



Prediction of Academic Burnout in College Students: A Comparative Analysis of Support Vector Machine and Random Forest Algorithms

Prediksi Academic Burnout pada Mahasiswa: Analisis Komparatif Algoritma Support Vector Machine dan Random Forest

Arbiansyah Adinegara^{1*}, Suryarini Widodo²

^{1,2}Magister Manajemen Sistem Informasi, Program Pascasarjana Magister Teknologi dan Rekayasa, Universitas Gunadarma, Indonesia

E-Mail: ¹arbiansyahadinegara@gmail.com, ²srini@staff.gunadarma.ac.id

Received Aug 01st 2025; Revised Oct 14th 2025; Accepted Oct 18th 2025; Available Online Oct 31th 2025

Corresponding Author: Arbiansyah Adinegara

Copyright © 2025 by Authors, Published by Institut Riset dan Publikasi Indonesia (IRPI)

Abstract

Academic burnout has become a significant problem among students, negatively impacting mental health and academic performance. This study aims to conduct a comparative analysis of the performance of Support Vector Machine (SVM) and Random Forest (RF) algorithms in predicting academic burnout among students at a university in DKI Jakarta. This quantitative research method utilizes primary data from the Burnout Assessment Tool – Student Version (BAT-S) questionnaire, which covers personal, academic, and psychological factors, as well as secondary academic data of students. The raw data then undergoes a preparation stage that includes cleaning, handling outliers with capping techniques, standardization, and class balancing using BorderlineSMOTE to address the imbalanced data distribution. The modeling results show very high predictive performance for both algorithms after hyperparameter optimization, with SVM achieving 98.75% accuracy and RF achieving 97.50% on the test data. Although RF showed superiority in ranking-based metrics such as ROC-AUC, SVM is recommended as the final model because it has a more clinically acceptable error risk profile, i.e., it does not produce high-risk false negatives. This study demonstrates that machine learning can be an effective tool for early detection of burnout risk, but further research is recommended to explore more complex algorithms such as gradient boosting and perform feature importance analysis for deeper understanding.

Keyword: *Academic Burnout, Machine Learning, Random Forest, Support Vector Machine, University Students.*

Abstrak

*Academic burnout telah menjadi masalah signifikan di kalangan mahasiswa, berdampak negatif pada kesehatan mental dan kinerja akademik. Penelitian ini bertujuan untuk melakukan analisis komparatif terhadap kinerja algoritma Support Vector Machine (SVM) dan Random Forest (RF) dalam memprediksi *academic burnout* pada mahasiswa di salah satu universitas di DKI Jakarta. Metode penelitian kuantitatif ini menggunakan data primer dari kuesioner *Burnout Assessment Tool – Student Version* (BAT-S) yang mencakup faktor pribadi, akademik, dan psikologis, serta data sekunder akademik mahasiswa. Data mentah kemudian melalui tahap persiapan yang meliputi pembersihan, penanganan outlier dengan teknik *capping*, standarisasi, dan penyeimbangan kelas menggunakan *BorderlineSMOTE* untuk mengatasi distribusi data yang tidak seimbang. Hasil pemodelan menunjukkan performa prediktif yang sangat tinggi untuk kedua algoritma setelah optimasi *hyperparameter*, dengan SVM mencapai akurasi 98,75% dan RF sebesar 97,50% pada data uji. Meskipun RF menunjukkan keunggulan pada metrik berbasis peringkat seperti ROC-AUC, SVM direkomendasikan sebagai model akhir karena memiliki profil risiko kesalahan yang lebih dapat diterima secara klinis, yakni tidak menghasilkan *false negative* yang berisiko tinggi. Penelitian ini membuktikan bahwa *machine learning* dapat menjadi alat yang efektif untuk deteksi dini risiko *burnout*, namun penelitian selanjutnya disarankan untuk mengeksplorasi algoritma yang lebih kompleks seperti gradient boosting dan melakukan analisis kepentingan fitur untuk pemahaman yang lebih mendalam.*

Kata Kunci: *Academic Burnout, Machine Learning, Mahasiswa, Random Forest, Support Vector Machine.*

1. PENDAHULUAN

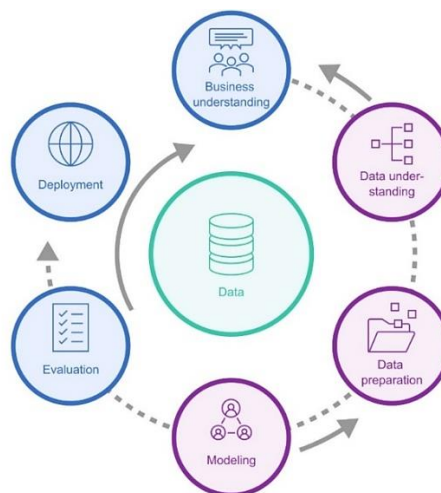
Perkembangan *Machine Learning* (ML) dalam dekade terakhir telah membuka peluang signifikan untuk memprediksi masalah psikologis dan akademik secara lebih akurat, termasuk *academic burnout* [1][2]. Fenomena *burnout* pada mahasiswa, yang secara formal didefinisikan sebagai sindrom kelelahan akibat tuntutan belajar, sikap sinis terhadap nilai sekolah, dan perasaan tidak mampu secara akademis, telah menjadi masalah yang mengkhawatirkan dengan prevalensi tinggi di berbagai negara [3][4]. Studi menunjukkan bahwa lebih dari separuh mahasiswa mengalami kelelahan akademis dengan berbagai tingkat keparahan [1], sementara penelitian lain mencatat prevalensi gejala *burnout* antara 33,4% hingga 55% pada mahasiswa kedokteran [3]. Kondisi ini memerlukan pendekatan analitis yang mampu mengintegrasikan variabel multidimensi seperti beban akademik, kesehatan mental, dan dukungan sosial [1]. Tantangan utama dalam memprediksi *burnout* terletak pada kompleksitas interaksi antar variabel tersebut, yang sering kali bersifat non-linier dan mengandung *noise*. Di sinilah algoritma ML seperti Support Vector Machine (SVM) dan Random Forest (RF) menawarkan solusi berbasis data, dengan kemampuan mengidentifikasi pola tersembunyi yang sulit diolah oleh metode statistik konvensional [5].

