma ini bekerja dengan cara membuat banyak pohon keputusan (decision trees) dan menggabungkan hasil dari semua pohon tersebut untuk memberikan prediksi akhir. Keuntungan dari metode ini adalah kemampuannya untuk mengurangi overfitting, yang sering terjadi pada pohon keputusan Tunggal (Pahlevi, Amrin and Handrianto, 2023).

Random Forest telah digunakan dalam berbagai aplikasi. Misalnya, dalam penelitian yang dilakukan untuk mendeteksi diabetes secara dini, Random Forest berhasil memberikan tingkat akurasi tertinggi 97,88% dibandingkan dengan algoritma lain seperti Support Vector Machine dan Naive Bayes. Dalam penelitian lain, Random Forest digunakan untuk klasifikasi kelayakan kredit dan memberikan tingkat akurasi sebesar 78,60% (Apriliah et al., 2021).

AdaBoost (Adaptive Boosting) juga merupakan algoritma klasifikasi yang termasuk dalam kategori ensemble learning. Algoritma ini bekerja dengan cara membuat banyak model prediksi yang lemah (weak learners) dan menggabungkan hasil dari semua model tersebut untuk memberikan prediksi akhir. Setiap model dibuat dengan mempertimbangkan kesalahan yang dibuat oleh model sebelumnya, sehingga model berikutnya berusaha untuk memperbaiki kesalahan tersebut (Majid and Dwi Miharja, 2022).

Berdasarkan latar belakang yang telah diuraikan sebelumnya, algoritma Random Forest dan AdaBoost dapat dimanfaatkan untuk melakukan klasifikasi terhadap tingkat stres pada individu, terutama dalam konteks pemuda. Pemanfaatan algoritma ini dapat membantu dalam mengidentifikasi faktor-faktor yang menyebabkan stres serta memberikan wawasan tentang tingkat stres yang dihadapi oleh seseorang. Dalam penyusunan program, digunakan *dataset* terkait tingkat stress sebagaimana dapat diakses pada Kaggle (rxnach, 2023). Dataset ini mencakup berbagai faktor yang dapat mempengaruhi tingkat stres, termasuk lingkungan, faktor psikologis, faktor biologis, dan faktor social.

Dengan penerapan program klasifikasi tingkat stres menggunakan algoritma Random Forest dan AdaBoost, diharapkan dapat memberikan kontribusi signifikan terhadap pencapaian Sustainable Development Goals (SDGs) nomor 3 yang berkaitan dengan kesehatan dan kesejahteraan. Program ini diharapkan mampu memberikan pemahaman yang lebih baik terkait faktor-faktor yang berkontribusi pada tingkat stres individu, khususnya pada kalangan pemuda. Dengan memahami sumber-sumber stres dan tingkat stres yang dialami, intervensi yang lebih tepat dapat diarahkan untuk meningkatkan kesejahteraan psikologis dan fisik. Dengan demikian, kontribusi program ini diharapkan dapat menciptakan masyarakat yang lebih sehat dan sejahtera, sejalan dengan tujuan pembangunan berkelanjutan (Nations, 2023).

**1.2. Rumusan Masalah**

Berdasarkan latar belakang yang telah disampaikan, rumusan masalah yang dapat diangkat adalah sebagai berikut.

1. Bagaimana efektivitas algoritma klasifikasi Random Forest dan AdaBoost dalam mendeteksi dan mengklasifikasikan tingkat stres pada pemuda, khususnya yang disebabkan oleh faktor lingkungan, faktor psikologis, faktor biologis, dan faktor sosial?
2. Bagaimana proses *deployment* dari program klasifikasi tingkat stress dengan algoritma Random Forest dan AdaBoost?

**1.3. Tujuan Penelitian**

Berdasarkan rumusan masalah yang telah disampaikan, penulisan makalah ini dirincikan sebagai berikut.

1. Mengukur efektivitas algoritma klasifikasi Random Forest dan AdaBoost dalam mendeteksi dan mengklasifikasikan tingkat stres pada pemuda, khususnya yang disebabkan oleh faktor lingkungan, faktor psikologis, faktor biologis, dan faktor sosial.
2. Mengetahui proses *deployment* dari program klasifikasi tingkat stress dengan algoritma Random Forest dan AdaBoost.

1.4. Manfaat Penelitian

Manfaat yang diharapkan pada penelitian ini dirincikan sebagai berikut.

