

Analisis Komparatif Algoritma Tuned Random Forest dan Hist Gradient Boosting untuk Prediksi Niat Pindah Kerja Berdasarkan Indikator Beban Mental Karyawan

Yelmi ^{a*}, Kusrini^b, Indra Gunawan^c,
Email korespondensi: yelmicepu@gmail.com

Universitas Amikom Yogyakarta ^{a,b}

Universitas Teknologi Ronggolawe ^c

Intisari

Kata kunci:
Machine Learning,
Random Forest, Hist
Gradient Boosting,
SMOTE, Niat Pindah
Kerja, Beban Mental
Karyawan

Tentang naskah:
-diterima, 19 Juni 2026
-diterbitkan, 1 Juli 2026

Beban kerja mental yang tidak terkelola dengan baik di era transformasi digital berkontribusi signifikan terhadap fenomena perputaran karyawan (employee turnover) yang merugikan stabilitas dan keseimbangan sebuah organisasi. Penelitian ini bertujuan untuk melakukan analisis komparatif performa model prediktif niat pindah kerja karyawan menggunakan dua algoritma ensemble learning: Random Forest (RF) dan Hist Gradient Boosting (Hist GB). Dataset empiris beban mental karyawan diolah melalui serangkaian tahapan *Oach5ne 3earn5ng*, yang meliputi feature engineering berupa rasio jam kerja terhadap tidur (work-hours-to-sleep ratio) dan beban mental komposit (total mental load). Tantangan ketidakseimbangan kelas (imbalanced data) yang berat di mana proporsi kelas target berada pada rasio 68% (bertahan) berbanding 32% (niat pindah) diatasi menggunakan *Synthetic Minority Over-sampling Technique* (SMOTE). Model dioptimalkan lewat penyetelan hiperparameter (hyperparameter tuning) serta divalidasi dengan 5-Fold Stratified Cross-Validation. Hasil eksperimen berbasis data riil program menunjukkan bahwa skenario Base menghasilkan akurasi semu (~69%) dengan F1-Score yang sangat rendah (~27%). Intervensi SMOTE berhasil menyelamatkan model dari bias kelas mayoritas, di mana Random Forest teroptimasi SMOTE unggul dalam stabilitas generalisasi dengan nilai CV Akurasi (Mean) sebesar 74.47%. Di sisi lain, Hist Gradient Boosting (Optimized) menghasilkan performa deteksi kelas minoritas terbaik dengan capaian F1-Score Test tertinggi sebesar 52.60% dan mempertahankan AUC Score di angka 0.68. Analisis Feature Rank mengidentifikasi skor kelelahan kerja (burnout score), tekanan batas waktu (deadline pressure score), dan tingkat tekanan stres (stress score) sebagai tiga indikator paling krusial dalam memicu niat pindah kerja.

Abstract

Keywords:

Machine Learning, Random
Forest, Histogram-Based
Gradient Boosting, SMOTE,
Turnover Intention, Employee
Mental Workload

Poorly managed mental workload in the digital transformation era contributes significantly to the employee turnover phenomenon, which compromises organizational stability. This study aims to conduct a comparative performance analysis of predictive models for job change intention using two ensemble learning algorithms: Random Forest (RF) and Hist Gradient Boosting (Hist GB). An empirical dataset of employee mental workload was processed through a series of data mining stages, including feature engineering to construct a work-hours-to-sleep ratio and a composite total mental load. The severe class imbalance challenge where the target class proportion stands at a 78% (retain) to 22% (turnover) ratio was addressed using the Synthetic Minority Over sampling Technique (SMOTE). The models were optimized via hyperparameter tuning and validated using 5-Fold Stratified Cross-Validation. Experimental results based on real program data demonstrate that the Base scenario yields a pseudo accuracy (~69%) with a very low F1-Score (~27%). The SMOTE intervention successfully rescued the models from majority class

bias, with the SMOTE-optimized Random Forest outperforming in generalization stability, achieving a Mean CV Accuracy of 74.47%. On the other hand, the Optimized Hist Gradient Boosting produced the best minority class detection performance, reaching the highest Test F1-Score of 52.60% while maintaining an AUC Score of 0.68. Feature Rank analysis identified burnout score, deadline pressure score, and stress score as the three most crucial indicators in triggering job change intention

