

Identifikasi Tingkat Stres Mahasiswa Menggunakan Algoritma XGBoost dengan Data Psikologis, Fisik, Akademik, dan Lingkungan

¹Muhamad Isa Firdaus, ²Hilmy Aliy Andra Putra, ³Muhammad Encep
^{1,2,3}Program Studi Ilmu Komputer, Universitas Djuanda, Bogor

E-mail: ¹muhamadisafirdaus123@gmail.com, ²hilmyaliy@unida.ac.id,
³ahmadpoac@unida.ac.id

ABSTRAK

Stres merupakan permasalahan kesehatan yang banyak dialami mahasiswa akibat tuntutan akademik, psikologis, fisik, dan lingkungan yang saling berkaitan dan bersifat multidimensi, sehingga pendekatan identifikasi konvensional dinilai kurang memadai untuk mengungkap pola yang kompleks. Penelitian ini menerapkan algoritma *Extreme Gradient Boosting* (XGBoost) untuk mengidentifikasi tingkat stres mahasiswa berdasarkan *StressLevelDataset*, mengetahui kinerja algoritma tersebut, serta mengidentifikasi faktor yang paling berpengaruh dalam proses klasifikasi. Model dibangun melalui dua tahap, yaitu model *baseline* dengan parameter bawaan dan model hasil *hyperparameter tuning* menggunakan *Grid Search*. Hasil penelitian menunjukkan bahwa model XGBoost hasil *tuning* yang dievaluasi menggunakan *Stratified 10-Fold Cross-Validation* memperoleh akurasi 88,91%, presisi 88,95%, *recall* 88,92%, dan *F1-score* 88,93%, meningkat dibandingkan model *baseline* (akurasi 87,27%). Analisis *feature importance* mengungkap bahwa *blood_pressure*, *sleep_quality*, dan *safety* merupakan tiga variabel paling berpengaruh berdasarkan nilai *gain*, dengan *blood_pressure* sebagai faktor paling dominan ($\pm 45,6\%$). Hasil ini menunjukkan bahwa algoritma XGBoost efektif dan dapat diandalkan untuk mengidentifikasi tingkat stres mahasiswa secara objektif berbasis data.

Kata kunci : *feature importance, klasifikasi, machine learning, stres, XGBoost*

ABSTRACT

Stress is a health problem widely experienced by students due to interrelated and multidimensional academic, psychological, physical, and environmental demands, making conventional identification approaches inadequate for revealing complex patterns. This study applies the *Extreme Gradient Boosting* (XGBoost) algorithm to identify student stress levels based on the *StressLevelDataset*, evaluate the performance of the algorithm, and identify the most influential factors in the classification process. The model was built in two stages, namely a baseline model with default parameters and a model resulting from hyperparameter tuning using *Grid Search*. The results show that the tuned XGBoost model, evaluated using *Stratified 10-Fold Cross-Validation*, achieved an accuracy of 88.91%, a precision of 88.95%, a recall of 88.92%, and an *F1-score* of 88.93%, an improvement over the baseline model (87.27% accuracy). Feature importance analysis revealed that *blood_pressure*, *sleep_quality*, and *safety* were the three most influential variables based on gain values, with *blood_pressure* as the most dominant factor ($\pm 45.6\%$). These results indicate that the XGBoost algorithm is effective and reliable for objectively identifying student stress levels based on data.

Keyword : *classification, feature importance, machine learning, stress, XGBoost*

1. PENDAHULUAN

Stres merupakan respons non-spesifik tubuh terhadap setiap tuntutan atau tekanan yang dihadapinya, baik yang berasal dari sumber fisik maupun psikologis (Selye 1976). Laporan *State of the World's Emotional Health 2025* dari Gallup mencatat bahwa 37% dari lebih 145.000 responden dewasa di 144 negara merasakan stres berat, dan mahasiswa menjadi salah satu kelompok dengan beban stres tinggi. *Healthy Minds Study 2024-2025* mencatat 32% dari lebih 84.000 mahasiswa di 135 institusi Amerika Serikat mengalami kecemasan tingkat sedang hingga berat. Beban studi, tekanan ujian, dan pengelolaan waktu yang buruk menjadi pemicu utama stres akademik mahasiswa (Misra and Michelle 2000), yang jika tidak dikelola dapat memicu gejala kecemasan dan depresi. Karena faktor psikologis, fisik, akademik, dan lingkungan saling berkaitan dan bersifat multidimensi serta nonlinear, identifikasi stres yang hanya mengandalkan survei atau statistik deskriptif sederhana menjadi kurang memadai untuk mengungkap pola kompleks tersebut.

