

Prediction of Academic Performance of Student Cadets at the Republic of Indonesia Defense University Using Random Forest

Triananda Marsya Harahap¹⁾, Muhammad Zidan Hafidz Ervia Nur Hakim²⁾, Nadiza Lediwara^{3*)}, Aulia Khamas Heikhmakhtiar⁴⁾, Sembada Denrineksa Bimorogo⁵⁾

¹⁾²⁾³⁾⁴⁾⁵⁾Informatika, Universitas Pertahanan Republik Indonesia

^{*)}Correspondence author: nadizalediwara@gmail.com, Kabupaten Bogor, Indonesia

DOI: <https://doi.org/10.37012/jtik.v12i2.3795>

Abstract

The academic readiness of cadets at the Indonesian Defense University (UNHAN RI) needs proactive monitoring, as cadets carry both academic workloads and military discipline simultaneously. This study aims to build a classification model for the academic performance of UNHAN RI cadets into three categories, Low, Moderate, and High, using the Random Forest algorithm within the CRISP-DM framework. Data were collected through an online questionnaire distributed to 100 cadet respondents from cohorts 5 and 6. After data cleaning, 95 valid records were obtained. The academic performance label was constructed from academic grades using a mean $\pm$ standard deviation categorization method, yielding a distribution of 20.00% Low, 57.89% Moderate, and 22.11% High. Six daily-habit variables, namely study hours, assignment load, sleep hours, exercise frequency, academic fatigue level, and perceived task burden, were used as predictors without including the grade itself, to avoid data leakage. Data were split into 80% training and 20% test sets using stratified sampling. A Random Forest model with 100 decision trees achieved 57.89% accuracy on the test set and an average of 55.79% across 5-fold cross-validation. Study hours, exercise frequency, and assignment load were identified as the three most influential features. An experiment removing the least important feature, perceived task burden, produced comparable accuracy and cross-validation mean but a higher cross-validation standard deviation, so the original six features were retained in the final model. The model's accuracy falls into the Poor category under common interpretation guidelines and is close to the majority-class baseline, indicating that daily-habit data alone is not yet sufficient to reliably predict academic performance at this limited sample size. The model is intended as a starting point for the institution to develop a more mature early-detection system for cadets' academic performance.

Keywords: Academic Performance, Cadet Students, Classification, Random Forest, UNHAN RI.

Abstrak

Kesiapan akademik kadet mahasiswa Universitas Pertahanan Republik Indonesia (UNHAN RI) perlu dipantau secara proaktif karena kadet menjalankan beban akademik dan kedisiplinan militer secara bersamaan. Penelitian ini bertujuan membangun model klasifikasi performa akademik kadet mahasiswa UNHAN RI ke dalam tiga kategori, yaitu Rendah, Sedang, dan Tinggi, menggunakan algoritma Random Forest dalam kerangka kerja CRISP-DM. Data dikumpulkan melalui kuesioner daring terhadap 100 responden kadet cohort 5 dan 6. Setelah pembersihan data, diperoleh 95 data valid. Label performa akademik dibentuk dari nilai akademik menggunakan metode kategorisasi mean $\pm$ standar deviasi, menghasilkan proporsi 20,00% Rendah, 57,89% Sedang, dan 22,11% Tinggi. Enam variabel kebiasaan harian, yaitu jam belajar, jumlah tugas, jam tidur, frekuensi olahraga, tingkat kelelahan akademik, dan persepsi beban tugas, digunakan sebagai prediktor tanpa melibatkan nilai akademik itu sendiri untuk menghindari kebocoran data. Data dibagi menjadi 80% data latih dan 20% data uji secara stratifikasi. Model Random Forest dengan 100 pohon keputusan menghasilkan akurasi 57,89% pada data uji dan rata-rata 55,79% pada validasi silang 5-fold. Jam belajar, frekuensi olahraga, dan jumlah tugas teridentifikasi sebagai tiga fitur paling berpengaruh. Eksperimen penghapusan fitur dengan importance terkecil, yaitu persepsi beban tugas, menghasilkan akurasi dan rata-rata cross-validation yang setara namun standar deviasi cross-validation yang lebih tinggi, sehingga keenam fitur awal dipertahankan sebagai model akhir. Akurasi model tergolong kategori Kurang menurut panduan interpretasi umum dan mendekati baseline kelas mayoritas, yang mengindikasikan bahwa kebiasaan harian saja belum cukup untuk memprediksi performa akademik secara andal pada ukuran sampel yang terbatas ini. Model ini diharapkan menjadi titik awal bagi institusi dalam mengembangkan sistem deteksi dini performa akademik kadet yang lebih matang.

Kata Kunci: Kadet Mahasiswa, Klasifikasi, Performa Akademik, Random Forest, UNHAN RI.

PENDAHULUAN

Performa akademik merupakan salah satu indikator utama keberhasilan proses pembelajaran, sehingga pemantauannya menjadi perhatian penting dalam pengelolaan pendidikan tinggi. Selama ini, pemantauan performa akademik cenderung bersifat reaktif baru dilakukan setelah nilai akhir terbit sehingga upaya intervensi bagi mahasiswa yang berisiko performa rendah kerap terlambat dilakukan. Bidang educational data mining sendiri tumbuh pesat sepanjang dekade terakhir, mencakup ragam tugas prediksi dari performa akademik hingga retensi mahasiswa (Bakhshinategh et al., 2018).

Di sinilah data mining menawarkan alternatif yang lebih proaktif: dengan mengolah data kebiasaan sehari-hari seperti jam belajar, jam tidur, frekuensi olahraga, dan beban tugas, model prediktif dapat dibangun untuk mengenali individu berisiko bahkan sebelum nilai akhir keluar. Pendekatan prediktif semacam ini relevan diterapkan pada permasalahan performa akademik karena memungkinkan deteksi dini tanpa harus menunggu hasil evaluasi formal, sehingga intervensi dapat dilakukan lebih awal.