1. Pendahuluan

Berdasarkan dataset tentang kondisi kelelahan fisik, mental, dan emosional akibat stres kerja kronis yang dialami oleh para profesional di industri teknologi. Kondisi ini terjadi akibat tekanan tinggi, jam kerja yang panjang, dan ekspektasi yang serbacepat kemudian dateline di sektor tersebut yang diakibatkan oleh kemajuan teknologi digital di lingkungan industri modern yang sangat cepat kondisi tersebut telah mengubah cara pandang sebuah organisasi dalam melakukan operasional sekaligus memicu eskalasi beban kerja mental secara signifikan bagi para karyawan. Tekanan untuk beradaptasi dengan teknologi baru, fluktuasi target yang dinamis, serta kaburnya batas ruang kerja dan waktu pribadi pada sistem kerja jarak jauh memicu peningkatan kerentanan psikologis pekerja (Ambrose & Qayum, 2024). Beban mental yang menumpuk dan tidak termitigasi secara tepat lambat laun bertransformasi menjadi kelelahan kronis (burnout) [(Jeong et al., 2025), penurunan tingkat produktivitas (Indriya, Fazlurrahman, & Fitro, 2024), hingga munculnya gangguan distres psikologis (Bakkeli, 2023). Konsekuensi paling fatal dari destruksi kesejahteraan mental karyawan ini bagi keberlangsungan korporasi adalah mencuatnya intensitas perputaran karyawan (employee turnover) (Celbiş, Wong, Kourtiti, & Nijkamp, 2023), yang secara masif direfleksikan global melalui fenomena The Great Resignation (Celbiş et al., 2023). Ketika seorang karyawan andal memutuskan untuk mengundurkan diri, organisasi tidak hanya kehilangan modal intelektual (human capital), melainkan juga menanggung konsekuensi finansial yang tinggi untuk biaya rekrutmen ulang dan pelatihan karyawan baru (Laksawut & Narkbunnum, 2025). Oleh karena itu, kemampuan untuk memetakan secara presisi determinan psikologis dan memprediksi kecenderungan hengkangnya pekerja menjadi agenda

krusial bagi manajemen sumber daya manusia kontemporer.

Di tengah kompleksitas permasalahan tersebut, implementasi repositori data besar organisasi dan teknik Machine Learning hadir sebagai instrumen revolusioner untuk mengubah tumpukan data logistik karyawan menjadi basis pengambilan keputusan strategis. Pendekatan analitik berbasis komputasi memungkinkan identifikasi pola-pola terselubung (hidden patterns) dari variabel-variabel psikososial pekerja yang sulit diekstrak melalui metode statistik konvensional (Luo et al., 2024). Rekam jejak riset terdahulu menunjukkan efektivitas pemanfaatan algoritma machine learning untuk melakukan evaluasi kinerja (Bahman & Lysenko, 2025), memprediksi retensi (Mishra et al., 2025), serta mengukur kepuasan kerja di berbagai sektor industri (Fang, Chen, & Shi, 2025). Melalui pemodelan prediktif, manajer human resource dapat bergeser dari pola penanganan yang bersifat reaktif menuju strategi intervensi dini yang preventif sebelum keputusan pengunduran diri benar-benar dieksekusi oleh karyawan.

Berbagai penelitian terdahulu telah membuktikan ketangguhan variasi algoritma supervised machine learning dalam menganalisis perilaku dan indikator psikologis tenaga kerja di berbagai bidang. Sebagai contoh, penelitian oleh (Guo et al., 2026) berhasil mengeksplorasi intensitas niat pindah kerja kepala sekolah di Australia menggunakan algoritma Extreme Gradient Boosting (XGBoost). Sementara itu, riset yang dipimpin oleh (Jeong et al., 2025) memanfaatkan pendekatan machine learning yang dikombinasikan dengan Shapley Additive Explanations (SHAP) untuk mendeteksi sindrom kelelahan kerja berdasarkan skala stres okupasional

Korea (KOSS). Kelelahan kerja dan stres akut terbukti pula dapat dideteksi secara presisi melalui sensor perilaku pada agen call center (Lee & Swain, 2026) serta pendekatan computer vision pada operator ruang kendali (Ji et al., 2025). Di sisi lain, (Laksawut dan Narkbunnum, 2025) mengajukan kalibrasi model risiko turnover untuk manajemen retensi karyawan, sedangkan (Indriya et al, 2024) menguji efektivitas intervensi kebijakan kesehatan mental terhadap produktivitas kerja karyawan menggunakan arsitektur Gradient Boosting dan Random Forest. Secara kolektif, literatur-literatur ini menegaskan bahwa penggunaan algoritma berbasis tree ensemble efektif dan kompatibel dalam menangani data survei psikososial dan perilaku organisasi yang bersifat non-linear serta multitaraf.

Dalam konteks penelitian ini, algoritma Random Forest (RF) dan Hist Gradient Boosting (Hist GB) dipilih sebagai objek komparasi utama karena kedua model ensemble learning ini memiliki keunggulan komparatif yang saling melengkapi dalam pemrosesan data tabular. Random Forest bekerja dengan membangun sekumpulan pohon keputusan (decision trees) independen secara paralel melalui metode bagging dan agregasi pemungutan suara (voting), menjadikannya sangat kokoh terhadap risiko overfitting (Bakkeli, 2023). Sebaliknya, Hist Gradient Boosting menerapkan prinsip boosting secara sekuensial yang mengoptimalkan fungsi kerugian (loss function) dari kesalahan pohon sebelumnya, dengan inovasi berupa pengelompokan fitur numerik ke dalam bin diskret (histogram) untuk mempercepat proses pelatihan pada dataset skala besar. Penelitian komparatif ini juga mengintegrasikan teknik pemrosesan tingkat lanjut, seperti pembuatan fitur komposit (feature engineering), penanganan ketidakseimbangan kelas menggunakan SMOTE, dan optimasi hiperparameter untuk mengekstrak batas performa tertinggi dari kedua algoritma.