Penelitian ini didasari oleh kebutuhan untuk memprediksi *academic burnout* pada mahasiswa menggunakan ML, di mana dua algoritma klasifikasi utama, yaitu SVM dan RF, dipilih dan dianalisis secara komparatif. Pemilihan SVM didukung oleh keunggulan teoretisnya dalam mencari *hyperplane* optimal, efisiensi pada data multidimensi, dan penerapan kernel trick untuk pemisahan non-linier [6][7]. Sementara itu, RF dipilih sebagai metode ensemble yang menggabungkan prediksi dari ratusan decision tree untuk mengurangi varians, membuatnya tangguh terhadap *overfitting* dan dikenal karena stabilitasnya dalam klasifikasi [8]. Riset-riset terdahulu telah menjadi landasan metodologis bagi studi ini, di mana efektivitas algoritma ML telah ditunjukkan untuk memprediksi kinerja akademik, dan penelitian tentang hubungan antara stres dan *academic burnout* juga telah dilakukan pada populasi mahasiswa [9]. Namun, penelitian ini membedakan diri dari riset-riset sebelumnya karena, meskipun dasar teoretis algoritma dan konteks akademik telah dibahas, implementasi model ML secara komparatif antara SVM dan RF untuk memprediksi *academic burnout* pada populasi mahasiswa di Jakarta masih terbatas, sehingga perlu dianalisis pola data unik yang mungkin dihasilkan oleh konteks lokal, seperti tekanan kompetitif yang tinggi dan mekanisme koping di lingkungan urban Jakarta.

Kebaruan penelitian dalam dokumen tesis ini terletak pada analisis komparatif mendalam terhadap kinerja algoritma SVM dan RF untuk memprediksi *academic burnout* secara spesifik pada populasi mahasiswa Jakarta guna mengisi kesenjangan pengetahuan kontekstual tersebut. Penelitian ini bertujuan untuk menyediakan model prediksi yang akurat dan stabil. Secara spesifik, studi ini akan menjawab tiga pertanyaan penelitian utama: (1) Bagaimana kinerja kuantitatif dan kualitatif algoritma SVM dan RF dalam memprediksi *academic burnout*? (2) Faktor-faktor prediktif apa yang paling berpengaruh dalam model? dan (3) Apakah terdapat perbedaan kinerja yang signifikan antara kedua algoritma yang dapat memberikan justifikasi untuk pemilihan model implementasi?

2. BAHAN DAN METODE

Penelitian ini menggunakan kerangka kerja *Cross-Industry Standard Process for Data Mining* (CRISP-DM) untuk memastikan proses penambangan data yang terstruktur dan sistematis. Sebagai standar industri untuk proyek ilmu data, metodologi ini bersifat fleksibel dan memandu penelitian melalui enam fase iteratif, mulai dari pemahaman bisnis hingga implementasi, untuk menghasilkan wawasan yang relevan dan dapat ditindaklanjuti [10]. Siklus dan alur kerja dari keenam fase tersebut diilustrasikan pada Gambar 1.



Gambar 1. Metodologi CRISP-DM [11]

2.1. Business & Data Understanding

Tahap awal *Business Understanding*, berfokus pada definisi masalah dan tujuan penelitian, yaitu mengembangkan dan membandingkan model prediktif untuk deteksi dini *academic burnout* pada mahasiswa. Keberhasilan proyek ini diukur dari kemampuan model untuk mengklasifikasikan tingkat risiko *burnout* dengan akurasi tinggi dan, yang lebih penting, dengan profil kesalahan yang dapat diterima secara klinis.

Tahap *Data Understanding* melibatkan pengumpulan dan eksplorasi data awal. Populasi penelitian adalah seluruh mahasiswa aktif angkatan 2019 di salah satu universitas di Jakarta. Dari populasi tersebut, sampel sebanyak 160 mahasiswa dipilih menggunakan teknik stratified random sampling untuk memastikan representasi yang proporsional dari setiap fakultas. Pengumpulan data primer dilakukan secara daring menggunakan instrumen *Burnout Assessment Tool – Student Version (BAT-S)*, sebuah kuesioner yang telah divalidasi secara komprehensif pada konteks internasional di Chile dan lokal di Indonesia, menunjukkan reliabilitas dan validitas yang kuat [12][13]. Instrumen ini mengukur enam dimensi inti *burnout* melalui 33 item dengan skala *Likert* 5 poin: *Exhaustion*, *Mental Distance*, *Cognitive Impairment*, *Emotional Impairment*, *Psychological Distress*, dan *Psychosomatic Complaints*. Data sekunder berupa catatan akademik mahasiswa (IPK, Semester, dan program studi) juga diperoleh dari pusat data akademik universitas untuk melengkapi analisis.

2.2. Tinjauan Literatur Metode

Pemilihan metode didasarkan pada landasan teoritis yang kuat dan rekam jejak keberhasilan dalam domain serupa.

2.2.1. Support Vector Machine (SVM)

SVM adalah algoritma supervised learning yang berakar pada Teori Pembelajaran Statistik. Berbeda dengan metode lain yang berfokus pada minimalisasi kesalahan empiris, tujuan utama SVM adalah meminimalkan risiko struktural, yaitu mencari keseimbangan optimal antara kompleksitas model dan kemampuannya untuk generalisasi pada data baru [7]. Mekanisme intinya adalah menemukan *hyperplane* atau batas keputusan optimal yang secara maksimal memisahkan kelas-kelas data dalam ruang fitur. *Hyperplane* ini ditentukan oleh titik-titik data terdekat dari setiap kelas, yang dikenal sebagai *support vectors*. Kemampuan SVM untuk menangani hubungan non-linier yang kompleks ditingkatkan secara signifikan melalui kernel trick, yang memetakan data ke ruang berdimensi lebih tinggi di mana pemisahan linier menjadi mungkin [14]. Aplikasinya dalam memprediksi stres dan masalah kesehatan mental pada mahasiswa telah terbukti efektif dan mampu mencapai akurasi tinggi [15].