Penelitian mengenai prediksi performa akademik mahasiswa melalui pendekatan komputasional berkembang cukup pesat dalam beberapa tahun terakhir. Random Forest secara konsisten dilaporkan sebagai salah satu algoritma dengan performa kompetitif untuk memprediksi capaian akademik mahasiswa dibandingkan algoritma tunggal lain, baik dikombinasikan dengan XGBoost maupun optimizer tambahan (Chen & Jin, 2024; Kumar et al., 2024). Studi berskala besar turut mengonfirmasi peran variabel non-akademik, seperti jam belajar dan jam tidur, sebagai prediktor performa akademik pada populasi mahasiswa perguruan tinggi (Ahmed, 2024; Murtaza & Bhaumik, 2025). Secara lebih spesifik, jam tidur dan waktu bangun tidur terbukti berasosiasi dengan performa akademik mahasiswa, baik pada populasi mahasiswa di Indonesia maupun pada periode ujian di perguruan tinggi lain (Sofyana et al., 2022; Suardiaz-Muro et al., 2023), sementara frekuensi dan keteraturan aktivitas fisik turut dilaporkan berasosiasi dengan capaian akademik mahasiswa melalui pendekatan data mining (Du et al., 2023). Pada konteks penerapan spesifik, kombinasi Random Forest dengan teknik optimisasi hyperparameter seperti Bayesian optimization juga terbukti mampu meningkatkan akurasi klasifikasi performa akademik pada dataset mahasiswa yang relatif kecil (Begum & Padmannavar, 2023). Selain untuk klasifikasi performa akademik, Random Forest juga banyak diterapkan pada permasalahan prediksi pendidikan lain, seperti prediksi putus studi (Devi & Ratnoo, 2022) dan prediksi keterlibatan

mahasiswa pada lingkungan pembelajaran daring dengan penanganan ketidakseimbangan kelas melalui teknik data balancing (Jawad et al., 2022), yang relevan dengan karakteristik distribusi kelas performa akademik pada penelitian ini yang juga tidak seimbang.

Merujuk pada rekam jejak penelitian tersebut, Random Forest dipilih sebagai algoritma utama pada penelitian ini bukan karena diklaim selalu unggul dibandingkan algoritma lain, melainkan karena karakteristiknya yang sesuai dengan kondisi data penelitian, yaitu ketahanannya terhadap data berjumlah sampel terbatas, kemampuannya menangani fitur campuran (numerik dan ordinal terencode), serta kemampuannya menghasilkan ukuran feature importance yang mudah diinterpretasikan oleh pihak akademik (Bellaj et al., 2024).

Sebagai institusi pendidikan tinggi kedinasan, UNHAN RI (Universitas Pertahanan Republik Indonesia) mewajibkan setiap mahasiswa atau dalam isitilah di UNHAN RI yaitu kadet mahasiswa, mereka menjalani dua peran sekaligus: mengejar capaian akademik berupa nilai, tugas, dan ujian, sekaligus memenuhi tuntutan disiplin militer dan latihan fisik yang berjalan nyaris setiap hari. Kondisi beban ganda semacam ini berisiko menekan performa akademik kadet, baik lewat jalur langsung berupa terpotongnya waktu belajar, maupun jalur tidak langsung berupa kelelahan fisik dan terganggunya waktu istirahat pola yang juga teramati pada kadet militer di negara lain, di mana intensitas latihan yang tinggi terkait dengan menurunnya kualitas tidur, meningkatnya kelelahan, dan terhambatnya proses pemulihan tubuh (Bulmer et al., 2022).

Berdasarkan uraian tersebut, penelitian ini secara khusus mengarahkan fokus kajian pada performa akademik kadet mahasiswa UNHAN RI. Research gap yang diisi oleh penelitian ini terletak pada dua hal. Pertama, sebagian besar studi performa akademik sejenis menggunakan variabel yang berkorelasi kuat dengan nilai (misalnya nilai semester sebelumnya) sebagai prediktor, sehingga rentan terhadap kebocoran data (data leakage) apabila label juga dibentuk dari nilai. Penelitian ini secara eksplisit mengeluarkan variabel nilai akademik dari himpunan fitur prediktor karena variabel tersebut menjadi dasar pembentukan label, sehingga model murni diuji kemampuannya memprediksi performa dari kebiasaan harian semata. Kedua, penelitian ini menyasar populasi spesifik kadet mahasiswa institusi pendidikan tinggi militer, yang karakteristik beban aktivitasnya berbeda dari mahasiswa sipil pada umumnya, sehingga hasilnya melengkapi literatur yang sebagian besar berbasis populasi mahasiswa sipil (Ahmed, 2024; Begum & Padmannavar, 2023; Chen & Jin, 2024; Kumar et al., 2024; Murtaza & Bhaumik, 2025).

Pendekatan klasifikasi berbasis data mining semacam ini telah lazim diterapkan pada berbagai konteks pengambilan keputusan di Indonesia, termasuk pada bidang non-akademik seperti penentuan kelayakan pendonor darah (Coastera et al., 2020), serta sejalan dengan tradisi penelitian terapan berbasis data di lingkungan Universitas Pertahanan RI, misalnya pada bidang pemetaan dan teknologi pertahanan (Gemilang Gultom et al., 2023). Dengan demikian, penelitian ini memberikan kontribusi berupa penerapan pendekatan berbasis data pada isu performa akademik yang belum banyak dikaji secara khusus pada populasi kadet mahasiswa, sekaligus melengkapi rekam jejak penelitian terapan berbasis data yang telah berkembang di lingkungan UNHAN RI pada bidang lain.

Penelitian ini menerapkan Random Forest dalam kerangka kerja CRISP-DM untuk membangun model klasifikasi performa akademik kadet mahasiswa UNHAN RI, sekaligus menyediakan analisis feature importance dan eksperimen feature selection sebagai dasar evaluasi kontribusi masing-masing variabel kebiasaan harian terhadap performa akademik.

METODE PENELITIAN

Secara metodologis, penelitian ini berbentuk penelitian kuantitatif yang mengikuti tahapan kerangka kerja CRISP-DM (Cross-Industry Standard Process for Data Mining) mulai dari pemahaman bisnis, pemahaman data, persiapan data, pemodelan, hingga evaluasi (Handoko et al., 2025). Pada tahap pemahaman bisnis, tujuan yang ditetapkan adalah mengelompokkan performa akademik kadet mahasiswa ke dalam tiga kategori, sebagai instrumen deteksi dini yang dapat digunakan institusi sebelum nilai akhir semester diumumkan. Alur penelitian secara keseluruhan mengikuti tahapan: pengumpulan data → pra-pemrosesan data → pembentukan label → pembagian data latih/uji → pemodelan Random Forest → evaluasi model → analisis feature importance → eksperimen feature selection → evaluasi ulang.