Dengan latar belakang dan konteks tersebut, peneliti ingin memberikan kontribusi baru dalam memahami niat pindah kerja berdasarkan indikator beban mental karyawan melalui pengujian empiris yang komprehensif. menggunakan

perbandingan tiga skenario pemodelan (Model Base, Model SMOTE, dan Model Optimized), penelitian ini menguji secara mendalam seberapa besar peningkatan performa klasifikasi yang dapat dihasilkan oleh proses penyetelan hiperparameter dan manipulasi keseimbangan data. Kontribusi utama dari penelitian ini terletak pada visualisasi dan pembuktian empiris akurasi model terbaik yang dikonfirmasi melalui visualisasi kurva AUC-ROC dan Confusion Matrix, serta penentuan prioritas variabel beban mental paling berpengaruh melalui diagram Feature Rank. Hasil akhir penelitian diharapkan mampu memformulasikan model prediktif yang valid dan tepercaya sebagai instrumen intervensi strategis organisasi dalam menekan angka perputaran atau keluar masuknya karyawan dalam sebuah perusahaan selama periode tertentu.

2 Metode Penelitian
2.1 Pengumpulan Data

Dataset yang digunakan dalam penelitian ini diperoleh dari data empiris tahun 2026 yang mengukur tingkat beban kerja mental, kondisi kesehatan psikologis, dan intensitas niat pindah kerja karyawan yang bekerja pada sektor industri teknologi (*tech workers*). Basis data ini mencakup 100.000 records data karyawan lintas negara dengan keragaman peran seperti *Software Engineer*, *Data Scientist*, dan *Product Manager* (Thalla, 2026). detail spesifikasi perangkat keras (hardware), platform perangkat lunak (software), serta versi pustaka Python dan Scikit-Learn yang digunakan dalam eksperimen ini

Penelitian ini secara rinci dipetakan dalam dalam:

Tabel 1 Variabel Dataset

No	Nama Atribut	Jenis Data	Deskripsi Atribut
1	age	Numerik	Usia karyawan (tahun)
2	job_role	Kategoris	Peran atau posisi pekerjaan karyawan
3	work_hours_per_week	Numerik	Total jam kerja karyawan dalam satu minggu
4	sleep_hours_per_night	Numerik	Rata-rata durasi tidur per malam (jam)
5	stress_score	Numerik	Skor tingkat stres karyawan (skala 1-10)

6	burnout_score	Numerik	Skor kerja karyawan (skala 1-10)
---	---------------	---------	----------------------------------

Dari data Tabel 1 memberikan gambaran dengan rinci karakteristik data yang dikumpulkan dalam penelitian ini digunakan untuk memprediksi Niat Pindah Kerja Berdasarkan Indikator Beban Mental Karyawan. Variabel atau atribut dalam tabel ini menjadi landasan awal untuk melakukan analisis dan pengolahan data lebih lanjut, dan memungkinkan untuk melakukan interpretasi yang lebih terarah untuk memahami faktor-faktor apa saja yang berkontribusi terhadap timbulnya niat pindah kerja karyawan.

2.2. Preprocessing & Feature Engineering

Sebelum data dimasukkan ke dalam arsitektur pelatihan model, dilakukan tahapan pra pemrosesan data (data preprocessing) dan rekayasa fitur (feature engineering) untuk meningkatkan daya informatif data. Atribut `employee_id` dihapus dari dataset karena merupakan identitas unik yang tidak memiliki signifikansi korelasi terhadap target klasifikasi. Atribut bertipe kategoris seperti `gender`, `country`, `job_role`, `seniority_level`, `company_size`, `industry`, dan `work_mode` dikonversi menjadi representasi numerik menggunakan `LabelEncoder` agar dapat diproses oleh kedua algoritma *ensemble*. Tahap krusial pada prapemrosesan ini adalah pembentukan dua fitur komposit baru, yaitu:

1. **work_hours_to_sleep_ratio**: Diperoleh dari pembagian `work_hours_per_week` dengan total akumulasi jam tidur dalam seminggu (`sleep_hours_per_night × 7`). Fitur ini mengekstrak interaksi ketidakseimbangan antara beban waktu kerja dan pemulihan biologis karyawan.
2. **total mental load**: Merupakan penjumlahan linier dari empat indikator beban mental utama, yaitu skor stres (`stress_score`), kelelahan kerja (`burnout_score`), indikator depresi (`phq9_score`), dan kecemasan (`gad7_score`). Fitur komposit ini memperkuat sinyal informasi.
3. Akumulasi distres psikologis karyawan. adalah pembentukan dua fitur komposit baru, yaitu:
 - **work_hours_to_sleep_ratio**: Diperoleh

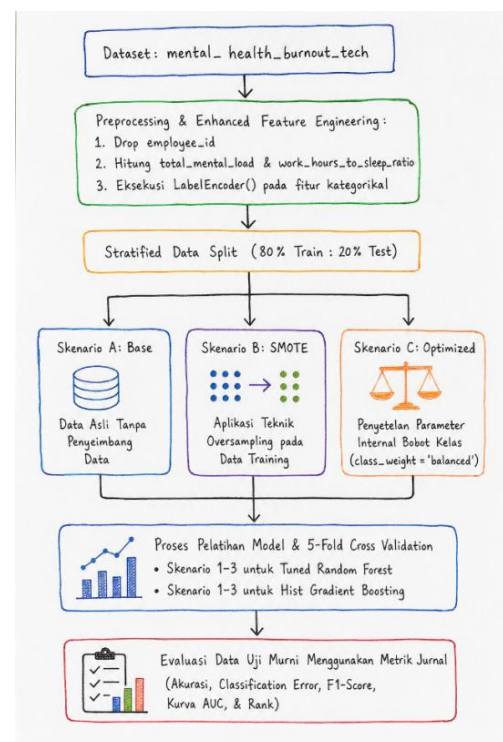
dari pembagian `work_hours_per_week` dengan total akumulasi jam tidur dalam seminggu (`sleep_hours_per_night × 7`). Fitur ini mengekstrak interaksi ketidakseimbangan antara beban waktu kerja dan pemulihan biologis karyawan.