2.2.2. Random Forest (RF)

RF adalah metode ensemble learning yang kuat, yang bekerja dengan membangun sejumlah besar pohon keputusan (decision trees) selama fase pelatihan dan menggabungkan hasilnya untuk prediksi akhir [8]. Kekuatan RF terletak pada dua sumber kerandoman: setiap pohon dilatih pada sampel data acak yang diambil melalui *bootstrap aggregating (bagging)*, dan pada setiap titik pemisahan (*split*) di dalam pohon, hanya subset acak dari total fitur yang dipertimbangkan. Proses ganda ini secara efektif mengurangi varians model dan menurunkan korelasi antar pohon, sehingga menghasilkan model yang sangat tahan terhadap *overfitting* dan stabil [16]. RF sangat cocok untuk data psikososial yang kompleks karena kemampuannya menangani data berdimensi tinggi tanpa memerlukan eliminasi fitur yang ekstensif. Penelitian komparatif oleh Lara-González et al. secara spesifik menemukan bahwa RF menunjukkan kinerja yang lebih superior dibandingkan SVM dalam memproyeksikan risiko *burnout* [17].

2.3. Deskripsi Dataset dan Variabel

Dataset awal terdiri dari 160 sampel. Setelah pra-pemrosesan dengan *BorderlineSMOTE*, total data yang digunakan untuk pemodelan menjadi 399 sampel. *Dataset* ini terdiri dari enam atribut prediktor (fitur) dan satu atribut target. Cuplikan *dataset* awal dapat dilihat pada Tabel 1.

Table 1. Cuplikan *Dataset* Awal

<i>Exhaustion</i>	<i>Mental Distance</i>	<i>Cognitive Impairment</i>	<i>Emotional Impairment</i>	<i>Psychological Distress</i>	<i>Psychosomatic Complaints</i>
3.555	3.50	3.6	3.8	3.6	3.50
3.666	3.50	3.4	3.4	4.4	3.50
3.555	3.50	3.8	3.6	4.2	3.50

Untuk pelatihan dan pengujian, *dataset* yang telah seimbang (399 sampel) dibagi menjadi data latih (319 sampel, ~80%) dan data uji (80 sampel, ~20%) menggunakan *stratified split* untuk menjaga proporsi kelas di kedua set.

2.4. Data Preparation

Tahap *Data Preparation* adalah fase paling intensif dalam penelitian ini, mencakup semua kegiatan untuk menyusun *dataset* akhir dari data mentah.

2.4.1. Pembersihan Data

Integritas *dataset* awal diperiksa. Dilakukan pengecekan untuk nilai yang hilang (*missing values*) pada seluruh variabel. Hasil pemeriksaan menunjukkan bahwa *dataset* memiliki kualitas yang sangat baik dan tidak memerlukan proses imputasi.

2.4.2. Penanganan Outlier

Outlier atau nilai ekstrem dideteksi menggunakan metode *Interquatile Range* (IQR), di mana titik data di luar rentang $1.5 \times \text{IQR}$ dari kuartil pertama dan ketiga dianggap sebagai *outlier*. Sebanyak 56 titik data ekstrem diidentifikasi pada berbagai fitur. Untuk menangani *outlier* ini tanpa menghilangkan data, teknik *capping* (juga dikenal sebagai *Winsorization*) diterapkan. Nilai *outlier* di bawah batas bawah diganti dengan nilai batas bawah, dan nilai yang melebihi batas atas diganti dengan nilai batas atas.

2.4.3. Standardisasi Data

Untuk memastikan setiap fitur memiliki skala yang seragam dan berkontribusi secara setara pada model, data distandardisasi menggunakan *StandardScaler* dari *library scikit-learn*. Proses ini mentransformasi setiap fitur sehingga memiliki rata-rata nol dan standar deviasi satu. Ini adalah langkah krusial, terutama untuk algoritma yang sensitif terhadap skala seperti SVM.

2.4.4. Penanganan Ketidakseimbangan Kelas

Analisis distribusi variabel target menunjukkan adanya ketidakseimbangan kelas yang ekstrem, dengan rasio kelas mayoritas terhadap minoritas mencapai 11.1:1. Kondisi ini dapat menyebabkan model menjadi bias. Untuk mengatasinya, teknik *oversampling* canggih, *BorderlineSMOTE*, diterapkan. Teknik ini secara cerdas mensintesis sampel baru untuk kelas-kelas minoritas dengan berfokus pada sampel yang berada di sekitar batas keputusan,

2.5. Modelling & Evaluation

Tahap pemodelan dilakukan dengan menerapkan dua prinsip ML yang berbeda: SVM, yang bekerja dengan mencari *hyperplane* optimal untuk memisahkan kelas data (14–16), dan RF, yang menggabungkan prediksi dari banyak pohon keputusan untuk meningkatkan akurasi dan stabilitas (17–19). Untuk menemukan konfigurasi model yang paling optimal, proses pencarian *hyperparameter* secara ekstensif dilakukan menggunakan *GridSearchCV*. Teknik ini secara sistematis menguji berbagai kombinasi *hyperparameter* yang telah ditentukan (misalnya, C, kernel, gamma untuk SVM; n_estimators, max_depth untuk RF) yang dikombinasikan dengan validasi silang 5-lipatan (5-fold cross-validation) untuk mencegah *overfitting* dan memilih model dengan kemampuan generalisasi terbaik.

Pada tahap Evaluation, kinerja kedua model yang telah dioptimalkan dievaluasi pada *dataset* uji yang terpisah. Metrik evaluasi yang digunakan mencakup akurasi, presisi, *recall*, *F1-score*, serta kurva *Receiver Operating Characteristic* (ROC) dan *Area Under the Curve* (AUC) untuk menilai kinerja model secara komprehensif di berbagai ambang batas klasifikasi.

3. HASIL DAN PEMBAHASAN

3.1. Data Preprocessing

Hasil dari tahap persiapan data menunjukkan transformasi *dataset* yang signifikan. Analisis awal pada 160 sampel mentah mengidentifikasi adanya ketidakseimbangan kelas yang ekstrem. Kelas "*Red* (*Burnout Tinggi*)" sangat mendominasi dengan 133 sampel (83.1%), sementara kelas "*Green* (*Tidak Burnout*)" dan "*Orange* (*Risiko Burnout*)" merupakan kelas minoritas yang sangat kecil, masing-masing hanya terdiri dari 15 sampel (9.4%) dan 12 sampel (7.5%). Setelah penerapan *BorderlineSMOTE*, distribusi kelas berhasil dinormalisasi menjadi seimbang secara sempurna. Setiap kelas ("*Green*", "*Orange*", dan "*Red*") kini memiliki 133 sampel, sehingga total *dataset* yang digunakan untuk melatih dan menguji model menjadi 399 sampel. Transformasi ini sangat krusial untuk mencegah model menjadi bias terhadap kelas mayoritas dan memastikan kemampuan model untuk mempelajari pola dari setiap kelas secara merata.