Data Penelitian

Pengumpulan data pada penelitian ini dilakukan melalui kuesioner daring berbasis Google Form, yang disebarkan kepada kadet mahasiswa UNHAN RI dari cohort 5 dan 6, sehingga terkumpul 100 responden secara keseluruhan. Variabel yang diukur meliputi jam belajar per hari, jumlah tugas per minggu, nilai rata-rata akademik (skala 0–100), jam tidur per hari, frekuensi olahraga per minggu, frekuensi kelelahan akademik (skala ordinal lima poin), dan persepsi beban tugas (skala ordinal tiga poin). Penelitian ini menggunakan nilai akademik

sebagai dasar pembentukan label performa akademik, sehingga variabel tersebut tidak diikutsertakan sebagai prediktor guna menghindari kebocoran data (data leakage).

Pra-pemrosesan Data

Pra-pemrosesan data dilakukan melalui beberapa tahapan berurutan sebelum pembentukan label, sebagaimana dirangkum pada Tabel 1.

Tabel 1.
Tahapan Pra-pemrosesan Data

No	Tahapan	Kriteria/Deskripsi
1	Pemeriksaan usia	Menghapus entri dengan usia ≥ 100 tahun (dianggap tidak wajar untuk populasi mahasiswa)
2	Penanganan nilai kosong	Menghapus baris dengan nilai kosong pada kolom fitur utama (JamBelajar, Tugas, Nilai, Tidur, Olahraga)
3	Pemeriksaan skala nilai akademik	Menghapus entri dengan nilai akademik ≤ 10 (terindikasi menggunakan skala IPK, tidak konsisten dengan skala 0–100 yang menjadi acuan)
4	Encoding variabel ordinal	Kelelahan akademik dikonversi ke skala numerik 0 (Tidak pernah) – 4 (Sangat sering); persepsi beban tugas dikonversi ke skala numerik 0 (Tidak) – 2 (Ya)

Setelah pra-pemrosesan, diperoleh 95 data bersih dari 100 data awal, dengan rincian 1 baris dihapus karena usia tidak wajar dan 4 baris dihapus karena nilai akademik berskala IPK; tidak ditemukan nilai kosong pada kolom fitur.

Pembentukan Label Performa Akademik

Dataset kuesioner tidak menyediakan label performa akademik secara langsung, sehingga label dibentuk dari variabel nilai akademik menggunakan metode kategorisasi mean $\pm$ standar deviasi (mean $\pm$ 1 SD), yang lazim digunakan dalam penelitian pendidikan dan psikometri Indonesia untuk mengubah skor kontinu menjadi kategori ordinal tiga tingkat (Azwar, 2021). Kadet dengan nilai lebih besar dari mean + 1 SD dikategorikan Performa Tinggi, kadet dengan nilai lebih kecil dari mean – 1 SD dikategorikan Performa Rendah, dan kadet dengan nilai di antara kedua batas tersebut dikategorikan Performa Sedang. Metode ini dipilih karena bersifat objektif (murni berdasar pada distribusi statistik data, bukan penilaian subjektif peneliti), dapat direproduksi, dan lazim dijadikan rujukan pada penelitian sejenis di lingkungan akademik Indonesia.

Pembagian Data

Pembagian dataset dilakukan dengan proporsi 80% data latih (76 sampel) dan 20% data uji (19 sampel) menggunakan teknik stratified split untuk menjaga proporsi setiap kategori performa akademik pada kedua subset.

Pemodelan Random Forest

Sebagai salah satu metode ensemble learning yang diperkenalkan pertama kali oleh Breiman (2001), Random Forest hingga saat ini masih banyak dipakai di berbagai bidang prediksi, termasuk untuk memprediksi performa akademik (Bellaj et al., 2024). Cara kerjanya adalah membangun banyak pohon keputusan secara independen pada tahap pelatihan, lalu menggabungkan seluruh hasil prediksi tersebut melalui mekanisme voting mayoritas ketika menyelesaikan tugas klasifikasi. Karena mengombinasikan banyak pohon sekaligus, pendekatan ensemble ini memiliki varians yang lebih rendah dibandingkan pohon keputusan tunggal, sehingga model menjadi lebih tahan terhadap overfitting meskipun jumlah sampel data terbatas.

Model diimplementasikan menggunakan pustaka Scikit-Learn pada bahasa Python (Scikit-learn Developers, 2026). Fitur yang digunakan sebagai prediktor adalah jam belajar, jumlah tugas per minggu, jam tidur, frekuensi olahraga, tingkat kelelahan (terencode), dan persepsi beban tugas (terencode) enam fitur kebiasaan harian, tanpa melibatkan nilai akademik. Konfigurasi model dan pembagian data dirangkum pada Tabel 2.

Tabel 2.
Konfigurasi Model Random Forest

Parameter	Nilai
Algoritma	RandomForestClassifier (Scikit-Learn)
n_estimators	100
random_state	42
n_jobs	-1
test_size	0,2 (20%)
stratify	Ya, berdasarkan label performa akademik
Jumlah data latih	76 sampel
Jumlah data uji	19 sampel

Model dilatih pada 76 sampel data latih, dievaluasi pada 19 sampel data uji, dan divalidasi lebih lanjut menggunakan 5-fold cross-validation untuk menilai stabilitas performa.

Evaluasi Model

Penilaian performa model bertumpu pada confusion matrix, yang menjadi dasar penghitungan metrik akurasi, presisi, recall, dan F1-Score. Kombinasi beberapa metrik ini diperlukan mengingat distribusi kelas pada dataset penelitian tidak seimbang; akurasi saja berisiko memberi gambaran yang menyesatkan tentang kinerja model, terutama untuk kelas dengan jumlah sampel minoritas (Han et al., 2022). Persamaan metrik evaluasi yang digunakan adalah sebagai berikut.

$$Akurasi = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

$$Presisi = TP / (TP + FP) \quad (2)$$

$$Recall = TP / (TP + FN) \quad (3)$$

$$F1-Score = 2 \times (Presisi \times Recall) / (Presisi + Recall) \quad (4)$$

dengan TP (True Positive) adalah sampel positif yang diprediksi benar, TN (True Negative) adalah sampel negatif yang diprediksi benar, FP (False Positive) adalah sampel negatif yang diprediksi positif, dan FN (False Negative) adalah sampel positif yang diprediksi negatif. Di luar evaluasi pada data uji, stabilitas performa model juga diuji melalui 5-fold cross-validation, yang memberikan gambaran lebih andal dibandingkan hanya mengandalkan satu kali evaluasi.