- **total mental load**: Merupakan penjumlahan linier dari empat indikator beban mental utama, yaitu skor stres (`stress_score`), kelelahan kerja (`burnout_score`), indikator depresi (`phq9_score`), dan kecemasan (`gad7_score`). Fitur komposit ini memperkuat sinyal informasi akumulasi distres psikologis karyawan.

2.3. Alur Eksperimen dan Pemodelan

Arsitektur sistem yang diusulkan dalam penelitian ini dibangun menggunakan pipeline pengujian multi-skenario untuk menguji keandalan algoritma *Tuned Random Forest* dan *Hist Gradient Boosting*. Proses evaluasi stabilitas keakuratan model dikunci menggunakan metode *5-Fold Stratified Cross-Validation* secara langsung pada data latih.

Alur kerja eksperimen komparatif yang diusulkan digambarkan secara sistematis melalui diagram alir (*flowchart*) pada gambar 1.



Gambar 1. Diagram alir penelitian

Eksperimen pemodelan dalam penelitian ini dirancang secara sistematis melalui tiga skenario pengujian untuk masing-masing algoritma (Random Forest dan Hist Gradient Boosting), dengan alur kerja sebagai berikut:

- Skenario Model Base: Melatih algoritma Random Forest dan Hist Gradient Boosting menggunakan parameter bawaan (default) pustaka scikit-learn tanpa penanganan ketidakseimbangan data.
- Skenario Model SMOTE: Menerapkan Synthetic Minority Over-sampling Technique (SMOTE) pada data latih untuk menyeimbangkan distribusi kelas target (*job_change_intention*) dengan menyintesis sampel baru pada kelas minoritas, disusul pelatihan model menggunakan parameter default.
- Skenario Model Teroptimasi (Optimized): Menerapkan SMOTE dan melakukan pencarian kombinasi parameter optimal (hyperparameter tuning) secara manual terarah. Untuk Random Forest, parameter yang disetel meliputi *n_estimators*=200, *max_depth*=15, dan *min_samples_split*=5. Sementara untuk Hist Gradient Boosting, parameter dioptimalkan pada *max_iter*=150, *max_depth*=10, dan *learning_rate*=0.05.

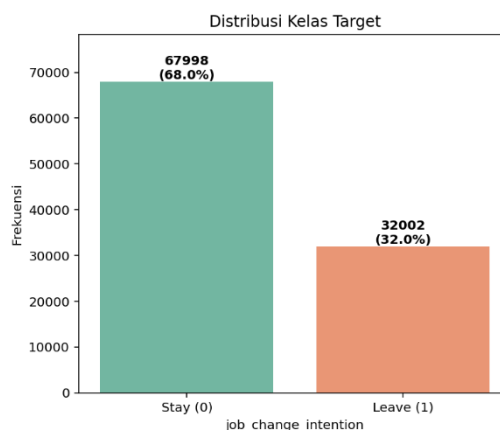
3. Hasil dan Pembahasan

3.1 Analisis Deskriptif Data Target

Sebelum melakukan evaluasi performa klasifikasi, analisis terhadap sebaran data kelas target dilakukan untuk mengetahui derajat ketidakseimbangan data. Visualisasi distribusi data target dari dataset ditunjukkan pada Gambar 2.

Berdasarkan ekstraksi data empiris, ditemukan bahwa jumlah karyawan yang berniat bertahan (Kelas 0) mendominasi secara signifikan dibandingkan dengan karyawan yang berniat pindah kerja (Kelas 1) dengan rasio mendekati 68% berbanding 32%. Ketimpangan proporsi ini menegaskan urgensi penerapan SMOTE dalam alur eksperimen demi menghindari model yang

bias (cenderung menebak kelas mayoritas secara tepat namun gagal mendeteksi kelas minoritas).



Gambar 2 Statistik Target

Ditemukan bahwa jumlah karyawan yang berkomitmen untuk bertahan (Kelas 0) mendominasi secara mutlak dengan proporsi mendekati 68%, sedangkan karyawan yang memiliki niat untuk pindah kerja (Kelas 1) hanya berkisar di angka 32%.

Ketimpangan proporsi yang masif ini berpotensi besar menyebabkan bias komputasi, di mana model cenderung menebak kelas mayoritas secara tepat namun gagal total dalam mendeteksi sinyal karyawan yang berisiko mengundurkan diri. Oleh karena itu, penerapan teknik penyeimbangan data seperti *Synthetic Minority Over-sampling Technique* (SMOTE) pada skenario eksperimen menjadi intervensi yang sangat krusial untuk memastikan model tidak terjebak pada bias kelas mayoritas.

