

TESIS

**PREDIKSI KINERJA AKADEMIK MATEMATIKA SISWA
BERDASARKAN KEPERIBADIAN BIG FIVE MENGGUNAKAN
RANDOM FOREST DENGAN TEKNIK SYNTHETIC MINORITY
OVER-SAMPLING**



Disusun oleh:

Annisa Nurul Pratiwi

23.55.2465

Digital Transformation Intelligence

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA
2025**

TESIS

**PREDIKSI KINERJA AKADEMIK MATEMATIKA SISWA BERDASARKAN
KEPRIBADIAN BIG FIVE MENGGUNAKAN RANDOM FOREST DENGAN
TEKNIK SYNTHETIC MINORITY OVER-SAMPLING**

**PREDICTING STUDENTS' ACADEMIC PERFORMANCE IN
MATHEMATICS BASED ON BIG FIVE PERSONALITY TRAITS USING
RANDOM FOREST WITH SYNTHETIC MINORITY OVER-SAMPLING
TECHNIQUE**

Diajukan untuk memenuhi salah satu syarat mencapai derajat Pascasarjana

Program Studi PJJ Informatika



Disusun oleh:

Annlsa Nurul Pratiwi

23.55.2465

Digital Transformation Intelligence

**PROGRAM STUDI S2 TEKNIK INFORMATIKA
PROGRAM PASCASARJANA UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2025

HALAMAN PERSETUJUAN

**PREDIKSI KINERJA AKADEMIK MATEMATIKA SISWA BERDASARKAN
KEPRIBADIAN BIG FIVE MENGGUNAKAN RANDOM FOREST DENGAN
TEKNIK SYNTHETIC MINORITY OVER-SAMPLING**

**PREDICTING STUDENTS' ACADEMIC PERFORMANCE IN
MATHEMATICS BASED ON BIG FIVE PERSONALITY TRAITS USING
RANDOM FOREST WITH SYNTHETIC MINORITY OVER-SAMPLING
TECHNIQUE**

yang disusun dan diajukan oleh

Annisa Nurul Pratiwi

23.55.2465

telah disetujui oleh Dosen Pembimbing Tesis
pada tanggal 03 September 2025

Dosen Pembimbing



Prof. Dr. Ema Utami, S.Si., M.Kom.
NIK. 190302037

HALAMAN PENGESAHAN

**PREDIKSI KINERJA AKADEMIK MATEMATIKA SISWA BERDASARKAN
KEPRIBADIAN BIG FIVE MENGGUNAKAN RANDOM FOREST DENGAN
TEKNIK SYNTHETIC MINORITY OVER-SAMPLING**

**PREDICTING STUDENTS' ACADEMIC PERFORMANCE IN
MATHEMATICS BASED ON BIG FIVE PERSONALITY TRAITS USING
RANDOM FOREST WITH SYNTHETIC MINORITY OVER-SAMPLING
TECHNIQUE**

yang disusun dan diajukan oleh

Annisa Nurul Pratiwi

23.55.2465

Telah dipertahankan di depan Dewan Penguji
pada tanggal 03 September 2025

Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Dr. Kumara Ari Yuana, S.T., M.T
NIK. 190302575

Hanafi, S.Kom., M.Eng., Ph.D.
NIK. 190302024

Prof. Dr. Ema Utami, S.Si., M.Kom.
NIK. 190302037



Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer
Tanggal 03 September 2025

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : Annisa Nurul Pratiwi
NIM : 23.55.2465
Konsentrasi : Digital Transformation Intelligence

Menyatakan bahwa Tesis dengan judul berikut:
Prediksi Kinerja Akademik Matematika Siswa Berdasarkan Kepribadian Big Five Menggunakan Random Forest Dengan Teknik Synthetic Minority Over-Sampling

Dosen Pembimbing Utama : Prof. Dr. Ema Utami, S.Si., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 03 September 2025
Yang Menyatakan,