Feature Importance

Tingkat kontribusi setiap fitur terhadap prediksi model diperoleh melalui atribut feature importance bawaan Random Forest pada Scikit-Learn, yang mengukur kontribusi setiap fitur berdasarkan penurunan impurity rata-rata di seluruh pohon keputusan. Nilai feature importance keenam fitur kemudian diurutkan dari yang terbesar hingga terkecil untuk mengidentifikasi fitur dengan kontribusi paling rendah.

Eksperimen Feature Selection

Sebagai bagian dari evaluasi lanjutan, dilakukan eksperimen tambahan berupa penghapusan fitur dengan feature importance terkecil pada model awal, kemudian seluruh tahap pelatihan, pengujian, dan validasi silang dijalankan ulang menggunakan fitur yang tersisa. Model hasil eksperimen dibandingkan dengan model awal berdasarkan tiga kriteria, yaitu akurasi data uji, rata-rata akurasi cross-validation, dan standar deviasi cross-validation; model dengan performa lebih baik atau setara pada kriteria tersebut, dengan standar deviasi cross-

validation yang lebih kecil, dipilih sebagai model akhir penelitian. Hasil numerik dan pembahasan eksperimen ini disajikan pada bagian Hasil dan Pembahasan.

HASIL DAN PEMBAHASAN

Skenario Pra-pemrosesan Data

Proses pembersihan data dijalankan mengikuti tahapan pada Tabel 1 terhadap 100 data mentah hasil kuesioner. Ringkasan tahapan dan jumlah baris yang terdampak pada setiap langkah ditampilkan pada Tabel 3.

Tabel 3.
Ringkasan Tahapan Pra Proses Data

Tahapan	Deskripsi	Baris Terdampak
Data awal	Hasil kuesioner sebelum pra proses	100
Cleaning usia	Penghapusan entri dengan usia tidak wajar (≥ 100 tahun)	1 dihapus
Cleaning skala nilai	Penghapusan entri nilai akademik berskala IPK (≤ 10)	4 dihapus
Penanganan nilai kosong	Penghapusan baris dengan nilai kosong pada fitur	0 dihapus
Pembentukan label	Label Performa Akademik dibentuk dari nilai ($\text{mean} \pm 1 \text{ SD}$)	95 baris dilabel
Data bersih	Data final yang digunakan untuk pemodelan	95

Dari 100 data awal, lima baris dikeluarkan karena tidak memenuhi kriteria kebersihan data: satu baris akibat usia yang tidak wajar untuk populasi mahasiswa, dan empat baris karena nilai akademik dilaporkan pada skala IPK (0–4) alih-alih skala 0–100 yang menjadi acuan variabel Nilai. Tidak ditemukan nilai kosong pada kolom fitur, sehingga seluruh pengurangan data murni berasal dari dua kriteria cleaning tersebut.

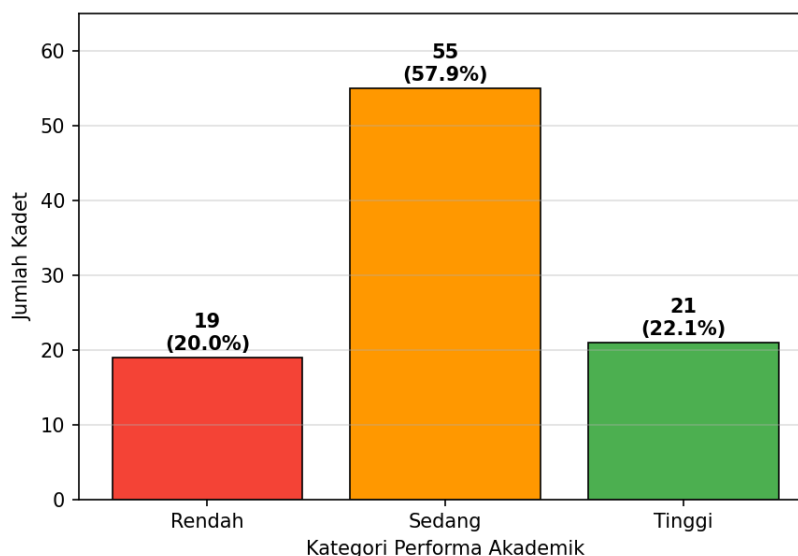
Deskripsi Dataset dan Pembentukan Label Performa Akademik

Nilai akademik pada 95 data bersih memiliki mean 88,19 dan standar deviasi 6,66, sehingga batas kategori Rendah adalah Nilai $< 81,53$ dan batas kategori Tinggi adalah Nilai $> 94,85$; kadet dengan nilai di antara kedua batas tersebut dikategorikan Sedang. Penerapan aturan ini menghasilkan distribusi label performa akademik sebagaimana ditampilkan pada Tabel 4 dan Gambar 1.

Tabel 4.

Distribusi Kelas Performa Akademik Setelah Pembentukan Label

Kategori	Kriteria Nilai	Jumlah Sampel	Persentase (%)
Rendah	< 81,53	19	20,00
Sedang	81,53 – 94,85	55	57,89
Tinggi	> 94,85	21	22,11
Total	-	95	100



Gambar 1. Distribusi Kategori Performa Akademik

Karakteristik variabel prediktor (kebiasaan harian) ditampilkan pada Tabel 5. Rata-rata jam belajar kadet adalah 6,07 jam per hari dengan rentang cukup lebar (1 hingga 18 jam), yang mengindikasikan variasi kebiasaan belajar yang tinggi antarresponden. Rata-rata jam tidur hanya 3,64 jam per hari, jauh di bawah rekomendasi kebutuhan tidur orang dewasa, konsisten dengan pola kehidupan kadet yang padat oleh kegiatan akademik dan kedisiplinan. Pada variabel ordinal, tingkat kelelahan akademik memiliki mean 2,44 dari skala 0–4 (mendekati kategori "Sering"), dengan mayoritas responden berada pada kategori Sering (40 orang) dan Kadang-kadang (26 orang), sedangkan persepsi beban tugas memiliki mean 0,89 dari skala 0–2, dengan mayoritas responden memilih kategori Kadang (51 orang). Kedua pola ini menegaskan bahwa beban akademik dan kelelahan dirasakan cukup intensif oleh sebagian besar kadet dalam sampel penelitian.