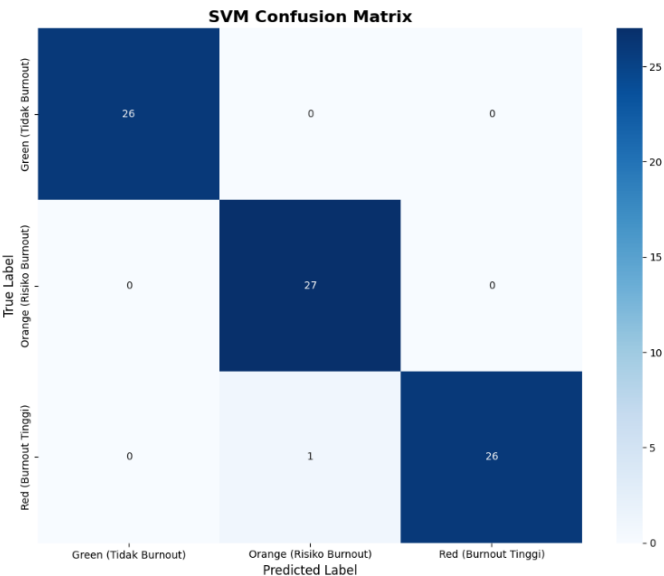
3.2. Kinerja Model Support Vector Machine (SVM)

Setelah proses optimasi *hyperparameter* yang ekstensif, konfigurasi terbaik untuk model SVM ditemukan pada {'C': 1, 'gamma': 'auto', 'kernel': 'poly'}. Pemilihan kernel 'poly' (polinomial) mengindikasikan bahwa hubungan antara fitur-fitur *burnout* dan kelas target bersifat non-linear namun terstruktur. Model ini mencapai akurasi pelatihan sebesar 100.00% dan akurasi pengujian yang sangat tinggi yaitu 98.75%. Laporan klasifikasi model SVM pada data uji dapat dilihat pada Tabel 2.

Table 2. Laporan Klasifikasi Model SVM pada Data Uji

Kategori	Precision	Recall	F1-Score	Support
Green (Tidak Burnout)	1.00	1.00	1.00	26
Orange (Risiko Burnout)	0.96	1.00	0.98	27
Red (Burnout Tinggi)	1.00	0.96	0.98	27
accuracy			0.99	80
macro avg	0.99	0.99	0.99	80
weighted avg	0.99	0.99	0.99	80

Analisis matriks konfusi (Gambar 2) memberikan pandangan granular terhadap kinerja model. Dari 80 sampel pada data uji, model SVM hanya membuat satu kesalahan klasifikasi. Kesalahan tersebut adalah mengklasifikasikan satu sampel yang seharusnya "Red (Burnout Tinggi)" sebagai "Orange (Risiko Burnout)".



Gambar 2. Matriks Konfusi untuk Model SVM pada Data Uji

3.3. Kinerja Model Random Forest (RF)

Model RF dioptimalkan dan menghasilkan konfigurasi terbaik dengan {'n_estimators': 100, 'max_depth': None, 'min_samples_split': 2, 'min_samples_leaf': 4, 'max_features': 'sqrt'}. Konfigurasi ini mencerminkan keseimbangan antara kompleksitas dan regularisasi. Model ini mencapai akurasi pelatihan 99.06% dan akurasi pengujian 97.50%. Laporan klasifikasi model RF pada data uji dapat dilihat pada Tabel 3.

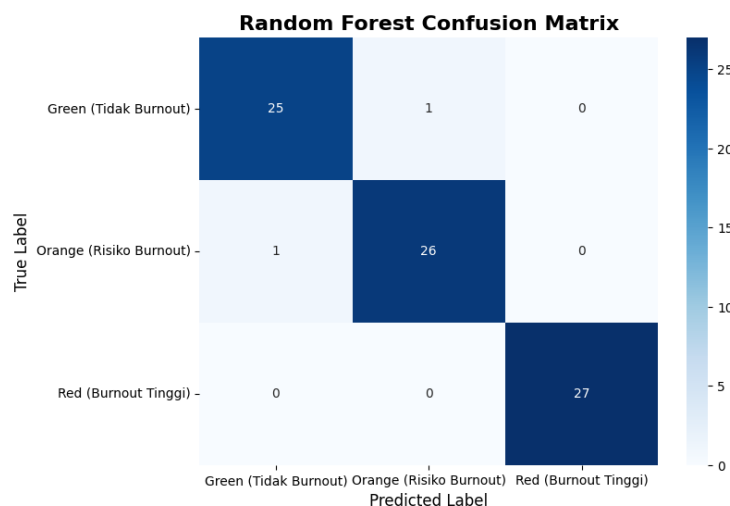
Table 3. Laporan Klasifikasi Model RF pada Data Uji

Kategori	Precision	Recall	F1-Score	Support
Green (Tidak Burnout)	0.96	0.96	0.96	26
Orange (Risiko Burnout)	0.96	0.96	0.96	27
Red (Burnout Tinggi)	1.00	1.00	1.00	27
accuracy			0.97	80
macro avg	0.97	0.97	0.97	80
weighted avg	0.97	0.97	0.97	80

Matriks konfusi untuk model RF (Gambar 3) menunjukkan dua kesalahan klasifikasi. Kesalahan yang paling signifikan secara praktis adalah misklasifikasi satu sampel "Orange (Risiko Burnout)" sebagai "Green (Tidak Burnout)".

3.4. Perbandingan Kinerja Model

Bagian ini menyajikan perbandingan langsung antara model SVM dan RF menggunakan serangkaian metrik evaluasi standar dan lanjutan. Tujuannya adalah untuk mendapatkan pemahaman yang komprehensif mengenai kekuatan dan kelemahan relatif dari setiap model, melampaui sekadar perbandingan skor akurasi tunggal.



Gambar 3. Matriks Konfusi untuk Model RF pada Data Uji

3.4.1. Perbandingan Metrik Klasifikasi Standar

Evaluasi awal dilakukan dengan membandingkan metrik-metrik klasifikasi standar yang dihitung dari hasil prediksi pada data uji. Metrik-metrik ini, yaitu Akurasi, Presisi, *Recall*, dan *F1-Score*, memberikan gambaran performa model pada ambang batas klasifikasi default. Tabel 4 merangkum perbandingan ini.