Penelitian ini memanfaatkan *machine learning* untuk mengidentifikasi tingkat stres mahasiswa secara objektif berdasarkan *StressLevelDataset*, yang berisi 1.100 data survei mahasiswa dengan 20 variabel fitur dan 1 variabel target mencakup faktor psikologis, fisik, akademik, dan lingkungan. Dari berbagai algoritma yang telah diuji pada kasus serupa, *Extreme Gradient Boosting (XGBoost)* dipilih karena sesuai untuk data multidimensi dan nonlinear, dilengkapi regularisasi L1/L2 yang menekan *overfitting* serta efisien pada data berdimensi tinggi (Chen and Guestrin 2016). *XGBoost* juga terbukti kompetitif pada studi stres mahasiswa sejenis yang mencapai akurasi 86,8% pada dataset serupa oleh (Yeler and Sürücü 2025), dan 96,17% pada data fisiologis *wearable* oleh

(Gedam, Dutta, and Jha 2025), serta dapat diinterpretasikan menggunakan SHAP untuk mengidentifikasi fitur paling berkontribusi.

Berbeda dari penelitian terdahulu yang menggabungkan *XGBoost* dengan skema *ensemble* dan seleksi fitur tambahan, penelitian ini berfokus pada optimasi algoritma *XGBoost* secara mandiri melalui *hyperparameter tuning* langsung pada seluruh fitur multidimensi dalam *StressLevelDataset*, guna mengetahui sejauh mana kinerjanya dapat dioptimalkan sebagai model klasifikasi tunggal.

2. LANDASAN TEORI

Stres

Menurut Selye (Selye 1976), stres adalah respons non-spesifik tubuh terhadap tuntutan atau tekanan fisik maupun psikologis, yang muncul sebagai bagian dari *General Adaptation Syndrome* (reaksi siaga, perlawanan, hingga kelelahan). Di Indonesia, prevalensi stres akademik mahasiswa berkisar 36,7%--71,6% (Yuda, Mawarti, and Mutmainnah 2023), dipengaruhi oleh interaksi faktor psikologis, fisik, akademik, dan lingkungan yang bersifat multidimensi.

StressLevelDataset

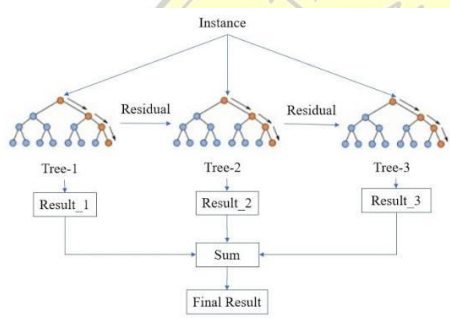
Dataset publik pada platform *Kaggle* ini berisi 1.100 data survei mahasiswa dengan 20 variabel fitur (psikologis, fisik, akademik, lingkungan) dan 1 variabel target berupa kategori tingkat stres (rendah, sedang, tinggi).

Machine Learning dan Supervised Learning

Machine learning adalah cabang kecerdasan buatan yang memungkinkan sistem "belajar" dari data tanpa diprogram secara eksplisit (Mitchell 1997). *Supervised learning* mempelajari pola dari data berlabel untuk memprediksi label pada data baru, dan pada penelitian ini dirumuskan sebagai masalah klasifikasi (James et al. 2013).