Tabel 5.
Statistik Deskriptif Variabel Prediktor

Variabel	Min	Maks	Mean	Std. Dev	Median
Jam Belajar/hari	1	18	6,07	3,28	6
Tugas/minggu	3	18	6,51	2,62	6
Jam Tidur/hari	1	8	3,64	1,19	4
Olahraga/minggu	1	14	4,82	2,46	5
Tingkat Kelelahan Akademik (skala ordinal 0–4)	0	4	2,44	1,04	3
Persepsi Beban Tugas (skala ordinal 0–2)	0	2	0,89	0,68	1

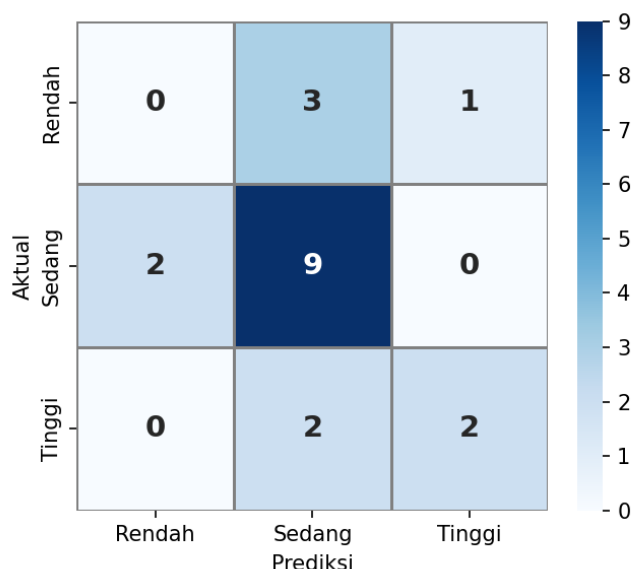
Kondisi dataset ini memiliki dua implikasi penting. Pertama, ukuran sampel yang relatif kecil (95 data) membatasi kemampuan model dalam mempelajari pola pada kelas minoritas, khususnya kategori Rendah yang hanya diwakili oleh 19 sampel. Kedua, karena label dibentuk langsung dari nilai akademik, variabel tersebut sengaja dikeluarkan dari himpunan fitur prediktor untuk menghindari kebocoran data (data leakage); konsekuensinya, model harus memprediksi performa akademik murni dari pola kebiasaan harian, yang secara inheren lebih menantang dibandingkan model yang masih memasukkan nilai sebagai salah satu prediktornya.

Evaluasi Model

Model Random Forest yang dilatih pada 76 data latih dan diuji pada 19 data uji menghasilkan akurasi sebesar 57,89%. Metrik presisi, recall, dan F1-Score per kelas ditampilkan pada Tabel 6 sebagai pelengkap gambaran performa model.

Tabel 6.
Hasil Klasifikasi Per Kelas (Data Uji) — Model Akhir (6 Fitur)

Kategori	Presisi	Recall	F1-Score	Support
Rendah	0,00	0,00	0,00	4
Sedang	0,64	0,82	0,72	11
Tinggi	0,67	0,50	0,57	4
Macro Avg	0,44	0,44	0,43	19
Weighted Avg	0,51	0,58	0,54	19



Gambar 2. Confusion Matrix Model pada Data Uji

Confusion matrix pada Gambar 2 menggambarkan distribusi prediksi model terhadap label aktual pada data uji, dengan baris mewakili kelas aktual dan kolom mewakili kelas hasil prediksi. Model tidak berhasil mengklasifikasikan satu pun sampel kategori Rendah dengan benar; dari 4 sampel kategori tersebut, 3 diprediksi sebagai kategori Sedang dan 1 diprediksi sebagai kategori Tinggi. Kondisi ini konsisten dengan proporsi sampel Rendah yang terbatas pada data latih, sehingga model tidak memiliki cukup pola pembelajaran untuk mengenalinya. Kategori Sedang dikenali dengan recall tertinggi (0,82), sedangkan kategori Tinggi hanya dikenali pada separuh sampelnya (recall 0,50).

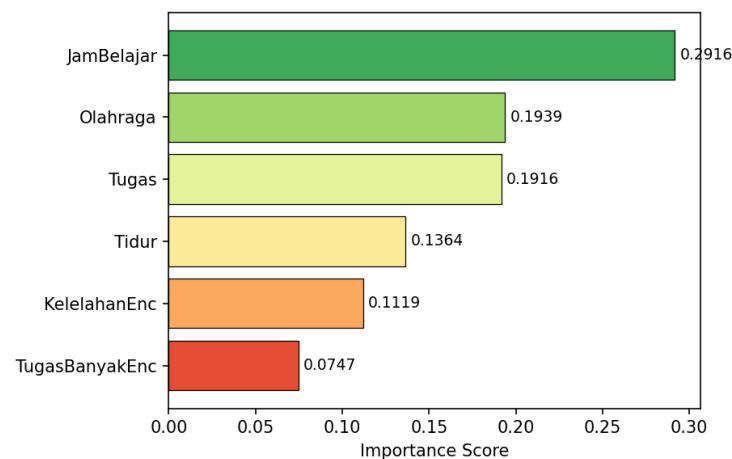
Feature Importance dan Eksperimen Feature Selection

Skor feature importance yang dihasilkan Random Forest menunjukkan kontribusi relatif setiap variabel terhadap keputusan klasifikasi, sebagaimana ditampilkan pada Tabel 7 dan Gambar 3. Jam belajar menjadi fitur paling berpengaruh (29,16%), diikuti oleh frekuensi olahraga (19,39%) dan jumlah tugas (19,16%), sementara persepsi beban tugas memberikan kontribusi terkecil (7,47%).

Tabel 7.