Table 4. Perbandingan Metrik Klasifikasi

Metrik	SVM	RF
<i>Test Accuracy</i>	0.9875	0.9750
<i>Macro Precision</i>	0.9900	0.9700
<i>Macro Recall</i>	0.9900	0.9700
<i>Macro F1-Score</i>	0.9900	0.9700

Hasil perbandingan ini secara konsisten menunjukkan bahwa model SVM sedikit lebih unggul daripada RF pada keempat metrik standar. SVM mencapai akurasi 98.75% dibandingkan dengan 97.50% pada RF, dan juga menunjukkan nilai yang lebih tinggi untuk Presisi, *Recall*, dan *F1-Score* rata-rata makro.1 Keunggulan ini, meskipun tipis, menunjukkan bahwa pada ambang batas klasifikasi standar, prediksi akhir yang dihasilkan oleh SVM lebih akurat. Visualisasi perbandingan akurasi dan metrik lainnya dapat dilihat pada Gambar 4 dan 5

3.4.2. Perbandingan Analisis Mendalam: Kinerja Berbasis Peringkat vs Klasifikasi

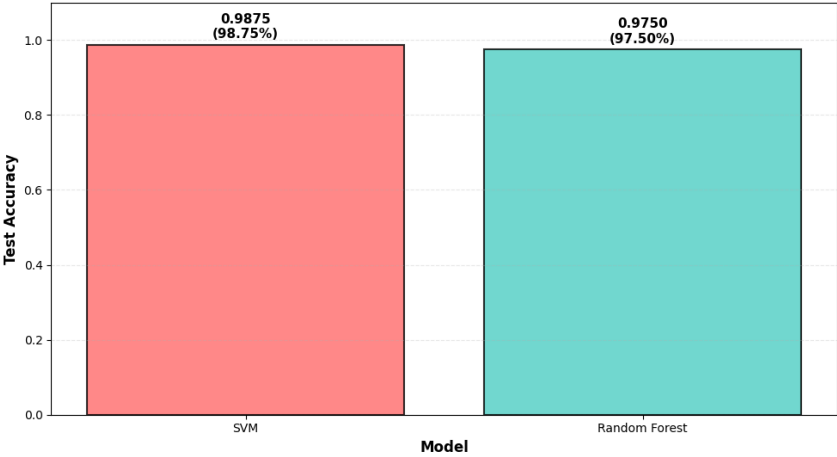
Untuk evaluasi yang lebih mendalam dan tidak bergantung pada satu ambang batas klasifikasi, analisis dilanjutkan dengan menggunakan kurva ROC dan kurva *Precision-Recall* (PR). Metrik turunan dari kurva ini, yaitu AUC dan *Average Precision* (AP), mengukur kemampuan model dalam membedakan antar kelas di semua kemungkinan ambang batas.

Gambar 6 dan 7 menyajikan evaluasi kinerja komparatif antara model SVM dan RF menggunakan kurva ROC dan *Precision-Recall*. Pada Gambar 6, kurva ROC untuk kedua model menunjukkan performa yang sangat luar biasa, dengan bentuk kurva yang mendekati sudut kiri atas dan nilai AUC untuk setiap kelas hampir sempurna (berkisar antara 0.992 hingga 1.000). Perbandingan mikro-rata-rata juga mengonfirmasi keunggulan tipis dari model RF (AUC = 0.996) dibandingkan SVM (AUC = 0.993). Temuan ini diperkuat lebih lanjut pada Gambar 7, di mana kurva *Precision-Recall* menunjukkan kemampuan kedua model untuk mempertahankan presisi tinggi di berbagai tingkat perolehan (*recall*). Sekali lagi, model RF menunjukkan kinerja yang sedikit lebih unggul dengan AP sebesar 0.993, dibandingkan dengan SVM yang memiliki AP sebesar 0.987. Secara keseluruhan, kedua visualisasi ini secara konsisten menunjukkan bahwa meskipun kedua model sangat akurat dan andal, RF memiliki keunggulan performa marginal dalam tugas klasifikasi ini.

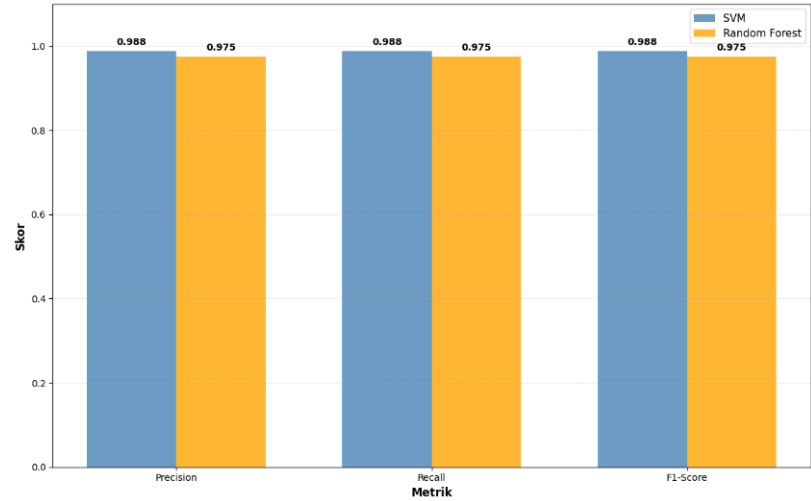
Meskipun RF unggul dalam metrik ini, perlu dipertimbangkan adanya potensi kontradiksi jika dibandingkan dengan metrik lain seperti akurasi. Kontradiksi ini memiliki makna yang penting: kemenangan SVM pada metrik seperti akurasi (jika ada) berarti bahwa prediksi kelas akhirnya (misalnya, "Green", "Orange", atau "Red") pada ambang batas probabilitas standar sedikit lebih baik. Namun, keunggulan RF pada metrik ROC-AUC dan AP menunjukkan bahwa skor probabilitas yang dihasilkannya lebih terkalibrasi dengan baik. Artinya, RF lebih andal dalam menyusun peringkat mahasiswa berdasarkan tingkat risiko,

karena model ini secara konsisten memberikan skor probabilitas yang lebih tinggi untuk kasus positif yang sebenarnya dibandingkan dengan kasus negatif.