Ensemble Learning dan XGBoost

Ensemble learning menggabungkan prediksi beberapa model untuk menghasilkan prediksi yang lebih kuat dan stabil (Dietterich 2000), dengan tiga pendekatan utama: *bagging*, *boosting*, dan *stacking*. XGBoost adalah algoritma ensemble berbasis *boosting* yang membangun serangkaian *decision tree* secara bertahap, dengan setiap pohon baru memperbaiki *residual* pohon sebelumnya, dilengkapi regularisasi L1/L2 serta teknik komputasi seperti *weighted quantile sketch* dan *sparsity-aware split finding*. Parameter utama yang memengaruhi kinerjanya antara lain *n_estimators*, *learning_rate*, *max_depth*, dan *subsample* (Wade 2020).



Gambar 1. Konsep dasar algoritma XGBoost

Hyperparameter Tuning

Proses pencarian kombinasi *hyperparameter* optimal, salah satunya melalui *Grid Search* yang menguji seluruh kombinasi nilai secara sistematis (Kohavi and John 1995), dikombinasikan dengan *k-fold cross-validation* untuk estimasi kinerja yang stabil.

Evaluasi Model

Kinerja model klasifikasi dievaluasi menggunakan *confusion matrix* yang menghasilkan metrik *accuracy*, *precision*, *recall*, dan *F1-score*, sebagaimana dirumuskan berikut:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1 - score = \frac{2 \times (P \times R)}{P + R} \quad (4)$$

Uji McNemar

Uji McNemar (McNemar 1947) digunakan untuk menguji signifikansi statistik perbedaan performa dua model berpasangan (*baseline* vs *tuned*) pada sampel data yang sama yang berfokus pada *discordant pairs* di antara kedua model:

$$\chi^2 = \frac{(n_{01} - n_{10})^2}{n_{01} + n_{10}} \quad (5)$$

dengan n_{01} adalah jumlah data yang salah diklasifikasikan oleh Model A tetapi benar oleh Model B, dan n_{10} sebaliknya.

Feature Importance

Feature importance mengukur kontribusi setiap fitur terhadap prediksi model (Breiman 2001). Pada XGBoost, kontribusi diukur melalui metrik *gain*, dan dilengkapi metode SHAP (SHapley Additive exPlanations) yang mengukur kontribusi fitur berdasarkan nilai Shapley dari teori permainan (Lundberg and Lee 2017).

3. METODOLOGI

Objek penelitian ini adalah tingkat stres mahasiswa yang direpresentasikan dalam *StressLevelDataset*, diunduh dari platform Kaggle ("Student Stress Monitoring Datasets") dan dianalisis pada platform Google Colab menggunakan bahasa pemrograman Python. Dataset berisi 1.100 data survei mahasiswa berusia 18-21 tahun berbasis kuesioner skala Likert, dengan 20 variabel fitur (psikologis, fisik, akademik, lingkungan) dan 1 variabel target (*stress_level*). Sebagian variabel psikologis (*anxiety_level*, *depression*, *self_esteem*)

diukur menggunakan instrumen psikometri baku (GAD-7, PHQ-9, *Rosenberg Self-Esteem Scale*).

Tahapan penelitian mengacu pada kerangka *CRISP-DM* (Schröer, Kruse, and Gómez 2021), sebagai berikut:

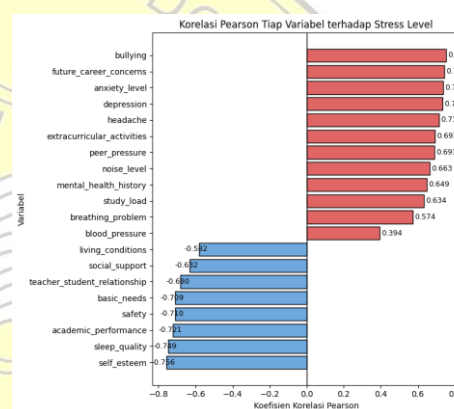
- Pengumpulan data yaitu tahap dataset dimuat ke *Google Colab* menggunakan pustaka *Pandas*.
- Pemahaman data yaitu tahap pemeriksaan struktur, statistik deskriptif, dan analisis korelasi *Pearson* antara setiap fitur dengan target.
- Pra-pemrosesan data yaitu tahap pengecekan *missing values*, data duplikat, dan deteksi/penanganan *outlier* menggunakan metode *Interquartile Range (IQR)*. Seluruh variabel telah numerik hasil skala *Likert* sehingga tidak memerlukan *encoding* maupun *feature scaling*, mengingat *XGBoost* berbasis pohon tidak sensitif terhadap skala.
- Cross-validation* merupakan evaluasi menggunakan *Stratified 10-Fold Cross-Validation* ($n_splits=10$, $shuffle=True$, $random_state=42$) agar proporsi kelas tetap seimbang pada setiap *fold*, menghasilkan prediksi *out-of-fold* yang merepresentasikan performa model pada seluruh dataset.
- Pelatihan model yaitu tahap model *XGBoost* dibangun dalam dua tahap: model *baseline* (parameter bawaan) dan model *tuned* (hasil *hyperparameter tuning* menggunakan *Grid Search* dengan kandidat $n_estimators$ (100, 200, 300), max_depth (3, 4, 5, 6), $learning_rate$ (0.01, 0.05, 0.1, 0.2), dan $subsample$ (0.7, 0.8, 1.0), dikombinasikan dengan *k-fold cross-validation*).
- Analisis dan evaluasi hasil yaitu tahap evaluasi kinerja menggunakan *confusion matrix* serta metrik akurasi, presisi, *recall*, dan *F1-score*; analisis *feature importance* berbasis *gain* dan *SHAP*; serta uji *McNemar* terhadap prediksi *out-of-fold* model *baseline*

dan *tuned* untuk menguji signifikansi statistik perbedaan performa.

4. HASIL DAN PEMBAHASAN

Pemahaman Data

Dataset terdiri atas 1.100 baris dan 21 kolom (20 fitur, 1 target). Hasil analisis korelasi *Pearson* menunjukkan hampir seluruh fitur berkorelasi cukup kuat dengan *stress_level* (0,39--0,76 nilai mutlak); *self_esteem* (-0,756), *bullying* (0,751), dan *sleep_quality* (-0,749) memiliki korelasi terkuat, sedangkan *blood_pressure* terlemah (0,394) meski tetap tergolong sedang, sebagaimana ditunjukkan pada Gambar 2.



Gambar 2. Korelasi *Pearson* setiap variabel fitur terhadap target

Pra-pemrosesan Data

Pemeriksaan *data cleaning* tidak menemukan *missing values* (0 dari 23.100 sel) maupun data duplikat (0 dari 1.100 baris). Distribusi kelas target relatif seimbang: 373 data (33,9%) kelas Rendah, 358 data (32,5%) kelas Sedang, dan 369 data (33,5%) kelas Tinggi.

Perbandingan Model

Proses *Grid Search* menghasilkan 144 kombinasi parameter yang diuji melalui *10-fold cross-validation* (1.440 kali pelatihan). Kombinasi terbaik diperoleh pada $n_estimators=300$, $max_depth=3$, $learning_rate=0,1$, dan $subsample=0,8$, dengan akurasi rata-rata *cross-validation*

88,91%, sebagaimana dirangkum pada Tabel 1.

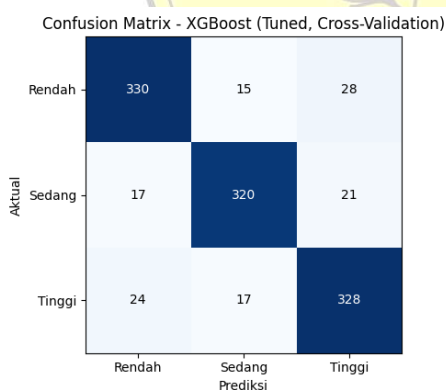
Tabel 1. Perbandingan kinerja model *XGBoost baseline* dengan *tuned*

Metrik	Baseline	Tuned	Peningkatan
Accuracy	87,27%	88,91%	+1,64 poin
Precision	87,36%	88,95%	+1,59 poin
Recall	87,28%	88,92%	+1,64 poin
F1-Score	87,31%	88,93%	+1,62 poin

Seluruh metrik model *tuned* meningkat secara konsisten dibandingkan *baseline*, mengindikasikan bahwa *hyperparameter tuning* memberikan kinerja klasifikasi yang lebih baik dan stabil, meskipun peningkatannya relatif kecil.