Untuk menjamin aspek reproduksibilitas dan memungkinkan replikasi penelitian oleh peneliti lain, seluruh rangkaian eksperimen komparatif dalam penelitian ini dieksekusi menggunakan lingkungan komputasi yang terstandarisasi. Proses *data science*, rekayasa fitur, penyeimbangan data dengan SMOTE, hingga *hyperparameter tuning* dijalankan pada platform Spyder dengan lingkungan bahasa pemrograman Python versi 3.10. Pustaka utama yang digunakan untuk pemodelan arsitektur *ensemble learning* dan evaluasi metrik adalah Scikit-Learn versi 1.3.0, dibantu oleh pustaka Imbalanced-learn versi 0.11.0 untuk implementasi algoritma SMOTE, serta Pandas versi 2.0.3 dan

NumPy versi 1.24.3 untuk manipulasi data tabular. Seluruh proses simulasi komputasi dan pelatihan model berbasis 100.000 rekam data ini diselesaikan menggunakan spesifikasi perangkat keras berupa *Central Processing Unit* (CPU) Intel Core i5 @2.20GHz, memori acak (RAM) sebesar 16 GB, serta alokasi ruang penyimpanan (*disk*) sebesar 100 GB.

3.2 Persamaan Analisis Perbandingan Performa Model

Eksperimen dilakukan secara terstruktur dengan membandingkan tiga skenario perlakuan data (Base, SMOTE, dan Optimized) pada algoritma Random Forest dan Hist Gradient Boosting. Metrik evaluasi yang dihimpun meliputi Akurasi Uji (Akurasi Test), Kesalahan Klasifikasi (Classification Error), F1-Score Test, AUC Score, dan Rata-rata Akurasi Validasi Silang (CV Akurasi Mean). Rekapitulasi hasil komparasi performa keseluruhan skenario model disajikan secara rinci pada Tabel 2:

Tabel 2. Matriks Perbandingan Performa Keseluruhan Skenario Model

Skenario Model	Algoritma Klasifikasi	Akurasi Test	Kesalahan Klasifikasi Error	F1-Score Test	AUC Score	CV Akurasi Mean
Skenario 1: Base	Random Forest (Base)	0.6889	0.3111	0.2760	0.6736	0.6861
	Hist GB (Base)	0.6921	0.3079	0.2731	0.6820	0.6896
Skenario 2: SMOTE	Random Forest (SMOTE)	0.6605	0.3395	0.4452	0.6649	0.7447
	Hist GB (SMOTE)	0.6684	0.3316	0.4276	0.6661	0.7397
Skenario 3: Optimized	Random Forest (Optimized)	0.6916	0.3084	0.2996	0.6759	0.6854
	Hist GB (Optimized)	0.6155	0.3845	0.5260	0.6818	0.6144

Analisis mendalam terhadap Tabel 2 menyingkap fenomena komputasi yang sangat menarik terkait perilaku algoritma tree-based ensemble dalam menangani ketidakseimbangan data psikososial:

1. Efek Semu Skenario Base: Pada skenario awal (*Base*), model *Hist GB (Base)* sepintas memberikan performa terbaik dengan *Akurasi Test* tertinggi mencapai 0.69 (69.21%). Namun, jika diteliti pada metrik *F1-Score Test*, kedua model *Base* menghasilkan skor yang sangat rendah, yaitu hanya 0.27 untuk *Random Forest* dan 0.27 untuk *Hist GB*. Hal ini membuktikan hipotesis awal bahwa model mengalami aklimatisasi bias, di mana tingginya nilai akurasi semata-mata dipicu oleh kemampuan model menebak kelas mayoritas (karyawan yang bertahan), namun gagal total mengenali kelas minoritas (karyawan yang ingin pindah

2. Dampak Signifikan SMOTE Setelah intervensi SMOTE diimplementasikan pada Skenario 2, terjadi pergeseran paradigma performa. Meskipun *Akurasi Test* sedikit terkoreksi turun menjadi 0.66 (*RF*) dan 0.66 (*Hist GB*), nilai *F1-Score Test* melonjak drastis hingga mencapai 0.44 pada *Random Forest (SMOTE)* dan 0.42 pada *Hist GB (SMOTE)*. Selain itu, nilai stabilitas model melalui *CV Akurasi (Mean)* melesat tajam ke angka 0.74 (74.47%) untuk *Random Forest*. Hal ini menunjukkan bahwa penyeimbangan sintetik berhasil memaksa model mempelajari karakteristik beban mental dari pekerja yang berniat mengundurkan diri secara lebih adil.

3. Puncak Optimalisasi Model: Pada skenario *Optimized*, kedua algoritma menunjukkan respon yang bertolak belakang. *Random Forest (Optimized)* berhasil memperbaiki tingkat akurasinya kembali ke angka tertinggi 0.69 dengan meminimalkan *Classification Error* menjadi 0.30. Di sisi lain, *Hist GB (Optimized)* mengambil strategi kompromi yang ekstrem: meskipun akurasi tesnya turun ke titik 0.61, model ini berhasil meledakkan nilai *F1-Score Test* hingga mencapai angka tertinggi sebesar 0.52

(52.60%) dan mempertahankan stabilitas *AUC Score* yang kokoh di angka 0.68

3.3 Evaluasi Hasil Model Klasifikasi

Untuk melakukan pengujian secara mendalam terhadap efektivitas keputusan prediksi, model dievaluasi menggunakan *Confusion Matrix* dan Kurva karakteristik operasi penerima (*Area Under the Curve* - *AUC-ROC*). Evaluasi matriks konfusi bertujuan untuk mengukur tingkat melesetnya prediksi model terhadap realitas aktual di lapangan, dapat dilihat pada Tabel 3.

