E. A. Putri dkk/ JIMI 11 (2) pp. 171-178

P-ISSN : 2549-7480
E-ISSN : 2549-6301

JURNAL ILMIAH INFORMATIKA
E-ISSN : 2549-6301
ISSN : 2549-7480
Faculty of Sains and Technology, Ibrahimy University
S4-Accredited : 105/E/KPT/2022
Journal page is available to : https://journal.ibrahimy.ac.id/index.php/JIMI

KOMPARASI ALGORITMA RANDOM FOREST, XGBOOST, DAN LIGHTGBM PADA KLASIFIKASI TINGKAT STRES SISWA BERDASARKAN FAKTOR AKADEMIK DAN LINGKUNGAN

Eva Andini Putri[1)], Avip Kurniawan[2)], Wiga Ariani[3)]

[1,2,3] *Program Studi Teknik Informatika, Fakultas Teknik, Universitas Krisnadwipayana, Jakarta*

email: [1] putrievaandini@gmail.com, [2] avip.kurniawan@unkris.ac.id, [3] wigaariani@unkris.ac.id

ARTICLE INFO

Article History:
Recieved : 04 Oktober 2026
Accepted : 22 Oktober 2026
Published : 12 November 2026

Keywords:
Stress Level Classification
Random Forest
XGBoost
LightGBM
Hyperparameter Tuning

IEEE style in citing this article:
E. A. Putri, A. Kurniawan, and W. Ariani, "Komparasi Algoritma Random Forest, XGBoost, dan LightGBM pada Klasifikasi Tingkat Stres Siswa Berdasarkan Faktor Akademik dan Lingkungan", JIMI, vol. x, no. x, pp. xx-xx, Bulan. Tahun.

Corresponding Author:
Eva Andini Putri
Teknik Informatika,
Universitas Krisnadwipayana

ABSTRACT

Stress among students can reduce learning motivation and academic achievement, yet early identification still relies largely on subjective questionnaires. This study aims to build a three-class student stress classification model (Low, Moderate, High) based on academic and environmental factors and to compare the performance of three ensemble algorithms, namely Random Forest, XGBoost, and LightGBM. The StressLevelDataset from Kaggle, containing 1,100 student records, was analysed following the CRISP-DM methodology; 994 records remained after duplicate removal and were divided into training and testing sets at an 80:20 ratio. Each algorithm was evaluated in a baseline scenario with default parameters and in a tuned scenario using grid search with three-fold cross-validation. Hyperparameter tuning increased the accuracy of XGBoost from 88.94% to 91.96% and that of LightGBM from 87.94% to 90.95%, whereas Random Forest decreased from 90.45% to 86.93% because its tree depth was restricted. The tuned XGBoost achieved the best overall performance, with a macro F1-score of 0.92, a ROC-AUC of 0.9928, and a prediction time below 0.01 second. Feature importance analysis indicates that academic performance, future career concerns, and extracurricular workload dominate the predictions, while environmental factors contribute little in this dataset. The tuned XGBoost model is proposed as the basis of a web-based early stress screening tool for school counselling units.

© 2026 Jurnal Ilmiah Informatika (Scientific Informatics Journal) with CC BY NC licence

P-ISSN : 2549-7480
E. A. Putri dkk/ JIMI 11 (2) pp. 171-178 E-ISSN : 2549-6301

1. PENDAHULUAN

Stres merupakan respons alami terhadap tuntutan yang melampaui kapasitas adaptasi seseorang. Pada siswa, tekanan psikologis yang berlebihan cenderung menurunkan motivasi belajar dan konsentrasi sehingga berdampak pada capaian akademik [1], [2]. Kondisi ini menjadi perhatian di Indonesia: survei I-NAMHS 2022 mencatat sekitar 15,5 juta remaja mengalami masalah kesehatan mental dalam satu tahun terakhir, tetapi hanya 2,6% di antaranya yang pernah memperoleh dukungan emosional atau konseling [3].

Pemicu stres siswa umumnya dikelompokkan menjadi faktor akademik, yaitu beban belajar, performa akademik, hubungan siswa dengan guru, kekhawatiran terhadap karier, dan keterlibatan dalam kegiatan ekstrakurikuler, serta faktor lingkungan, yaitu tingkat kebisingan, kondisi tempat tinggal, rasa aman, pemenuhan kebutuhan dasar, dukungan sosial, tekanan teman sebaya, dan perundungan [1]. Penilaian stres dengan kuesioner bersifat subjektif karena bergantung pada kejujuran dan kondisi psikologis responden saat mengisi, sehingga pendekatan berbasis data melalui *machine learning* dipandang dapat membantu deteksi dini yang lebih objektif dan terstandar [4].