Hasil Evaluasi per Kelas

Model *tuned* memperoleh *F1-score* 0,89 (Rendah), 0,90 (Sedang), dan 0,88 (Tinggi), menunjukkan performa yang relatif konsisten tanpa kecenderungan kuat mengutamakan salah satu kelas. Berdasarkan *confusion matrix* pada Gambar 3, dari 1.100 data, 978 data (88,91%) diklasifikasikan benar dan 122 data mengalami kesalahan klasifikasi, dengan kesalahan terbesar terjadi antara kelas Rendah dan Tinggi (28 dan 24 data).



Gambar 3. *Confusion matrix* model *XGBoost* hasil *tuning*

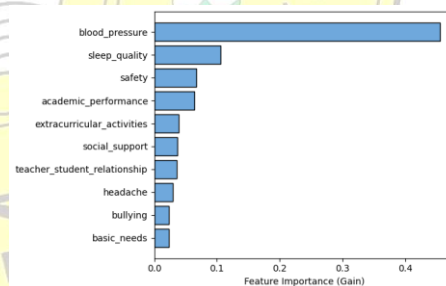
Uji McNemar

Berdasarkan tabel kontingensi prediksi *out-of-fold* kedua model, ditemukan 30 *discordant pairs* (6 data membaik, 24 data memburuk setelah *tuning*). Statistik uji McNemar diperoleh sebesar 3 dengan *p*-

value = 1,000 (*alfa* = 0,05), sehingga *H0* tidak ditolak. Artinya, meskipun secara numerik model *tuned* unggul, perbedaan performa antara model *baseline* dan *tuned* tidak signifikan secara statistik pada data yang dievaluasi.

Feature Importance (Gain)

Analisis *gain* pada model terbaik menunjukkan *blood_pressure* sebagai variabel paling dominan (45,60%), diikuti *sleep_quality* (10,61%) dan *safety* (6,67%), ketiganya berkontribusi kumulatif 62,88%, sebagaimana ditunjukkan pada Gambar 4. *Academic_performance* menempati posisi keempat (6,37%), sehingga empat variabel teratas mencapai 69,25% kontribusi kumulatif. Sebaliknya, variabel psikologis seperti *anxiety_level* (1,26%) dan *depression* (1,35%) berkontribusi rendah, menunjukkan bahwa tingkat kepentingan fitur pada model tidak selalu sejalan dengan asumsi teoritis umum.



Gambar 4. Variabel *feature importance (gain)* tertinggi

Interpretasi SHAP

Analisis SHAP menghasilkan urutan berbeda: *blood_pressure* (0,792397), *self_esteem* (0,389485), dan *sleep_quality* (0,370739) sebagai tiga variabel berkontribusi tertinggi. Perbedaan urutan antara *gain* dan SHAP mencerminkan perbedaan perspektif pengukuran, *gain* berbasis peningkatan kualitas split pohon, sedangkan SHAP berbasis kontribusi individual terhadap prediksi. Namun *blood_pressure* dan *sleep_quality* secara konsisten muncul sebagai variabel berkontribusi tinggi pada kedua metode.