**1.4.1. Manfaat Praktis**

Penelitian ini dapat membantu dalam mengidentifikasi tingkat stress berdasarkan algoritma klasifikasi Random Forest dan AdaBoost.

**1.4.2. Manfaat Teoritis**

Penelitian ini memberikan kontribusi dalam pengembangan dan penerapan algoritma klasifikasi Random Forest dan AdaBoost dalam konteks kesehatan mental, khususnya dalam mendeteksi dan mengklasifikasikan tingkat stress.

**1.5. Batasan Masalah**

Batasan masalah dalam penelitian ini adalah fokus pada penggunaan algoritma klasifikasi Random Forest dan AdaBoost untuk mendeteksi dan mengklasifikasikan tingkat stres pada pemuda. Selain itu, penelitian ini juga terbatas pada penggunaan dataset “Student Stress Factors: A Comprehensive Analysis” dari penyedia *open dataset* Kaggle (rxnach, 2023).

**Bab 2 Tinjauan Pustaka**

**2.1. Penelitian Terdahulu**

Penelitian terdahulu diperlukan untuk memperoleh bahan referensi dan perbandingan. Hal ini juga bertujuan untuk menghindari kesan kesamaan penelitian terdahulu dengan karya ilmiah yang dibuat. Oleh karena itu, dalam tinjauan pustaka ini, penulis menyertakan temuan-temuan dari penelitian sebelumnya yang dirincikan pada Tabel 2.1.

Tabel 2.1. Penelitian terdahulu

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Sumber** | **Tahun** | **Fokus** | **Data** | **Metode** | **Hasil** |
| [1](https://www.semanticscholar.org/paper/f8a46a718f64f9d398dcc40c74fd6fd68d677651" \t "_blank) | 2023 | Deteksi penyakit jantung | Dataset Kaggle Heart Disease Dataset (<https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset>) | SVM, Random Forest, Logistic Regression, AdaBoost | Random Forest, SVM, dan AdaBoost memiliki nilai terbaik dan sama pada hasil pengujian dengan akurasi 0.985366 |
| [2](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8289583/) | 2021 | Klasifikasi detak jantung (Arrhythmia) | MIT-BIH dataset | AdaBoost + Random Forest | AdaBoost + Random Forest mencapai akurasi 99.11% |
| [4](https://www.semanticscholar.org/paper/920b7f653e825ef3fc926ae87a1f03fcd9831242) | 2010 | Klasifikasi ekosistem | LANDSAT TM/ETM+ images dan GIS data layers | Random Forest dan Adaboost tree | Adaboost tree memberikan akurasi klasifikasi yang lebih tinggi dibandingkan Random Forest |
| [5](https://www.semanticscholar.org/paper/b5d345c69c244dec47bb75d21f09f328acb61510" \t "_blank) | 2021 | Deteksi website phishing | Tidak disebutkan | Support Vector Machine, Decision Tree, Random Forest, dan Multilayer Perceptron | Multilayer Perceptron memiliki performa terbaik dengan akurasi 93.15% |
| [7](https://www.semanticscholar.org/paper/acef0fa2474cd06f43f8634d5a68fa947db00fce) | 2021 | Prediksi dan klasifikasi tingkat stres | PhysioBank dataset | Gradient boost, KNN, Random Forest, dan SVM | Gradient boost memiliki performa terbaik dengan akurasi 83.33% |
| [9](https://www.semanticscholar.org/paper/51b7881da2240a0780827426e5c26b39dc1d91e3) | 2022 | Prediksi kesuksesan Kickstarter | 5723 data campaign kategori teknologi dari portal Kickstarter | Backpropagation ANN dan Random Forest | Random Forest memiliki akurasi tertinggi sebesar 98% |

**2.2. Kajian Teori**

The theoretical review for the research on "Comparative Analysis of Random Forest and AdaBoost Classification Algorithms in Detecting Stress Levels" can be constructed based on the understanding of these two algorithms and their applications in stress detection, as found in the search results.

## Random Forest Algorithm

Random Forest is a machine learning algorithm that operates by constructing multiple decision trees and outputting the class that is the mode of the classes or mean prediction of the individual trees[1][3][6]. It is known for its high accuracy, robustness to outliers, and ability to handle large datasets with high dimensionality[1][3][6].