Annisa Nurul Pratiwi

HALAMAN PERSEMBAHAN

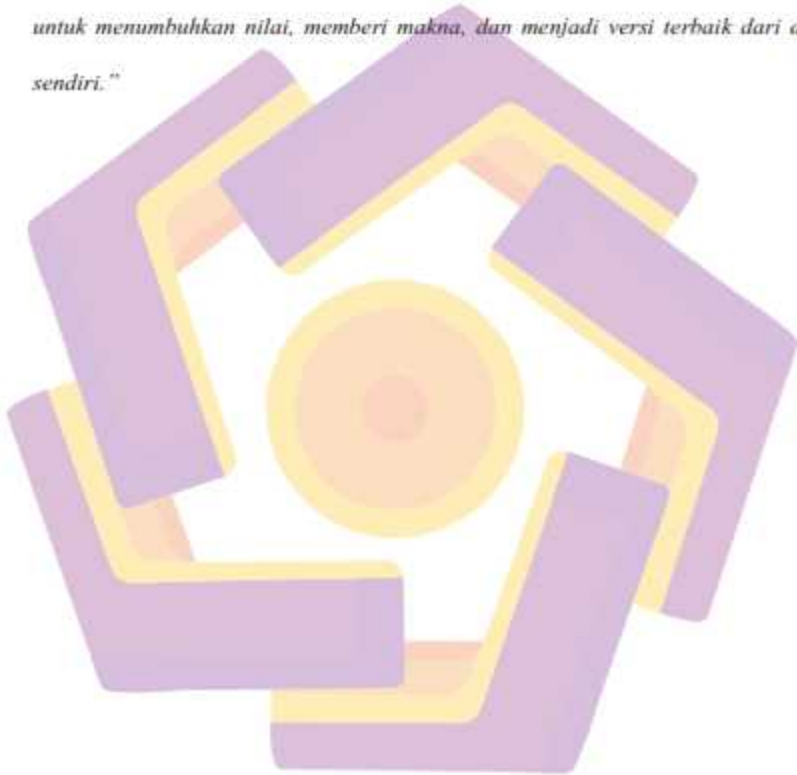
Dengan segala rasa syukur dan kerendahan hati, karya ini penulis persembahkan kepada :

1. Kedua orang tua tercinta atas doa, kasih sayang, dan dukungan yang tiada henti.
2. Kakak serta keluarga besar yang selalu memberikan semangat dan inspirasi.
3. Kepada Abah dan Kakek tercinta yang telah lebih dulu berpulang, namun kehadiran dan teladan mereka senantiasa menjadi motivasi untuk terus berbuat baik dan berjuang dalam kehidupan.
4. Pembimbing dan para pendidik yang telah menuntun dengan ilmu dan keteladanan.
5. Serta rekan-rekan PJJ Angkatan 10 Universitas AMIKOM yang menjadi bagian penting dalam perjalanan akademik ini.

Semoga karya ini menjadi wujud kecil dari rasa terima kasih dan dedikasi atas segala dukungan yang telah diberikan.

HALAMAN MOTTO

“Pernah merasa kehilangan arah dan jati diri, namun justru dari titik itu saya belajar arti bangkit. Sebab ketika hati terasa kosong, di sanalah ruang tercipta untuk menumbuhkan nilai, memberi makna, dan menjadi versi terbaik dari diri sendiri.”



KATA PENGANTAR

Puji syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat dan karunia-Nya, sehingga tesis ini dapat terselesaikan dengan baik. Segala kemudahan dan kekuatan yang diberikan selama proses penulisan merupakan anugerah yang tak ternilai.

Ucapan terima kasih penulis sampaikan kepada kedua orang tua tercinta, kakak, dan seluruh kerabat atas doa, dukungan, serta semangat yang tiada henti. Penghargaan dan rasa hormat yang mendalam juga penulis tujukan kepada Prof. Dr. Ema Utami, S.Si., M.Kom. selaku dosen pembimbing, atas bimbingan, ilmu, dan arahan yang sangat berharga. Tidak lupa terima kasih kepada rekan-rekan PJJ Angkatan 10 Universitas AMIKOM atas kebersamaan dan dukungan selama proses studi.

Akhirnya, penulis juga berterima kasih kepada diri sendiri atas ketekunan dan usaha yang telah dilakukan hingga karya ini terselesaikan. Semoga hasil ini dapat memberikan manfaat bagi pengembangan ilmu dan menjadi langkah menuju masa depan yang lebih baik.

Yogyakarta, 13 Oktober 2025

Penulis

DAFTAR ISI

HALAMAN JUDUL.....	ii
HALAMAN PERSETUJUAN.....	iii
HALAMAN PENGESAHAN.....	iv
HALAMAN PERNYATAAN KEASLIAN TESIS.....	Error! Bookmark not defined.
HALAMAN MOTTO.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL.....	xii
DAFTAR GAMBAR.....	xiv
INTISARI.....	xvi
<i>ABSTRACT</i>	xvii
BAB I PENDAHULUAN.....	1
1.1. Latar Belakang Masalah.....	1
1.2. Rumusan Masalah.....	12
1.3. Batasan Masalah.....	13
1.4. Tujuan Penelitian.....	15
1.5. Manfaat Penelitian.....	15
BAB II TINJAUAN PUSTAKA.....	17
2.1. Tinjauan Pustaka.....	17
2.2. Keaslian Penelitian.....	25
2.3. Landasan Teori.....	32