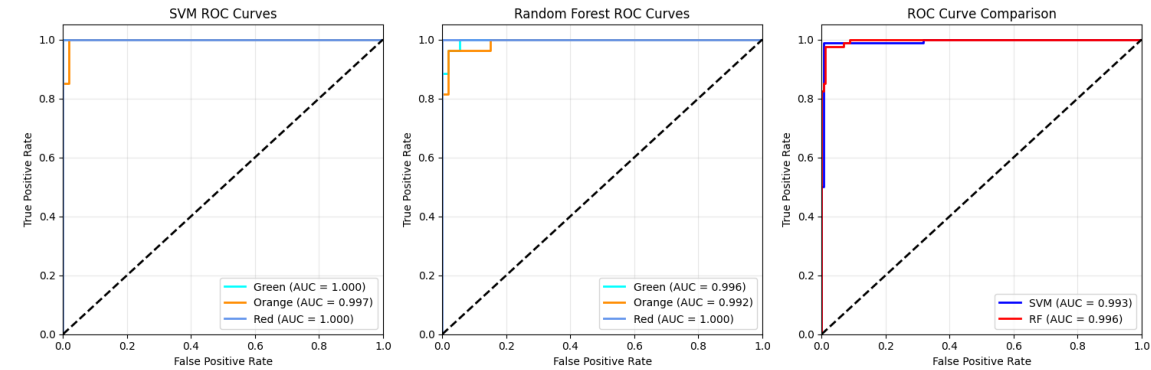
Implikasi praktis dari perbedaan ini adalah pemilihan model yang optimal bergantung pada tujuan akhir aplikasi. Untuk tujuan klasifikasi kategoris secara langsung, seperti menentukan status *burnout* mahasiswa, model SVM berpotensi menunjukkan keunggulan. Sebaliknya, jika tujuannya adalah untuk melakukan stratifikasi risiko misalnya menyusun daftar prioritas untuk intervensi dengan mengurutkan mahasiswa dari risiko tertinggi ke terendah maka model RF menghasilkan peringkat yang lebih reliabel. Perbedaan subtil ini menjadi pertimbangan krusial dalam justifikasi pemilihan model yang paling sesuai dengan kebutuhan implementasi.



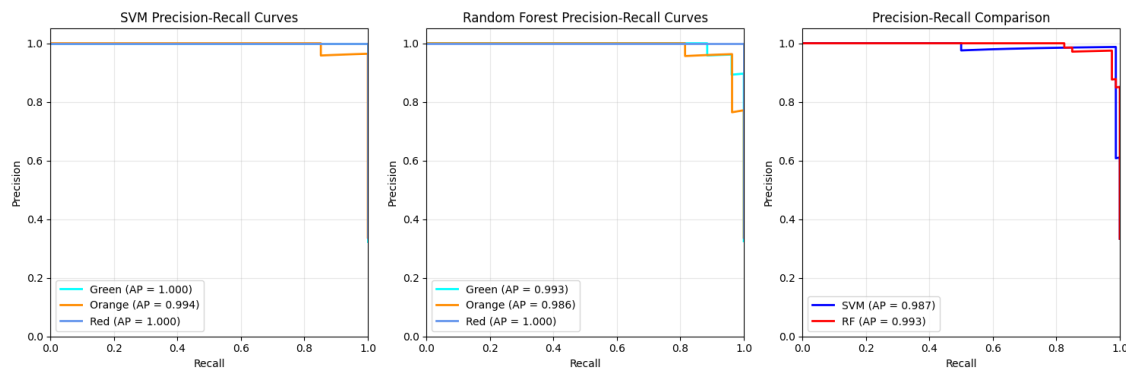
Gambar 4. Perbandingan Test Accuracy



Gambar 5. Perbandingan Precision, Recall, dan F1-Score



Gambar 6. Perbandingan Kurva ROC Micro-Average.



Gambar 7. Perbandingan Kurva *Precision-Recall Micro-Average*.

3.5. Pembahasan

Penelitian ini menghasilkan dua model prediktif dengan kinerja sangat tinggi, namun merekomendasikan SVM sebagai model final. Rekomendasi ini tidak hanya didasarkan pada keunggulan marginal dalam metrik akurasi, tetapi pada analisis kualitatif yang mendalam terhadap jenis dan dampak kesalahan. Keunggulan fundamental SVM dalam konteks aplikasi ini terletak pada profil risiko kesalahannya yang secara klinis lebih dapat diterima. SVM tidak menghasilkan *false negative* berisiko tinggi, yaitu kesalahan mengklasifikasikan mahasiswa yang memerlukan bantuan sebagai tidak berisiko. Satu-satunya kesalahan yang dibuat SVM adalah menurunkan tingkat risiko dari "tinggi" ke "sedang", yang dalam skenario praktis masih akan membuat mahasiswa tersebut berada dalam radar sistem intervensi. Sebaliknya, RF tercatat melakukan kesalahan tipe *false negative* yang paling kritis, yang dalam aplikasi dunia nyata dapat menyebabkan seorang mahasiswa terlewat dari sistem pendukung.

Tingginya performa RF dalam penelitian ini selaras dengan temuan konsisten dalam literatur. Kekuatan RF sebagai metode ensemble yang mengagregasi prediksi dari ratusan pohon keputusan membuatnya sangat tangguh dan stabil [18]. Hal ini didemonstrasikan secara empiris oleh *Lara-Gonzalez et al.*, yang menyimpulkan bahwa RF (akurasi 95,9%) secara signifikan melampaui SVM (91,5%) untuk proyeksi *burnout* [17]. Dukungan serupa datang dari *Pereira et al.*, di mana dalam analisis profil risiko *burnout*, RF (95,06%) juga menunjukkan kinerja lebih baik dibandingkan SVM (93,82%) [19]. Lebih lanjut, penelitian oleh *Nurdiansyah et al.* menyoroti keunggulan lain dari RF, yaitu stabilitas dan konsistensi kinerjanya di berbagai skenario pembagian data, sebuah atribut krusial untuk model yang akan diimplementasikan pada lingkungan produksi [20]. Kemampuan RF dalam menangani hubungan non-linear yang kompleks, sebuah ciri khas data psikologis [21], semakin memperkuat posisinya sebagai algoritma yang andal.