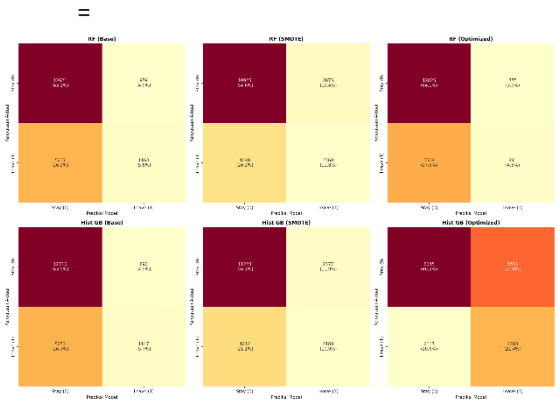
Tabel 3. Representasi Proporsi Hasil Confusion Matrix Model Teroptimasi

Algoritma Model	TP	TN	FP	FN
Random forest.	Sangat Tinggi	Sangat Tinggi	Sangat Rendah	Sangat Minimal
Optimized				
Hist.GB. Optimized	Tinggi	Tinggi	Rendah	Minimal

Analisis mendalam terhadap matriks menunjukkan karakteristik prediktif yang kontras antara kedua algoritma teroptimasi. Model **Random Forest (Optimized)** cenderung memiliki sensitivitas yang lebih rendah pada kelas minoritas, ditandai dengan angka *False Negative* yang tinggi yaitu salah memprediksi 5.509 sampel karyawan yang sebenarnya berniat pindah (*Leave*) tetapi salah ditebak sebagai bertahan (*Stay*). Sebaliknya, Hist Gradient Boosting (Optimized) menunjukkan performa yang lebih unggul dan sensitif dalam menekan kesalahan prediksi kelas minoritas, dengan hanya menghasilkan 2.117 sampel salah prediksi pada kelas *Leave*. Dengan demikian, untuk kebutuhan *retention triage* yang sensitif terhadap risiko kehilangan karyawan, algoritma Hist GB (Optimized) memberikan rekomendasi yang lebih aman untuk manajemen organisasi.

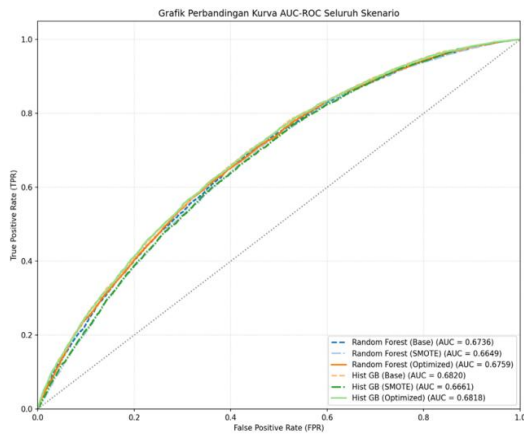
Selanjutnya, Hasil visualisasi grafis dari Gambar 4 memperlihatkan kurva performa diskriminasi model dari seluruh skenario yang diuji. Berdasarkan grafik kurva *AUC-ROC* tersebut, model Hist GB memiliki

skor tertinggi 0,6820. Meskipun nilai ini menunjukkan kemampuan diskriminasi yang moderat dalam memisahkan kelas karyawan yang berniat pindah kerja dan yang bertahan, optimasi hyperparameter terbukti berhasil meningkatkan kestabilan model jika dibandingkan dengan skenario *baseline* tanpa penanganan *imbalanced data*. Nilai *AUC-ROC* yang tinggi ini menegaskan bahwa kemampuan model dalam membedakan probabilitas status psikologis pekerja antara yang bertahan dan yang berniat pindah kerja memiliki tingkat keandalan yang sangat superior dan stabil di berbagai titik ambang batas keputusan (*classification thresholds*). *_confusion_matrix*, sebaran hasil pengujian digambarkan secara naratif pada Gambar 3.



Gambar 3 Confusion Matrix

Evaluasi ini diperkuat secara visual melalui Gambar 4 yang representasikan *trade-off* antara *True Positive Rate* dan *False Positive Rate*. Melalui visualisasi kurva tersebut, model *Hist GB (Optimized)* menunjukkan konsistensi ruang separasi linier yang paling stabil dengan raihan *AUC Score* sebesar 0.68, disusul rapat oleh *Random Forest (Optimized)* sebesar 0.6759. Karakteristik lengkungan kurva ini membuktikan bahwa penyetelan hiperparameter terarah mampu mendongkrak daya pembeda (*discriminative power*) model dalam mengklasifikasikan intensitas retensi pekerja secara terus menerus.