5. KESIMPULAN

Algoritma *XGBoost* berhasil diterapkan untuk mengidentifikasi tingkat stres mahasiswa berdasarkan variabel psikologis, fisik, akademik, dan lingkungan pada *StressLevelDataset*, melalui pembangunan model baseline dan model hasil *hyperparameter tuning* menggunakan *Grid Search*, dievaluasi dengan *Stratified 10-Fold Cross-Validation*. Model *tuned* memperoleh performa lebih tinggi dibandingkan baseline pada seluruh metrik: akurasi 88,91%, presisi 88,95%, *recall* 88,92%, dan *F1-score* 88,93% (peningkatan berkisar 1,59-1,64 poin persentase). Meskipun demikian, uji McNemar (*p-value* = 1,000) menunjukkan bahwa peningkatan tersebut tidak signifikan secara statistik pada taraf 5%. Analisis *feature importance* berbasis *gain* mengidentifikasi *blood_pressure* (45,60%), *sleep_quality* (10,61%), dan *safety* (6,67%) sebagai tiga variabel paling berpengaruh, sementara analisis SHAP menempatkan *blood_pressure*, *self_esteem*, dan *sleep_quality* sebagai tiga teratas, dengan *blood_pressure* dan *sleep_quality* konsisten dominan pada kedua metode. Secara keseluruhan, penelitian ini membuktikan bahwa *XGBoost* mampu mengidentifikasi tingkat stres mahasiswa secara akurat dan stabil berbasis data, dengan seluruh metrik evaluasi di atas 88%.

DAFTAR PUSTAKA

- Breiman, Leo. 2001. "Random Forest." *Machine Learning* 45(1): 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Chen, Tianqi, and Carlos Guestrin. 2016. "XGBoost: A Scalable Tree Boosting System." *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 13–17: 785–94. <https://doi.org/10.1145/2939672.2939785>.
- Dietterich, Thomas G. 2000. "Ensemble Methods in Machine Learning." *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 1857 LNCS: 1–15. https://doi.org/10.1007/3-540-45014-9_1.
- Gedam, Shruti, Sandip Dutta, and Ritesh Jha. 2025. "Analyzing Mental Stress in Indian Students through Advanced Machine Learning and Wearable Technologies." *Scientific Reports* 15(1): 1–18. <https://doi.org/10.1038/s41598-025-06918-6>.
- James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2013. Springer *An Introduction to Statistical Learning: With Applications in R*. 1st ed.
- Kohavi, Ron, and George H. John. 1995. "Automatic Parameter Selection by Minimizing Estimated Error." *Proceedings of the 12th International Conference on Machine Learning, ICML 1995*: 304–12. <https://doi.org/10.1016/B978-1-55860-377-6.50045-1>.
- Lundberg, Scott M., and Su In Lee. 2017. "A Unified Approach to Interpreting Model Predictions." *Arxiv* 2017-Decem(Section 2): 4766–75. <https://doi.org/10.48550/arXiv.1705.07874>.
- McNemar, Quinn. 1947. "Note On The Sampling Error Of The Difference Between Correlated Proportions Or Percentages." *Psychometrika* 12(2): 153–57. <https://doi.org/10.1007/BF02295996>.
- Misra, Ranjita, and McKean Michelle. 2000. "College Students' Academic Stress and Its Relation to Their Anxiety, Time Management, and Leisure Satisfaction." *American Journal of Health Studies*: 41–51.
- Mitchell, Tom M. 1997. McGraw-Hill *Machine Learning*. 1st ed.
- Schröer, Christoph, Felix Kruse, and Jorge Marx Gómez. 2021. "A Systematic Literature Review on Applying CRISP-DM Process Model." *Procedia Computer Science* 181(2019): 526–34. <https://doi.org/10.1016/j.procs.2021.01.199>.
- Selye, Hans. 1976. *The Stress Of Life*. 2nd ed. McGraw-Hill.

- Wade, Corey. 2020. *Hands-On Gradient Boosting with XGBoost and Scikit-Learn*. 1st ed. Packt Publishing. <https://www.oreilly.com/library/view/hands-on-gradient-boosting/9781839218354/>.
- Yeler, Kadircan, and Selim Sürücü. 2025. "Classification of Student Stress Levels Using Machine Learning Methods." *International Journal of Advanced Natural Sciences and Engineering Researches*: 210–14. <https://as-proceeding.com/index.php/ijanser/article/view/2863>.
- Yuda, Marta Prima, Indah Mawarti, and Muthia Mutmainnah. 2023. "Description of Students' Academic Stress Levels in Completing Their Final Thesis Assignments at the Faculty of Medicine and Health Sciences Universitas Jambi." *Jurnal Pinang Masak* 55(393): 298–305. <https://doi.org/10.22437/jpima.v2i1.27092>.