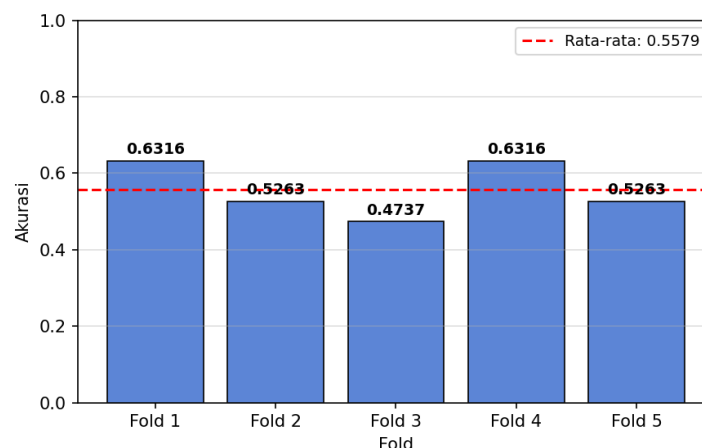
Feature Importance Model Awal (6 Fitur)

Peringkat	Fitur	Importance	Persentase (%)
1	Jam Belajar per Hari	0,2916	29,16
2	Frekuensi Olahraga	0,1939	19,39
3	Jumlah Tugas per Minggu	0,1916	19,16
4	Jam Tidur per Hari	0,1364	13,64
5	Tingkat Kelelahan Akademik	0,1119	11,19
6	Persepsi Beban Tugas	0,0747	7,47



Gambar 3. Feature Importance Model Awal

Validasi silang 5-fold dilakukan untuk menilai stabilitas performa model secara lebih andal dibandingkan evaluasi tunggal pada data uji. Hasil pada Gambar 4 menunjukkan akurasi rata-rata 55,79% dengan standar deviasi 6,32%, sedikit lebih rendah dibandingkan akurasi pada data uji (57,89%) namun berada pada rentang yang relatif konsisten, mengindikasikan bahwa performa model tidak bergantung secara berlebihan pada satu pembagian data tertentu, meskipun secara keseluruhan performanya masih terbatas.



Gambar 4. Akurasi Model pada Setiap Fold Cross-Validation

Mengacu pada Tabel 7, fitur dengan feature importance terkecil pada model awal adalah persepsi beban tugas (TugasBanyakEnc) dengan kontribusi 7,47%. Fitur tersebut dihapus dari himpunan prediktor, kemudian seluruh tahap pelatihan, pengujian, dan validasi silang dijalankan ulang menggunakan lima fitur tersisa. Prosedur evaluasi ulang ini dirangkum pada Tabel 8.

Tabel 8.
Prosedur Evaluasi Ulang dan Feature Importance

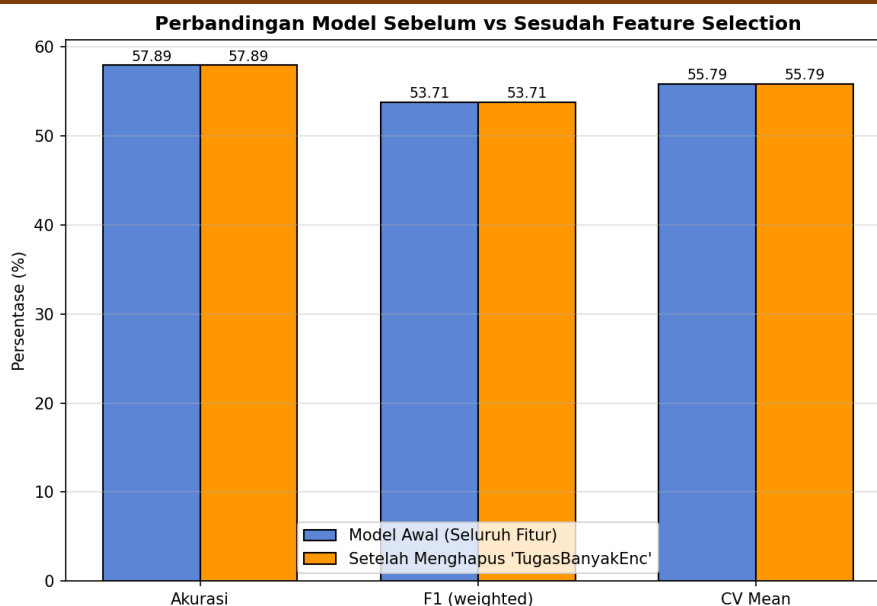
Langkah	Deskripsi
1	Memprediksi label pada data uji menggunakan model yang telah dilatih
2	Menghitung akurasi data uji dari hasil prediksi terhadap label aktual
3	Menghitung skor akurasi 5-fold cross-validation
4	Mengurutkan skor feature importance dari yang terbesar ke terkecil untuk mengidentifikasi fitur berkontribusi paling rendah

Sumber: Diadaptasi dari kode sumber Python, 2026

Hasil perbandingan kedua model ditampilkan pada Tabel 9. Model baru dengan lima fitur menghasilkan akurasi data uji dan F1-Score (weighted) yang identik dengan model awal, yaitu 57,89% dan 53,71%. Akan tetapi, standar deviasi cross-validation model baru (8,55%) lebih tinggi dibandingkan model awal (6,32%), sedangkan rata-rata akurasi cross-validation kedua model sama-sama sebesar 55,79%. Perbandingan visual kedua model, sebagaimana disajikan pada Gambar 5, memperkuat pola yang sama: akurasi dan F1-Score kedua model relatif setara, sementara variabilitas performa antarfold pada model lima fitur lebih tinggi. Karena tidak ada peningkatan performa yang diperoleh dari penghapusan fitur, sementara stabilitas model justru menurun (ditunjukkan oleh standar deviasi cross-validation yang lebih tinggi), keenam fitur awal dipertahankan sebagai model akhir penelitian.

Tabel 9.
Perbandingan Model Awal dan Model Setelah Penghapusan Fitur

Metrik	Model Awal (6 Fitur)	Model Baru (5 Fitur, tanpa Persepsi Beban Tugas)
Akurasi Data Uji	57,89%	57,89%
Presisi (Weighted)	51,25%	51,25%
Recall (Weighted)	57,89%	57,89%
F1-Score (Weighted)	53,71%	53,71%
F1-Score (Macro)	43,05%	43,05%
Rata-rata Akurasi CV 5-Fold	55,79%	55,79%
Standar Deviasi CV 5-Fold	6,32%	8,55%
Keputusan	Digunakan sebagai model akhir	Tidak digunakan



Gambar 5. Perbandingan Akurasi, F1-Score, dan Rata-rata CV Sebelum dan Sesudah Feature Selection

Berdasarkan hasil tersebut, model awal dengan enam fitur dipertahankan sebagai model akhir penelitian. Secara ilmiah, hal ini dapat dijelaskan oleh sifat Random Forest sebagai model ensemble yang memanfaatkan kombinasi banyak fitur, termasuk fitur dengan kontribusi individual kecil, untuk membentuk batas keputusan yang lebih stabil antar-kelas melalui interaksi antarfitur pada setiap pohon keputusan. Persepsi beban tugas, meskipun memiliki importance individual terendah, tetap memberikan informasi pelengkap yang berinteraksi dengan fitur lain seperti jumlah tugas dan tingkat kelelahan, sehingga penghapusannya tidak memperbaiki akurasi namun justru meningkatkan variabilitas performa model antarfold.