Algoritma *ensemble* berbasis pohon keputusan, seperti *Random Forest*, XGBoost, dan LightGBM, banyak digunakan pada klasifikasi multikelas [5], [6]. Pada domain stres siswa, penelitian terdahulu telah membandingkan *Random Forest*, SVM, dan KNN [7]; *Random Forest*, *Decision Tree*, dan *Logistic Regression* [8]; serta *Random Forest*, SVM, dan XGBoost [9]. Yeler dan Sürücü [10] melaporkan bahwa *Logistic Regression* memperoleh akurasi tertinggi sebesar 88,1% dibandingkan XGBoost dan *Random Forest*, sedangkan De Filippis dan Foysal [4] memakai *Random Forest* untuk mengidentifikasi faktor yang paling berpengaruh. Studi lain mengembangkan kerangka *explainable machine learning* untuk klasifikasi stres akademik mahasiswa pada beberapa *dataset* [11], membandingkan *Random Forest*, XGBoost, dan CatBoost pada pengguna media sosial [12], serta memanfaatkan *explainable AI* untuk pemodelan stres mahasiswa [13].

Meskipun demikian, sepengetahuan penulis, studi yang membandingkan ketiga algoritma tersebut secara bersamaan pada data siswa dengan pemisahan eksplisit antara faktor akademik dan faktor lingkungan, serta menilai pengaruh *hyperparameter tuning* terhadap tiap algoritma, masih terbatas. Padahal parameter bawaan tidak selalu menghasilkan kinerja terbaik, dan optimasi parameter banyak dikaji untuk meningkatkan kinerja model, termasuk XGBoost dan LightGBM [14], [15].

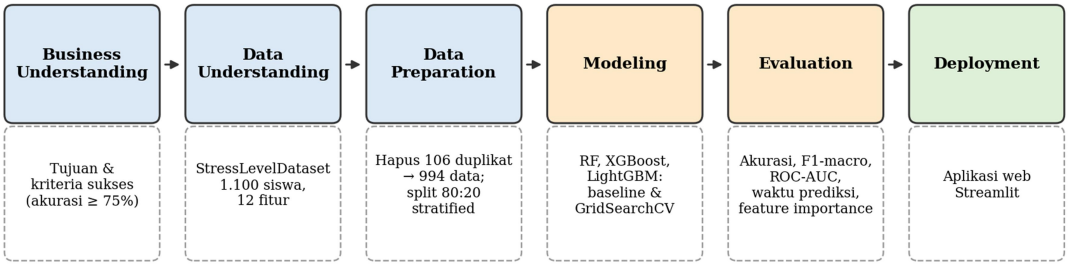
Penelitian ini bertujuan membangun model klasifikasi tingkat stres siswa tiga kelas (Rendah, Sedang, Tinggi) berdasarkan 12 fitur akademik dan lingkungan, membandingkan kinerja *Random Forest*, XGBoost, dan LightGBM pada skenario *baseline* dan *hyperparameter tuning*, serta mengidentifikasi faktor yang paling berpengaruh terhadap prediksi. Kontribusi penelitian ini adalah (1) komparasi enam model pada data uji dengan metrik akurasi, *F1-score macro*, ROC-AUC, dan waktu prediksi; (2) analisis *feature importance* antaralgoritma; dan (3) usulan implementasi model terbaik sebagai alat *screening* awal berbasis web untuk unit bimbingan konseling.

2. METODE PENELITIAN

Penelitian ini merupakan penelitian eksperimental kuantitatif yang mengikuti metodologi CRISP-DM [16] dengan enam

DOI : https://doi.org/10.35316/jimi.v11i2.171-178
© 2026 Jurnal Ilmiah Informatika with CC BY NC licence

tahap, yaitu *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*, sebagaimana ditunjukkan pada Gambar 1. Kriteria keberhasilan ditetapkan berupa akurasi minimal 75% pada data uji.