Meskipun demikian, rekomendasi SVM dalam penelitian ini bukanlah sebuah anomali dan didukung oleh studi lain yang menggarisbawahi prinsip bahwa tidak ada algoritma tunggal yang superior secara universal; pilihan optimal sangat bergantung pada karakteristik *dataset* dan tujuan spesifik aplikasi. Dukungan kuat untuk temuan penelitian ini datang dari studi oleh *Villar dan de Andrade*. Dalam penelitian mereka yang berfokus pada prediksi mahasiswa putus kuliah, sebuah konteks yang relevan dengan tekanan psikologis, disimpulkan bahwa SVM sering kali menjadi model dengan performa terbaik pada *dataset* asli sebelum rekayasa fitur atau penyeimbangan kelas [22]. Hal ini sejalan dengan temuan *Ahuja dan Banga* yang juga mencatat akurasi tertinggi pada SVM (85,71%) dibandingkan RF dalam mendeteksi stres pada mahasiswa [15]. Studi lain oleh *Mohd Shafiee dan Mutalib* juga menunjukkan bahwa SVM termasuk algoritma dengan akurasi tinggi (di atas 70%) untuk prediksi masalah kesehatan mental pada mahasiswa [23]. Dinamika ini menegaskan bahwa performa yang sangat tinggi dari kedua model dalam penelitian ini (di atas 97%) mengindikasikan keduanya adalah kandidat yang sangat valid, dan perbedaan tipis di antara keduanya lebih mencerminkan nuansa pada *dataset* yang digunakan daripada superioritas absolut salah satu algoritma [24].

Meskipun hasil yang diperoleh sangat menjanjikan, penelitian ini memiliki beberapa keterbatasan yang membuka peluang untuk pengembangan di masa depan. Pertama, terkait interpretasi model, penelitian ini berfokus pada kinerja prediktif, namun belum mengimplementasikan analisis kepentingan fitur (*feature importance*). Hal ini menjadi langkah krusial berikutnya untuk membuat hasil lebih interpretatif bagi psikolog dan akademisi. Dengan menggunakan teknik seperti SHAP (*SHapley Additive exPlanations*) atau Gini Importance pada model RF, penelitian selanjutnya dapat mengidentifikasi dan mengkuantifikasi faktor-faktor (misalnya, *Psychological Distress* atau *Exhaustion*) yang paling berpengaruh terhadap prediksi *burnout*, sehingga memberikan wawasan yang dapat ditindaklanjuti untuk intervensi.

Kedua, mengenai validitas dan potensi *overfitting*, model ini divalidasi hanya menggunakan data dari satu universitas. Untuk menilai robustisitas dan kemampuan generalisasinya, uji validasi eksternal dengan

data dari universitas berbeda sangat diperlukan. Meskipun akurasi tinggi pada data uji (98.75% untuk SVM) menunjukkan risiko *overfitting* yang rendah, potensi ini tetap ada, terutama karena akurasi pelatihan SVM mencapai 100%. Untuk mengatasinya secara lebih komprehensif, penelitian mendatang dapat menerapkan teknik regularisasi yang lebih lanjut atau menggunakan *dataset* lintas institusi (*cross-institutional dataset*) yang lebih beragam.

Ketiga, cakupan perbandingan algoritma perlu diperluas. Penelitian ini terbatas pada SVM dan RF. Mengingat tren saat ini dalam literatur ML, perluasan perbandingan dengan algoritma lain seperti Gradient Boosting (misalnya, XGBoost, LightGBM), atau bahkan pendekatan Deep Learning, sangat esensial. Sebagaimana ditunjukkan oleh studi Pereira et al. (2025), XGBoost berpotensi menawarkan kombinasi optimal antara akurasi prediktif yang lebih tinggi dan profil keamanan yang superior, sehingga validasi terhadap algoritma ini dapat memastikan pilihan model yang benar-benar optimal.

Selanjutnya, penggunaan ML untuk memprediksi kondisi psikologis mahasiswa membawa implikasi etis dan kebijakan yang signifikan. Dari perspektif etis, isu utama adalah penggunaan data psikologis mahasiswa yang sensitif. Implementasi sistem ini harus didasari oleh protokol yang ketat, termasuk perolehan persetujuan yang jelas (*informed consent*), anonimisasi data, dan kebijakan keamanan data yang transparan. Sangat penting untuk menekankan bahwa model ini harus berfungsi sebagai alat pendukung keputusan bagi para profesional (konselor, psikolog, dosen wali), bukan sebagai pengganti diagnosis klinis otomatis, untuk menghindari risiko stigmatisasi. Dari sisi kebijakan kampus, temuan ini dapat menjadi dasar bagi sistem deteksi dini yang proaktif, berfokus pada pencegahan dan intervensi suportif dengan menawarkan sumber daya dukungan secara tepat waktu kepada mahasiswa yang teridentifikasi berisiko.

4. KESIMPULAN

Berdasarkan hasil analisis komprehensif, dapat disimpulkan bahwa kedua algoritma, SVM dan RF, menunjukkan kinerja prediktif yang sangat tinggi dalam mengklasifikasikan *academic burnout* pada mahasiswa, dengan akurasi pengujian masing-masing sebesar 98,75% dan 97,50%. Meskipun terdapat perbedaan marginal pada metrik kuantitatif, perbedaan kinerja yang paling signifikan terletak pada analisis kualitatif terhadap jenis kesalahan klasifikasi. Model SVM direkomendasikan sebagai model akhir untuk implementasi praktis karena profil kesalahannya secara klinis lebih aman dan dapat diterima, dengan tidak menghasilkan *false negative* yang berisiko tinggi. Penelitian ini berhasil membuktikan bahwa ML dapat menjadi alat yang efektif untuk pengembangan sistem deteksi dini risiko *burnout*. Untuk penelitian di masa depan, disarankan untuk melakukan analisis kepentingan fitur (*feature importance*) guna mengidentifikasi faktor-faktor prediktor utama serta memperluas perbandingan dengan menyertakan algoritma *ensemble* yang lebih canggih seperti gradient boosting.