Pembahasan

Mengacu pada panduan kategorisasi performa model klasifikasi yang umum digunakan dalam literatur pembelajaran mesin (akurasi $\geq 90\%$ = Sangat Baik, 80–89% = Baik, 70–79% = Cukup, 60–69% = Sedang, $<60\%$ = Kurang), model yang dibangun dalam penelitian ini, dengan akurasi 57,89%, berada pada kategori Kurang. Sebagai perbandingan tambahan, proporsi kelas mayoritas (Sedang) pada data uji adalah 11 dari 19 sampel atau 57,89%, yang berarti akurasi model pada data uji setara dengan baseline naif berupa penebakan kelas mayoritas secara konsisten. Temuan ini perlu diungkapkan secara transparan: pada ukuran sampel dan komposisi fitur saat ini, model belum menunjukkan keunggulan prediktif yang

jelas dibandingkan baseline sederhana, meskipun secara kualitatif model tetap mampu membedakan sebagian kategori Sedang dan Tinggi, sebagaimana tercermin dari confusion matrix pada Gambar 2.

Temuan ini sejalan dengan temuan pada literatur bahwa jam belajar dan kebiasaan harian seperti olahraga dan beban tugas memang berasosiasi dengan performa akademik (Ahmed, 2024; Murtaza & Bhaumik, 2025), namun kekuatan asosiasi tersebut bervariasi antarpopulasi dan tidak selalu cukup kuat untuk menghasilkan model klasifikasi dengan akurasi tinggi tanpa melibatkan variabel akademik langsung, seperti nilai semester sebelumnya, yang justru banyak digunakan sebagai prediktor dominan pada studi performa akademik lain (Begum & Padmannavar, 2023; Chen & Jin, 2024; Kumar et al., 2024). Perbedaan utama penelitian ini dengan studi-studi tersebut terletak pada keputusan metodologis untuk secara ketat memisahkan sumber label (nilai akademik) dari himpunan fitur prediktor, sehingga hasil yang diperoleh mencerminkan kemampuan prediksi murni dari kebiasaan harian, bukan dari korelasi langsung antarukuran akademik.

Mengingat jam tidur kadet dalam sampel penelitian ini sangat rendah (rata-rata 3,64 jam per hari) sementara jam belajar dan frekuensi olahraga justru menjadi dua dari tiga fitur paling berpengaruh terhadap performa akademik, temuan ini mengindikasikan adanya ruang perbaikan konkret pada tataran kebiasaan harian kadet. Dari sisi institusi, kampus dapat mempertimbangkan penataan ulang jadwal harian yang memberi ruang waktu istirahat yang lebih memadai di sela kegiatan akademik dan latihan kedisiplinan, serta menyediakan program pendampingan akademik (akademik coaching atau konseling belajar) bagi kadet yang teridentifikasi berisiko performa rendah melalui model deteksi dini semacam ini. Dari sisi kadet itu sendiri, perbaikan manajemen waktu pribadi termasuk pengaturan jam tidur yang lebih teratur dan pemeliharaan frekuensi olahraga yang konsisten di tengah padatnya jadwal Latihan berpotensi mendukung performa akademik, sejalan dengan pola yang teramati pada variabel prediktor penelitian ini. Rekomendasi ini bersifat naratif berdasarkan temuan feature importance dan statistik deskriptif yang telah disajikan, bukan hasil dari eksperimen intervensi baru, sehingga penerapannya pada tataran kebijakan tetap memerlukan kajian lanjutan oleh institusi.

Penelitian ini memiliki beberapa keterbatasan yang perlu diungkapkan secara transparan. Pertama, jumlah sampel yang relatif kecil (95 data setelah pra proses) dan distribusi kelas yang tidak seimbang menyebabkan model tidak mampu mengenali kategori dengan sampel

minoritas, sebagaimana ditunjukkan oleh recall 0% pada kategori Rendah. Kedua, akurasi model yang setara dengan baseline kelas mayoritas mengindikasikan bahwa keenam variabel kebiasaan harian yang tersedia belum cukup informatif untuk memprediksi performa akademik secara andal, sehingga variabel tambahan, seperti performa akademik semester sebelumnya atau indikator keterlibatan akademik, perlu dipertimbangkan pada penelitian lanjutan. Ketiga, data dikumpulkan melalui self-report kuesioner yang berpotensi mengandung bias subjektivitas responden, khususnya pada variabel ordinal seperti tingkat kelelahan dan persepsi beban tugas. Keempat, penetapan label performa akademik menggunakan metode $\text{mean} \pm \text{standar deviasi}$ mengasumsikan distribusi nilai mendekati normal; pada populasi dengan distribusi nilai yang sangat menceng, metode kategorisasi alternatif, seperti persentil, perlu dipertimbangkan dan dibandingkan.

KESIMPULAN DAN REKOMENDASI

Penelitian ini telah berhasil membangun model Random Forest untuk mengklasifikasikan performa akademik kadet mahasiswa UNHAN RI ke dalam tiga kategori yang dibentuk dari nilai akademik melalui metode $\text{mean} \pm \text{standar deviasi}$. Enam variabel kebiasaan harian jam belajar, jumlah tugas, jam tidur, frekuensi olahraga, tingkat kelelahan, dan persepsi beban tugas berperan sebagai prediktor dalam model tersebut. Akurasi yang dicapai model ini adalah 57,89% pada data uji dan 55,79% pada cross validation 5-fold keduanya tergolong kategori Kurang dan mendekati baseline kelas mayoritas. Eksperimen penghapusan fitur dengan importance terkecil, yaitu persepsi beban tugas, tidak memperbaiki akurasi dan justru meningkatkan variabilitas performa antarfold, sehingga keenam fitur awal dipertahankan sebagai model akhir.

Jam belajar (29,16%), frekuensi olahraga (19,39%), dan jumlah tugas (19,16%) teridentifikasi sebagai tiga faktor paling berpengaruh dalam model, yang mengindikasikan bahwa alokasi waktu belajar dan aktivitas fisik kadet memberikan kontribusi lebih besar terhadap pola klasifikasi dibandingkan variabel kelelahan subjektif atau persepsi beban tugas. Keterbatasan utama model terletak pada akurasi yang belum melampaui baseline kelas mayoritas dan ketidakmampuannya mengenali kategori Rendah, sehingga penggunaannya sebagai alat bantu deteksi dini saat ini masih memerlukan penyempurnaan lebih lanjut sebelum diterapkan secara operasional oleh institusi.