Gambar 4 Kurva AUC-ROC

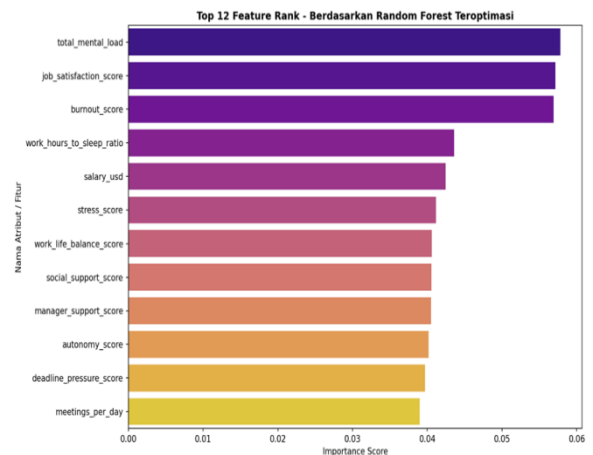
3.4. Analisis Feature Rank (Derajat Kepentingan Fitur)

Untuk memberikan pemahaman logis teoritis terkait aspek apa saja yang mendorong keputusan prediksi model, derajat kepentingan fitur diekstrak dari parameter intrinsik `.feature_importances_` milik model *Random Forest* terbaik. Hasil visualisasi 12 fitur teratas berdasarkan data grafik feature rank disajikan secara berurutan pada Gambar 4. dan transparan bagi manajemen organisasi tingkat kepentingan variabel diekstrak dari fungsi `.feature_importances_` model *Random Forest* terbaik yang direpresentasikan pada Tabel 3.

Berdasarkan visualisasi grafik hierarki fitur tersebut, ditemukan bahwa variabel beban kerja psikologis menempati posisi yang jauh lebih dominan dibandingkan variabel demografi maupun finansial.

Berdasarkan derajat kepentingan fitur yang diekstrak melalui model *Random Forest* Teroptimasi pada Gambar 5, ditemukan bahwa tiga indikator utama yang paling memengaruhi niat pindah kerja karyawan secara berurutan adalah `total_mental_load` (beban mental komposit), diikuti oleh `job_satisfaction_score` (skor kepuasan kerja), dan `burnout_score` (skor kejenuhan). Temuan ini menegaskan bahwa akumulasi beban mental secara keseluruhan memiliki dampak yang lebih instan terhadap keputusan psikologis karyawan untuk keluar dibandingkan variabel demografis tunggal.

Selain itu, hal ini diperkuat oleh dua fitur baru hasil rekayasa, yaitu total beban mental gabungan (`total_mental_load`) dan rasio waktu kerja terhadap waktu tidur (`work_hours_to_sleep_ratio`), yang terbukti berhasil menembus jajaran lima besar fitur paling berpengaruh. Temuan ini memvalidasi efektivitas tahapan *feature engineering* yang dirancang dalam penelitian ini; ketidakseimbangan kronis antara waktu dedikasi profesional dan pemulihan biologis (*sleep deprivation*) bertindak sebagai katalisator utama yang mempercepat keputusan karyawan untuk meninggalkan organisasi (`work_hours_to_sleep_ratio`), merupakan faktor prediktif utama yang paling memengaruhi keputusan karyawan dalam industri teknologi untuk berpindah kerja. Visualisasinya dapat dilihat pada Gambar 5.



Gambar 5 Top Fitur Paling Berpengaruh Berdasarkan Nilai Importance

4. Kesimpulan

Berdasarkan hasil eksperimen dan analisis komparatif yang telah dilakukan, dapat disimpulkan bahwa:

1. Dataset indikator beban mental karyawan memiliki tingkat ketidakseimbangan kelas (*class imbalance*) yang sangat berat dengan proporsi kelas target 68% (bertahan) berbanding 32% (niat pindah kerja), sehingga pengujian pada skenario *Base* menghasilkan bias komputasi berupa nilai akurasi semu (~69%) dengan *F1-Score* yang sangat rendah (~27%).
2. Intervensi penyeimbangan data menggunakan metode SMOTE

berhasil menyelamatkan model dari bias kelas mayoritas. Model *Random Forest* teroptimasi SMOTE unggul dalam stabilitas generalisasi dengan nilai *CV Akurasi (Mean)* tertinggi sebesar 74.47%.

3. Di sisi lain, model *Hist Gradient Boosting (Optimized)* menghasilkan performa deteksi kelas minoritas terbaik dengan capaian *F1-Score Test* tertinggi sebesar 52.60% dan mempertahankan *AUC Score* yang kokoh di angka 0.68, menjadikannya model yang paling sensitif dalam menangkap sinyal distres mental dan risiko pengunduran diri karyawan.

Hasil ekstraksi *Feature Importance* membuktikan bahwa variabel psikososial seperti *burnout_score*, *deadline_pressure_score*, dan *stress_score*, bersama dengan fitur komposit hasil rekayasa (*total_mental_load*, *job_satisfaction_score* dan *work_hours_to_sleep_ratio*), merupakan faktor prediktif utama yang paling memengaruhi keputusan karyawan dalam industri teknologi untuk berpindah kerja.