Gambar 1. Alur penelitian berdasarkan CRISP-DM

2.1 Dataset dan Variabel

Data yang digunakan adalah data sekunder *StressLevelDataset.csv* dari Kaggle yang memuat 1.100 siswa dari berbagai institusi pendidikan. Penelitian memakai 12 fitur yang dibagi menjadi 5 faktor akademik (*academic_performance*, *study_load*, *teacher_student_relationship*, *future_career_concerns*, *extracurricular_activities*) dan 7 faktor lingkungan (*noise_level*, *living_conditions*, *safety*, *basic_needs*, *social_support*, *peer_pressure*, *bullying*). Variabel target *stress_level* memiliki tiga kelas, yaitu Rendah (0), Sedang (1), dan Tinggi (2), dengan sebaran awal 373, 358, dan 369 siswa.

2.2 Persiapan Data

Pemeriksaan data tidak menemukan nilai hilang, tetapi terdapat 106 baris duplikat yang dihapus sehingga tersisa 994 baris. Data kemudian dipisah menjadi matriks fitur berukuran 994 × 12 dan vektor target, lalu dibagi dengan rasio 80:20 secara *stratified* (*random_state* = 42) sehingga diperoleh 795 data latih dan 199 data uji.

2.3 Pemodelan

Tiga algoritma, yaitu *Random Forest*, XGBoost, dan LightGBM, dilatih pada dua skenario sehingga terbentuk enam model. Pada skenario *baseline*, semua parameter memakai nilai bawaan pustaka. Pada skenario *tuned*, parameter dioptimasi menggunakan *GridSearchCV* dengan *3-fold cross-validation* dan akurasi sebagai metrik penilaian pada data latih. Ruang pencarian sengaja dibatasi pada pohon yang dangkal untuk menekan *overfitting*, dengan rincian pada Tabel 1.

P-ISSN : 2549-7480
E. A. Putri dkk/ JIMI 11 (2) pp. 171-178 E-ISSN : 2549-6301

Tabel 1. Ruang pencarian dan nilai terbaik *hyperparameter*

Algoritma	Hyperparameter	Nilai yang diuji	Nilai terbaik
Random Forest	*n_estimators*	50, 100	50
	max_depth	3, 5	3
	min_samples_split	2, 5	5
	min_samples_leaf	1, 2	1
XGBoost	*n_estimators*	50, 100	100
	learning_rate	0,01; 0,05	0,01
	max_depth	3, 4	3
LightGBM	*n_estimators*	50, 100	50
	learning_rate	0,01; 0,05	0,01
	max_depth	3, 4	4
	num_leaves	7, 15	7

2.4 Evaluasi

Kinerja model diukur pada data uji menggunakan *confusion matrix*, akurasi (Persamaan 1), presisi, *recall*, *F1-score* tiap kelas (Persamaan 2) beserta rata-ratanya (*F1-score macro*, Persamaan 3), ROC-AUC dengan pendekatan *one-vs-rest*, serta waktu prediksi.

$$\text{Akurasi} = \frac{\sum_{k=1}^{K} \text{TP}_k}{N} \quad (1)$$

$$F1_k = \frac{2 \times P_k \times R_k}{P_k + R_k} \quad (2)$$

$$\text{F1} - \text{macro} = \frac{1}{K} \sum_{k=1}^{K} F\,1_k \quad (3)$$

di mana TP adalah jumlah prediksi benar pada suatu kelas, N jumlah data uji, K jumlah kelas, serta P dan R masing-masing presisi dan *recall* kelas tersebut.

Kontribusi tiap fitur ditelaah melalui *feature importance* dengan *SelectFromModel* yang memakai nilai rata-rata (*mean threshold*) sebagai batas seleksi: fitur dengan skor di atas batas dinyatakan lolos, sedangkan fitur lain dieliminasi. Model terbaik kemudian diimplementasikan pada aplikasi web berbasis Streamlit.

3. HASIL DAN PEMBAHASAN

3.1 Hasil *Hyperparameter Tuning*

Hasil *grid search* pada Tabel 1 menunjukkan bahwa ketiga algoritma memilih konfigurasi yang konservatif. *Random Forest* memilih 50 pohon dengan kedalaman maksimum 3, XGBoost memilih *learning rate* 0,01 dengan kedalaman 3 dan 100 iterasi, sedangkan LightGBM memilih *learning rate* 0,01, kedalaman 4, dan 7 daun per pohon. Sebagian besar parameter jatuh pada nilai terkecil yang diuji, sehingga ruang pencarian yang lebih luas ke arah model yang lebih sederhana masih patut dicoba.