In the context of stress detection, Random Forest has been used effectively. For instance, a study used Random Forest for stress detection using a multimodal dataset, WESAD, which includes sensor data like electrocardiogram (ECG), body temperature, respiration, electromyogram (EMG), and electrodermal activity (EDA). The Random Forest model outperformed other models for both binary classification (stress vs. non-stress) and three-class classification (amusement vs. baseline vs. stress) with F1-scores of 83.34 and 65.73 respectively[5].

## AdaBoost Algorithm

AdaBoost (Adaptive Boosting) is an ensemble learning algorithm often used in boosting that can be combined with other classification algorithms to improve classification performance[2][4]. It is adaptive in nature, where subsequent classifiers are built to correct the classification errors made by the previous classifier[2]. AdaBoost is sensitive to noisy data and outliers. In some cases, AdaBoost is less susceptible to the overfitting problem compared to other learning algorithms[2].

In terms of stress detection, the AdaBoost algorithm has not been directly mentioned in the search results. However, AdaBoost has been used to optimize prediction accuracy in other contexts, such as smoke detection, where it increased the resulting accuracy value to 0.987 and precision value to 0.971[2].

## Comparative Analysis

Both Random Forest and AdaBoost are powerful machine learning algorithms used for classification tasks. While Random Forest is known for its high accuracy and robustness, AdaBoost is recognized for its ability to improve the performance of other classification algorithms. In the context of stress detection, Random Forest has been directly applied and shown promising results[5]. However, the application of AdaBoost in stress detection would need further exploration. The comparative analysis of these two algorithms in detecting stress levels would provide valuable insights into their performance, strengths, and weaknesses in this specific context.

Citations:

[1] https://www.semanticscholar.org/paper/c1e2163131070ed9240e27168e7f62b79a885cdd

[2] https://www.semanticscholar.org/paper/b8b55615c7308069422f6a10fb62663daf2b3713

[3] https://www.semanticscholar.org/paper/b123b0132fa2ca01e5769cc88c3036b2719f06e5

[4] https://www.semanticscholar.org/paper/6bfc2c1e1b3e3df9f622bfe144b7832d12b98a93

[5] https://www.semanticscholar.org/paper/dc16e948105c6e92dfea91513399122378eae00d

[6] https://www.semanticscholar.org/paper/6bfb934005d2e778a9e68d2250ed3b3ac1e2bd93

[7] https://www.semanticscholar.org/paper/8551e488332952e9d9f1d9402c75a6b5f1bb7fc9

[8] https://www.semanticscholar.org/paper/f768bbc4a4526e5f721ea06466434cd29a98aa47

[9] https://www.semanticscholar.org/paper/77dc3f93b63e3253cf450a7a4f183ad839b84fdc

[10] https://www.semanticscholar.org/paper/410a37bd8fa1264c81824bd055d3dcd44e91ca7e

[11] https://www.semanticscholar.org/paper/33e72162a181dfd00310e19c866c66821bf2a68b

[12] https://www.semanticscholar.org/paper/98ea113d1789b954c2f130af9e94a7c11e11596b

[13] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9955749/

[14] https://www.semanticscholar.org/paper/916472f19fadfbc417e45e32cf924ba2f19dd0bf

[15] https://www.semanticscholar.org/paper/c0a4350d65eda934f5728145152eef027176b76e

[16] https://www.semanticscholar.org/paper/19ebc352720cb04a4afb22525534ebf333a68d66

[17] https://www.semanticscholar.org/paper/340f5536d0b0a3c85dac280398da26e9d32372ea

[18] https://www.semanticscholar.org/paper/67eeef3e9b1e3944018749968aa79cacb8757f1f

[19] https://www.semanticscholar.org/paper/52457ba2cb4924f77cc498ed7ce640f4a3f2f070

[20] https://www.semanticscholar.org/paper/de3562f2bec1bbcbc0e4f2d9a7f8defd6ba01edf

[21] https://www.semanticscholar.org/paper/85decdb84ca0dc98fdb9a7420234daed5ff3d1ff

[22] https://www.semanticscholar.org/paper/4673168359abe58a8c1f6c1f3250f0f8d3457987

[23] https://www.semanticscholar.org/paper/b82db502c7b960169064c6af9f21225c0333f505

[24] https://www.semanticscholar.org/paper/f8a46a718f64f9d398dcc40c74fd6fd68d677651

[25] <https://www.semanticscholar.org/paper/955f4047cffba6f18ea94edb097593fc2b84c306>

DAFTAR PUSTAKA

Apriliah, W., Kurniawan, I., Baydhowi, M. and Haryati, T., 2021. Prediksi Kemungkinan Diabetes pada Tahap Awal Menggunakan Algoritma Klasifikasi Random Forest. *SISTEMASI*, 10(1). https://doi.org/10.32520/stmsi.v10i1.1129.