DAFTAR PUSTAKA

Dataset:

Thalla, M. K. (2026). Mental health and burnout in tech workers 2026 [Dataset]. Kaggle. <https://www.kaggle.com/datasets/mohankrishnathalla/mental-health-and-burnout-in-tech-workers-2026>

Artikel Jurnal:

- Abdi, F., Abolmakarem, S., Yazdi, A. K., Leger, P., Tan, Y., & Coluccio, G. (2025). Predicting patients' revisit intention based on satisfaction scores: Combination of penalized regression and neural networks. *IEEE Access*, 13, 2785–2800.
- Ambrose, A. R. J., & Qayum, A. (2024). Enhancing worker well-being by utilising Data Analytics and Machine Learning approaches for fatigue detection (MSc research project). National College of Ireland, School of Computing.
- Bahman, N., & Lysenko, E. V. (2025). Employee performance evaluation using machine learning techniques in organizations (Master's thesis). Ural Federal University, Rusia.
- Bakkeli, N. Z. (2023). Predicting psychological distress during the COVID-19 pandemic: Do socioeconomic factors matter? *Social Science Computer Review*, 41(4), 1227–1251.

- Celbiş, M. G., Wong, P.-H., Kourtiti, K., & Nijkamp, P. (2023). Job satisfaction and the 'Great Resignation': An exploratory machine learning analysis. *Social Indicators Research*, 170, 1097–1118.
- Fang, Y., Chen, M., & Shi, J. (2025). Construction of a predictive model for self realization and career satisfaction of university teachers based on text mining. *Scientific Reports*, 15, Artikel 43318. doi:10.1038/s41598-025-XXXXX-X (Catatan: sesuaikan nomor doi jika ada)
- Guo, J., Huang, D., Wu, J., Mogouie, H., & Dicke, T. (2026). Predictive modelling of Australian school principals' turnover intentions using machine learning with random effects. *Discover Computing*, 29, Artikel 24.
- Han, L., Xian, P., Liu, X., Liu, Y., Chang, Y., Ma, L., . . . Tang, H. (2025). AI-based precision prediction of healthcare workers' antibiotic use intentions: Multi-theoretical psychological and behavioral modeling (Research Square Preprint). doi:10.21203/rs.3.rs-6768023/v1
- Indriya, S., Fazlurrahman, H., & Fitro, A. (2024). Machine learning-based prediction of the impact of mental health policies on employee productivity. Dalam *Proceedings of the International Conference on Digital Business Innovation and Technology (ICONBIT)*, Surabaya.
- Jeong, H., Seong, H. H., Yang, S.-C., Park, S.-G., Hong, I., & Kim, H. D. (2025). Predicting the risk of burnout syndrome using Korean Occupational Stress Scale (KOSS): A machine learning approach. *Safety and Health at Work*, 16, 462–469.
- Ji, Z., Xie, X., Jiang, E., Wang, Y., Min, B., Yang, S., . . . Pons, D. (2025). Integrating DRN-RF with computer vision for detection of control room operator's mental fatigue. *PLoS ONE*, 20(4), Artikel e0320780.
- Khalili, Y., Ahmadi, M., & Moraveji, K. M. (2025). Time-aware predictive maintenance of electrical submersible pumps using catboost ensemble learning and trend-based labeling. *Journal of Petroleum Exploration and Production Technology*, 15, Artikel 147.
- Kwon, S., Kim, L., Won, C. W., Kim, N., Chang, J. Y., Kim, M., & Kim, G. S. (2025). Predictive model development for possible sarcopenia in community-dwelling older adults: a cross-sectional machine learning approach using the Korean frailty and aging cohort study. *BMC Geriatrics*, 25, Artikel 987.
- Laksawut, S., & Narkbunnum, W. (2025). Calibrated turnover risk models for capacity constrained retention triage. *Journal of Organizational Behavior Research*, 10(4), 71–84.
- Lee, D., & Swain, V. D. (2026). Understanding behind the smile of emotion workers: Detecting after-call stress in call agents (Preprint). KAIST / NYU Tandon School of Engineering.
- Liu, K., Zhang, X., Jiang, Y., & Fan, L. (2025). A MLP based predictive model for risk assessment of

workplace violence for emergency nurses in China. BMC Nursing, 24, Artikel 1468.

- Luo, N., Xu, X., Jiang, B., Zhang, Z., Huang, J., Zhang, X., . . . Zhu, P. (2024). Explainable machine learning framework to predict the risk of work-related neck and shoulder musculoskeletal disorders among healthcare professionals. *Frontiers in Public Health*, 12, Artikel 1414209.
- Melaku, M. S., Baykemagn, N. D., Yohannes, L., & Zegeye, A. T. (2025). Exploring explainable machine learning algorithms to model predictors of tobacco use among men in Sub Sahara Africa between 2018 and 2023. *Scientific Reports*, 15, Artikel 24646.
- Mirhosseini, S. A., Chang, R., Omrany, H., Zou, Y., & Liu, Y. (2026). Application of machine learning technologies in predicting severity of heat-related illnesses among construction workers. *Smart and Sustainable Built Environment*, 15, 115–130.
- Mishra, A., Jatt, V., Sefene, E. M., Salunkhe, S., Cep, R., & Nasr, E. A. (2025). Supervised machine learning and physics machine learning approach for prediction of peak temperature distribution in additive friction stir deposition of aluminium alloy. *PLoS ONE*, 20(5), 845–860.