3.2 Perbandingan Kinerja Model

Tabel 2 merangkum kinerja keenam model pada 199 data uji. Seluruh model melampaui kriteria keberhasilan 75%. XGBoost *tuned* memberikan hasil terbaik dengan akurasi 91,96%, *F1-score macro* 0,92, dan ROC-AUC 0,9928, diikuti LightGBM *tuned* dengan akurasi 90,95%. Pada skenario *baseline*, *Random Forest* paling unggul dengan akurasi 90,45%, mengungguli XGBoost (88,94%) dan LightGBM (87,94%).

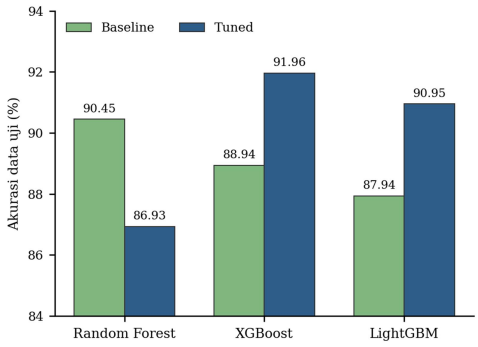
DOI : https://doi.org/10.35316/jimi.v11i2.171-178
© 2026 Jurnal Ilmiah Informatika with CC BY NC licence

P-ISSN: 2549-7480
E. A. Putri dkk/ JIMI 11 (2) pp. 171-178 E-ISSN: 2549-6301

Tabel 2. Kinerja enam model pada data uji (199 sampel)

Model	Skenario	Akurasi (%)	*F1-score macro*	ROC-AUC	Waktu prediksi (detik)
XGBoost	***Tuned***	**91,96**	**0,92**	**0,9928**	**0,0095**
LightGBM	*Tuned*	90,95	0,91	0,9909	0,0046
Random Forest	*Baseline*	90,45	0,90	0,9883	0,0138
XGBoost	*Baseline*	88,94	0,89	0,9883	0,0298
LightGBM	*Baseline*	87,94	0,88	0,9873	0,0127
Random Forest	*Tuned*	86,93	0,87	0,9870	0,0073

Pengaruh *tuning* berbeda antaralgoritma (Gambar 2). Akurasi XGBoost naik 3,02 poin persentase dan LightGBM naik 3,01 poin, sedangkan *Random Forest* turun 3,52 poin. Penurunan pada *Random Forest* tidak berarti bahwa *tuning* merugikan secara umum, melainkan mencerminkan ruang pencarian yang hanya menawarkan kedalaman 3 dan 5, sehingga model kehilangan kapasitas dibandingkan konfigurasi bawaan yang menumbuhkan pohon tanpa batas kedalaman.



Gambar 2. Perbandingan akurasi data uji pada skenario *baseline* dan *tuned*

Pada XGBoost *tuned*, presisi kelas Rendah (0,96) dan Sedang (0,97) tinggi, sedangkan *recall* kelas Tinggi mencapai 0,96. *Recall* kelas Tinggi yang tinggi penting untuk deteksi dini karena siswa bertekanan tinggi yang terlewat berisiko tidak mendapat penanganan. Konsekuensinya, presisi kelas Tinggi lebih rendah (0,85) dan *recall* kelas Rendah 0,84, artinya sebagian siswa berstres rendah diprediksi lebih tinggi; untuk *screening* awal, kesalahan ke arah ini masih dapat diterima.

Seluruh model *tuned* memprediksi dalam waktu di bawah 0,01 detik. LightGBM *tuned* paling cepat (0,0046 detik), sedangkan XGBoost *baseline* paling lambat (0,0298 detik). Nilai ROC-AUC keenam model berada pada rentang 0,987–0,993, sehingga kemampuan membedakan kelas relatif setara dan perbedaan utama terletak pada akurasi serta keseimbangan *recall*.

Perlu dicatat bahwa pada 199 data uji, satu sampel setara 0,5 poin persentase. Selisih XGBoost *tuned* (183 prediksi benar) dengan LightGBM *tuned* (181 prediksi benar) hanya dua sampel, sehingga keunggulan XGBoost belum dapat dinyatakan signifikan secara statistik. Pengujian berulang atau uji signifikansi seperti pada [11] diperlukan untuk memastikannya.

3.3 Faktor yang Berpengaruh

Pada XGBoost *tuned*, tiga fitur akademik lolos seleksi, yaitu *academic_performance*, *future_career_concerns*, dan *extracurricular_activities*, sedangkan sembilan fitur lain, termasuk seluruh fitur lingkungan, berkontribusi di bawah 0,05. *Random Forest tuned* mempertahankan tujuh fitur (*academic_performance*, *extracurricular_activities*, *safety*, *basic_needs*, *social_support*, *peer_pressure*, dan *bullying*),

P-ISSN : 2549-7480
E. A. Putri dkk/ JIMI 11 (2) pp. 171-178 E-ISSN : 2549-6301

sedangkan LightGBM *tuned* hanya tiga fitur (*extracurricular_activities*, *social_support*, dan *bullying*).