Bhargava, D. and Trivedi, H., 2018. A Study of Causes of Stress and Stress Management among Youth. *IRA-International Journal of Management & Social Sciences (ISSN 2455-2267)*, 11(3). https://doi.org/10.21013/jmss.v11.n3.p1.

Hayat, M.S., Sutarno, S. and Erwin, E., 2017. Lorong Waktu Revolusi Saintifik pada Era Eksponensial. *Titian Ilmu: Jurnal Ilmiah Multi Sciences*, 9(1). https://doi.org/10.30599/jti.v9i1.80.

Hutabarat, N.S., Elita, V. and Utomo, W., 2022. FAKTOR - FAKTOR PENYEBAB STRES AKADEMIK MAHASISWA KEPERAWATAN TAHUN PERTAMA SELAMA PEMBELAJARAN JARAK JAUH DI MASA PANDEMI COVID-19. *Coping: Community of Publishing in Nursing*, 10(5). https://doi.org/10.24843/coping.2022.v10.i05.p06.

Ismail Nur Maulid, Titi Saparina L and Noviati, 2022. Hubungan Kuliah Daring dengan Tingkat Stres Mahasiswa Kesehatan Masyarakat Universitas Mandala Waluya. *Jurnal Healthy Mandala Waluya*, 1(2). https://doi.org/10.54883/jhmw.v1i2.24.

Lupien, S.J., McEwen, B.S., Gunnar, M.R. and Heim, C., 2009. *Effects of stress throughout the lifespan on the brain, behaviour and cognition*. *Nature Reviews Neuroscience*, https://doi.org/10.1038/nrn2639.

Majid, A.M. and Dwi Miharja, M.N., 2022. PENERAPAN METODE DISCRETIZATION DAN ADABOOST UNTUK MENINGKATKAN AKURASI ALGORITMA KLASIFIKASI DALAM MEMPREDIKSI PENYAKIT JANTUNG. *Indonesian Journal of Business Intelligence (IJUBI)*, 5(2). https://doi.org/10.21927/ijubi.v5i2.2689.

Montero-Marin, J., Prado-Abril, J., Piva Demarzo, M.M., Gascon, S. and García-Campayo, J., 2014. Coping with stress and types of burnout: Explanatory power of different coping strategies. *PLoS ONE*, 9(2). https://doi.org/10.1371/journal.pone.0089090.

Nations, U., 2023. *Key Activities for Sustainable Development Goal 3: Good Health and Well-Being*. Available at: <https://indonesia.un.org/en/sdgs/3/key-activities>.

Pahlevi, O.-, Amrin, A.- and Handrianto, Y.-, 2023. Implementasi Algoritma Klasifikasi Random Forest Untuk Penilaian Kelayakan Kredit. *Jurnal Infortech*, 5(1). https://doi.org/10.31294/infortech.v5i1.15829.

Puteri, A.G. and Bhakti, R.M.H., 2019. Penggunaan Certainty Factor Dalam Sistem Pakar Diagnosa Penyakit Jerawat. *Jurnal Ilmiah Intech : Information Technology Journal of UMUS*, 1(02). https://doi.org/10.46772/intech.v1i02.72.

Putri, T.H. and Azalia, D.H., 2022. Faktor yang Memengaruhi Stres pada Remaja Selama Pandemi Covid-19. *Jurnal Keperawatan Jiwa*, 10(2). https://doi.org/10.26714/jkj.10.2.2022.285-296.

Riedl, R., 2022. On the stress potential of videoconferencing: definition and root causes of Zoom fatigue. *Electronic Markets*, 32(1). https://doi.org/10.1007/s12525-021-00501-3.

rxnach, 2023. *Student Stress Factors: A Comprehensive Analysis*. https://doi.org/10.31219/osf.io/x2k2t.

Williams, S.D. and Williams, J., 2020. Posttraumatic stress in organizations: Types, antecedents, and consequences. *Business and Society Review*, 125(1). https://doi.org/10.1111/basr.12192.