Penelitian selanjutnya disarankan untuk: (1) memperluas jumlah responden dan menerapkan teknik penyeimbangan kelas, seperti oversampling, untuk mengatasi ketidakmampuan model mengenali kategori Rendah; (2) menambahkan variabel akademik tidak langsung yang tetap independen dari label, seperti performa akademik semester sebelumnya atau tingkat kehadiran, untuk meningkatkan daya prediksi model; (3) membandingkan performa Random Forest dengan algoritma lain, seperti XGBoost, Support Vector Machine, atau jaringan saraf tiruan; dan (4) membandingkan metode kategorisasi label $\text{mean} \pm \text{standar deviasi}$ dengan metode alternatif, seperti kategorisasi berbasis persentil, untuk menguji sensitivitas hasil terhadap definisi label performa akademik.

REFERENSI

1. Ahmed, E. (2024). Student performance prediction using machine learning algorithms. *Applied Computational Intelligence and Soft Computing*, 2024, 4067721. <https://doi.org/10.1155/2024/4067721>
2. Azwar, S. (2021). *Penyusunan Skala Psikologi*. Edisi 3. Yogyakarta, Indonesia: Pustaka Pelajar.
3. Bakhshinategh, B., Zaïane, O. R., ElAtia, S., & Ipperciel, D. (2018). Educational data mining applications and tasks: A survey of the last 10 years. *Education and Information Technologies*, 23(1), 537–553. <https://doi.org/10.1007/s10639-017-9616-z>
4. Begum, S., & Padmannavar, S. S. (2023). Student performance analysis using Bayesian optimized random forest classifier and KNN. *International Journal of Engineering Trends and Technology*, 71(5), 132–140. <https://doi.org/10.14445/22315381/IJETT-V71I5P213>
5. Bellaj, M., Ben Dahmane, A., Boudra, S., & Lamarti Sefian, M. (2024). Educational Data Mining: Employing Machine Learning Techniques and Hyperparameter Optimization to Improve Students' Academic Performance. *International Journal of Online and Biomedical Engineering (IJOE)*, 20(3), 55–74. <https://doi.org/10.3991/ijoe.v20i03.46287>
6. Bulmer, S., Aisbett, B., Drain, J. R., Roberts, S., Gastin, P. B., Tait, J., & Main, L. C. (2022). Sleep of recruits throughout basic military training and its relationships with stress, recovery, and fatigue. *International Archives of Occupational and Environmental Health*, 95(6), 1331–1342. <https://doi.org/10.1007/s00420-022-01845-9>

7. Chen, Y., & Jin, K. (2024). Educational performance prediction with random forest and innovative optimizers: A data mining approach. *International Journal of Advanced Computer Science and Applications*, 15(3). <https://doi.org/10.14569/IJACSA.2024.0150308>
8. Coastera, F. F., Yusa, M., Lediwara, N., & Sari, J. P. (2020). Analisis Performa Algoritma Naïve Bayes Untuk Penentuan Kelayakan Pendonor Darah. *Jurnal VOI (Voice of Informatics)*, 9(1), 37–46. <https://voi.stmik-tasikmalaya.ac.id/index.php/voi/article/view/185>
9. Devi, K., & Ratnoo, S. (2022). Predicting student dropouts using random forest. *Journal of Statistics and Management Systems*, 25(7), 1579–1590. <https://doi.org/10.1080/09720510.2022.2130570>
10. Du, S., Hu, H., Cheng, K., & Li, H. (2023). Exercise makes better mind: a data mining study on effect of physical activity on academic achievement of college students. *Frontiers in Psychology*, 14, 1271431. <https://doi.org/10.3389/fpsyg.2023.1271431>
11. Gemilang Gultom, R. A., Lediwara, N., Kristijarso, Suhartotok, Hanifuddin, M., & Sandrika, H. P. (2023). IoT Drone Design for 3-Dimensional Mapping of the Republic of Indonesia Defense University (RIDU) Campus in the Sentul IPSC Area. *International Journal of Intelligent Systems and Applications in Engineering*, 11(1), 140–147. <https://ijisae.org/index.php/IJISAE/article/view/2452>
12. Han, J., Pei, J., & Tong, H. (2022). *Data Mining: Concepts and Techniques*. 4th ed. Cambridge, MA, USA: Morgan Kaufmann/Elsevier.
13. Handoko, M. B., Fauzi, Z. W., Agung, H., & Sofia, D. (2025). The Prediction of Science and Social Science Students Interests Using Decision Tree Algorithm with CRISP-DM. *Bit-Tech*, 8(1), 437–447. <https://doi.org/10.32877/bt.v8i1.2577>
14. Jawad, K., Shah, M. A., & Tahir, M. (2022). Students' Academic Performance and Engagement Prediction in a Virtual Learning Environment Using Random Forest with Data Balancing. *Sustainability*, 14(22), 14795. <https://doi.org/10.3390/su142214795>
15. Kumar, M., Singh, N., Wadhwa, J., Singh, P., Kumar, G., & Qtaishat, A. (2024). Utilizing random forest and XGBoost data mining algorithms for anticipating students' academic performance. *International Journal of Modern Education and Computer Science*, 16(2), 29–44. <https://doi.org/10.5815/ijmecs.2024.02.03>
16. Murtaza, G., & Bhaumik, A. (2025). Predicting academic outcomes in college students with machine learning: The role of study and sleep hours. *Journal of Media Horizons*,

-
- 6(3), 121–151. <https://jmhorizons.com/index.php/journal/article/view/280>.
<https://doi.org/10.5281/zenodo.15878011>
17. Sofyana, M., Wibowo, R. A., & Agustiningsih, D. (2022). Wake-up time and academic performance of university students in Indonesia: A cross-sectional study. *Frontiers in Education*, 7, 982320. <https://doi.org/10.3389/feduc.2022.982320>
18. Suardiaz-Muro, M., Ortega-Moreno, M., Morante-Ruiz, M., Monroy, M., A. Ruiz, M., Martin-Plasencia, P., Vela-Bueno, A. (2023). Sleep quality and sleep deprivation: relationship with academic performance in university students during examination period. *Sleep and Biological Rhythms*, 21(3), 377–383. <https://doi.org/10.1007/s41105-023-00457-1>
- 19.