Satu-satunya fitur yang lolos pada ketiga model adalah *extracurricular_activities*, sehingga beban kegiatan ekstrakurikuler tampak sebagai sinyal yang paling konsisten. Perbedaan daftar fitur antaralgoritma kemungkinan menandakan bahwa sebagian fitur saling berkorelasi dan dapat saling menggantikan, misalnya LightGBM *tuned* menempatkan *academic_performance* di bawah ambang seleksi tetapi tetap mencapai akurasi 90,95%. Dugaan ini perlu diuji lebih lanjut dengan analisis korelasi atau *ablation*. Peran performa akademik sejalan dengan temuan [11] pada *dataset* yang sama. Hasil ini menunjukkan bahwa pada data ini tekanan akademik lebih dominan daripada faktor lingkungan, dan perlu dipahami sebagai asosiasi prediktif, bukan hubungan sebab-akibat.

3.4 Implementasi

Model XGBoost *tuned* diimplementasikan pada aplikasi web Streamlit. Pengguna memasukkan skor tiap fitur melalui *slider*, kemudian sistem menampilkan kategori stres (Rendah, Sedang, atau Tinggi) berkode warna beserta probabilitas tiap kelas secara *real-time*. Aplikasi ini dapat dipakai guru bimbingan konseling sebagai alat *screening* awal secara berkala.

4. UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Program Studi Teknik Informatika, Fakultas Teknik, Universitas Krisnadwipayana atas dukungan selama penelitian ini.

5. KESIMPULAN

Penelitian ini berhasil membangun model klasifikasi tiga kelas tingkat stres siswa berdasarkan 12 fitur akademik dan lingkungan menggunakan *Random Forest*, XGBoost, dan LightGBM. XGBoost dengan *hyperparameter tuning* menjadi model terbaik dengan akurasi 91,96%, *F1-score macro* 0,92, ROC-AUC 0,9928, dan waktu prediksi 0,0095 detik. *Hyperparameter tuning* meningkatkan akurasi XGBoost dan LightGBM sekitar 3 poin persentase, tetapi menurunkan akurasi *Random Forest* sebesar 3,52 poin pada ruang pencarian yang dibatasi, sehingga dampak *tuning* bergantung pada algoritma dan rancangan ruang pencariannya. Fitur yang paling berpengaruh pada model terbaik adalah *academic_performance*, *future_career_concerns*, dan *extracurricular_activities*, sedangkan faktor lingkungan berkontribusi kecil pada *dataset* ini.

Keterbatasan penelitian ini meliputi pemakaian satu *dataset* sekunder yang kriteria pelabelannya tidak dijelaskan oleh sumbernya, jumlah data uji yang kecil (199), satu kali pembagian data, dan ruang pencarian parameter yang sempit. Penelitian lanjutan disarankan memakai validasi silang berulang dengan uji signifikansi, ruang pencarian yang lebih luas termasuk *Random Forest* tanpa batas kedalaman, data primer lintas sekolah dan jenjang, serta pembandingan dengan CatBoost atau *deep learning*. Secara praktis, model dapat diintegrasikan ke unit bimbingan konseling sebagai alat *screening* awal, dan aplikasinya dapat dikembangkan ke platform *mobile* atau sistem informasi akademik sekolah.

6. REFERENSI

[1] A. Rahmawati, A. Agustini, M. R. Ramadhan, N. M. Zulfaa, and H. Nisrina, “Gangguan kesehatan mental terhadap performa akademik mahasiswa tingkat awal,” *Educate: Journal of Education and Learning*, vol. 3, no. 1, pp. 1–15, 2025, doi: 10.61994/educate.v3i1.329.

DOI : https://doi.org/10.35316/jimi.v11i2.171-178
© 2026 Jurnal Ilmiah Informatika with CC BY NC licence

[2] B. Insantuan, H. Sihotang, E. Murniarti, and J. Buaton, "The impact of academic stress on students' mental health: A qualitative study of student experiences," *Journal La Sociale*, vol. 7, no. 2, pp. 624–636, 2026, doi: 10.37899/journal-la-sociale.v7i2.2897.