REFERENSI

- [1] Z. Liu, Y. Xie, Z. Sun, D. Liu, H. Yin, and L. Shi, "Factors associated with academic burnout and its prevalence among university students: a cross-sectional study," *BMC Med. Educ.*, vol. 23, no. 1, Dec. 2023, doi: 10.1186/s12909-023-04316-y.
- [2] J. G. Greener, S. M. Kandathil, L. Moffat, and D. T. Jones, "A guide to machine learning for biologists," *Nat. Rev. Mol. Cell Biol.*, vol. 23, no. 1, pp. 40–55, 2022, doi: 10.1038/s41580-021-00407-0.
- [3] A. Frajerman, Y. Morvan, M. O. Krebs, P. Gorwood, and B. Chaumette, "Burnout in medical students before residency: A systematic review and meta-analysis," *European Psychiatry*, vol. 55. Elsevier Masson SAS, pp. 36–42, Jan. 01, 2019, doi: 10.1016/j.eurpsy.2018.08.006.
- [4] J. Marôco et al., "Predictors of academic efficacy and dropout intention in university students: Can engagement suppress burnout?," *PLoS One*, vol. 15, no. 10 October, Oct. 2020, doi: 10.1371/journal.pone.0239816.
- [5] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, "Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13. Institute of Electrical and Electronics Engineers Inc., pp. 6308–6325, 2020, doi: 10.1109/JSTARS.2020.3026724.
- [6] S. Yu Hwanjo and Kim, "SVM Tutorial — Classification, Regression and Ranking," in *Handbook of Natural Computing*, T. and K. J. N. Rozenberg Grzegorz and Bäck, Ed. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 479–506.
- [7] Y. Zhang, Y. Liu, and X. Yang, "Review of Support Vector Machine Theory and Application Research," *Int. Core J. Eng.*, 2021, doi: 10.6919/ICJE.202106_7(6).0049.
- [8] G. Biau and E. Scornet, "A random forest guided tour," *Test*, vol. 25, no. 2, pp. 197–227, 2016, doi: 10.1007/s11749-016-0481-7.
- [9] J. Zhang, W. Yang, X. Luo, and L. Wang, "The Relationship Between Stress and Academic Burnout in College Students: Evidence from Longitudinal Data on Indirect Effects," *Front. Psychol.*, vol. 16, p.

- 1517920, 2025.
- [10] A. M. Shimaoka, R. C. Ferreira, and A. Goldman, "The evolution of CRISP-DM for Data Science: Methods, Processes and Frameworks," *SBC Rev. Comput. Sci.*, vol. 4, no. 1, pp. 28–43, 2024, doi: 10.5753/reviews.2024.3757.
 - [11] Fraunhofer IST, "Implementation strategy for a data-mining project with CRISP-DM in surface technology," 2025. .
 - [12] M. Carmona-Halty, K. Alarcón-Castillo, C. Semir-González, G. Sepúlveda-Páez, and W. B. Schaufeli, "Burnout Assessment Tool for Students (BAT-S): evidence of validity in a Chilean sample of undergraduate university students," *Front. Psychol.*, vol. 15, 2024, doi: 10.3389/fpsyg.2024.1434412.
 - [13] Muh Haritsah Hizbullah *et al.*, "Analisis Psikometri Burnout Assessment Tool (BAT) For Student Versi Bahasa Indonesia," *J. Penelit. dan Pengukuran Psikol. JPPP*, vol. 14, no. 1, pp. 30–38, Apr. 2025, doi: 10.21009/jppp.141.04.
 - [14] I. Malashin, V. Tynchenko, A. Gantimurov, V. Nelyub, and A. Borodulin, "Support Vector Machines in Polymer Science: A Review," *Polymers*, vol. 17, no. 4. Multidisciplinary Digital Publishing Institute (MDPI), Feb. 01, 2025, doi: 10.3390/polym17040491.
 - [15] R. Ahuja and A. Banga, "Mental stress detection in university students using machine learning algorithms," in *Procedia Computer Science*, 2019, vol. 152, pp. 349–353, doi: 10.1016/j.procs.2019.05.007.
 - [16] R. Rois, M. Ray, A. Rahman, and S. K. Roy, "Prevalence and predicting factors of perceived stress among Bangladeshi university students using machine learning algorithms," *J. Heal. Popul. Nutr.*, vol. 40, no. 1, Dec. 2021, doi: 10.1186/s41043-021-00276-5.
 - [17] L. R. Lara-González, M. A. Delgado-Luna, B. E. De León-Galván, and J. C. Venegas-Guerrero, "Comparison of Machine Learning algorithms for the Burnout projection," *ECORFAN Journal-Democratic Repub. Congo*, pp. 1–8, Jun. 2021, doi: 10.35429/ejdr.2021.12.7.1.8.
 - [18] R. C. Chen, C. Dewi, S. W. Huang, and R. E. Caraka, "Selecting critical features for data classification based on machine learning methods," *J. Big Data*, vol. 7, no. 1, Dec. 2020, doi: 10.1186/s40537-020-00327-4.
 - [19] M. G. Pereira, M. Santos, R. Magalhães, C. Rodrigues, O. Araújo, and D. Durães, "Burnout Risk Profiles in Psychology Students: An Exploratory Study with Machine Learning," *Behav. Sci. (Basel)*, vol. 15, no. 4, Apr. 2025, doi: 10.3390/bs15040505.
 - [20] N. Nurdiansyah, F. S. Febriyan, Z. G. D. Amanta, D. A. Saputra, and W. M. Baihaqi, "Analisis Kesehatan Mental untuk Mencegah Gangguan Mental pada Mahasiswa Menggunakan Algoritma K-Nearest Neighbor (K-NN) dan Random Forest," *MALCOM Indones. J. Mach. Learn. Comput. Sci.*, vol. 5, no. 1, pp. 1–9, Nov. 2024, doi: 10.57152/malcom.v5i1.1537.
 - [21] M. Grządzielewska, "Using Machine Learning in Burnout Prediction: A Survey," *Child and Adolescent Social Work Journal*, vol. 38, no. 2. Springer, pp. 175–180, Apr. 01, 2021, doi: 10.1007/s10560-020-00733-w.
 - [22] A. Villar and C. R. V. de Andrade, "Supervised machine learning algorithms for predicting student dropout and academic success: a comparative study," *Discov. Artif. Intell.*, vol. 4, no. 1, Dec. 2024, doi: 10.1007/s44163-023-00079-z.
 - [23] N. S. Mohd Shafiee and S. Mutalib, "Prediction of Mental Health Problems among Higher Education Student Using Machine Learning," *Int. J. Educ. Manag. Eng.*, vol. 10, no. 6, pp. 1–9, Dec. 2020, doi: 10.5815/ijeme.2020.06.01.
 - [24] P. A. Novianti, W. Yudiana, and S. Novita, "Can Learning Agility Predict Students' Academic Burnout During Distance Learning," *J. Educ. Heal. Community Psychol.*, vol. 1, no. 1, p. 17, Apr. 2023, doi: 10.12928/jehcp.v1i1.25098.