[3] M. Rahma, M. Fikry, and Y. Afrillia, "Prediksi kesehatan mental remaja berdasarkan faktor lingkungan sekolah menggunakan machine learning," *Jurnal Informatika: Jurnal Pengembangan IT*, vol. 10, no. 2, pp. 382–390, 2025, doi: 10.30591/jpit.v10i2.8556.

[4] R. De Filippis and A. Al Foysal, "Comprehensive analysis of stress factors affecting students: a machine learning approach," *Discover Artificial Intelligence*, vol. 4, no. 1, 2024, doi: 10.1007/s44163-024-00169-6.

[5] M. Erkamim, S. Suswadi, M. Z. Subarkah, and E. Widarti, "Komparasi algoritme Random Forest dan XGBoosting dalam klasifikasi performa UMKM," *Jurnal Sistem Informasi Bisnis*, vol. 13, no. 2, pp. 127–134, 2023, doi: 10.21456/vol13iss2pp127-134.

[6] R. Saha, S. Saif, T. Khan, and A. Alabdultif, "A critical systematic review of ensemble learning methods for predictive healthcare," *Discover Computing*, vol. 29, no. 1, 2026, doi: 10.1007/s10791-026-10005-3.

[7] P. Manjunath, T. S, P. Shreya, V. Ashok, and S. Sultana, "Predictive analysis of student stress level using machine learning," *International Journal of Engineering Research and Technology (IJERT)*, vol. 9, no. 12, pp. 76–80, 2021, doi: 10.17577/IJERTCONV9IS12019.

[8] M. Anbarasi, S. Thakkilapati, V. R. Vas, and S. V. B. Viswath, "Predictive analysis of student stress level using machine learning," *International Journal of Engineering Research and Technology*, vol. 12, no. 5, 2023.

[9] S. S. Shahapur *et al.*, "Decoding minds: Estimation of stress level in students using machine learning," *Indian Journal of Science and Technology*, vol. 17, pp. 2002–2012, 2024, doi: 10.17485/IJST/v17i19.2951.

[10] K. Yeler and S. Sürücü, "Classification of student stress levels using machine learning methods," *International Journal of Advanced Natural Sciences and Engineering Researches*, vol. 10, no. 10, 2025.

[11] A. Omarbekova *et al.*, "An explainable machine learning framework for academic stress classification among university students: A comparative multi-dataset study with ablation analysis and statistical significance testing," *Frontiers in Computer Science*, vol. 8, Art. no. 1886274, 2026, doi: 10.3389/fcomp.2026.1886274.

[12] D. N. Fahria, K. D. Tania, and R. D. Kurnia, "Analisis komparatif algoritma Random Forest, XGBoost, dan CatBoost untuk klasifikasi tingkat stres pengguna media sosial," *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 11, no. 1, pp. 1843–1853, 2026, doi: 10.36341/rabit.v11i1.7449.

[13] R. Tariq, M. G. Orozco-del-Castillo, M. T. Zamir, M. S. Ramírez-Montoya, and T. Wilberforce, "Explainable artificial intelligence for predictive modeling of student stress in higher education," *Scientific*

DOI : https://doi.org/10.35316/jimi.v11i2.171-178
© 2026 Jurnal Ilmiah Informatika with CC BY NC licence

P-ISSN : 2549-7480
E. A. Putri dkk/ JIMI 11 (2) pp. 171-178 E-ISSN : 2549-6301

Reports, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-22171-3.

[14] M. Imani and H. R. Arabnia, "Hyperparameter optimization and combined data sampling techniques in machine learning for customer churn prediction: A comparative analysis," *Technologies*, vol. 11, no. 6, 2023, doi: 10.3390/technologies11060167.

[15] D. Demirtürk, Ö. Mintemur, and A. Arslan, "Optimizing LightGBM and XGBoost algorithms for estimating compressive strength in high-performance concrete," *Arabian Journal for Science and Engineering*, 2025, doi: 10.1007/s13369-025-10217-7.

[16] A. E. Putri, M. Z. Yumarlin, and J. E. Bororing, "Implementasi K-Means clustering dan model CRISP-DM untuk pengelompokan daerah rawan tindak pidana narkoba di DIY," *Jurnal Informatika Teknologi dan Sains (Jinteks)*, vol. 7, no. 2, pp. 606–615, 2025.

DOI : https://doi.org/10.35316/jimi.v11i2.171-178
© 2026 Jurnal Ilmiah Informatika with CC BY NC licence