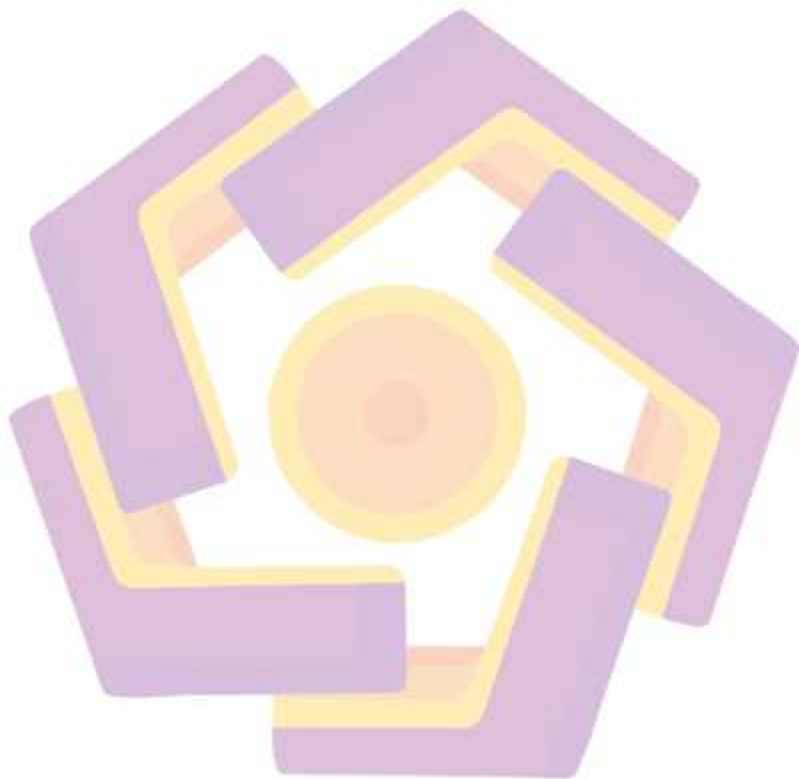
2.3.1 Educational Data Mining.....	32
2.3.2 Kinerja Akademik	33
2.3.3 Kepribadian Big Five	33
2.3.4 Klasifikasi	36
2.3.5 Random Forest.....	37
2.3.6 Hyperparameter	41
2.3.7 Grid Search.....	42
2.3.8 Teknik SMOTE	42
2.3.9 Evaluasi Metode	44
2.3.10 ROC-AUC.....	45
2.3.11 Cross Validation.....	47
BAB III METODE PENELITIAN.....	49
3.1. Jenis, Sifat, dan Pendekatan Penelitian.....	49
3.2. Metode Pengumpulan Data.....	49
3.3. Metode Analisis Data.....	56
3.4. Alur Penelitian	58
a. Topik Penelitian.....	59
b. Pengumpulan Dataset	59
c. Preprocessing Data	60
d. Pemodelan Random Forest.....	61
e. Hasil dan Pembahasan.....	62

BAB IV HASIL PENELITIAN DAN PEMBAHASAN	64
4.1. Pembuatan Dataset	64
4.1.1 Pengumpulan Data	64
4.1.2 Uji Validitas Isi Kuesioner	66
4.1.3 Uji Validitas Konstruk Kuesioner	70
4.1.4 Uji Reliabilitas	76
4.1.5 Preprocessing Data	77
4.2. Analisis Hubungan Faktor	89
4.3. Pemodelan dan Evaluasi	100
4.3.1 Skenario 1	100
4.3.2 Skenario 2	106
4.3.3 Skenario 3	113
4.3.4 Skenario 4	119
4.4. Analisis Perbandingan	126
BAB V PENUTUP	141
5.1. Kesimpulan	141
5.2. Saran	142
DAFTAR PUSTAKA	144
LAMPIRAN	149

DAFTAR TABEL

Tabel 2. 1. Matriks Literature Review dan Posisi Penelitian	25
Tabel 3. 1. Informasi Data Demografi dan Akademik Siswa	50
Tabel 3. 2. Kuesioner Kepribadian Model Big Five berdasarkan Big Five Inventory (BFI)	53
Tabel 4. 1. Informasi Atribut dan Value Data Akademik, Demografis, dan Kepribadian Big Five.....	65
Tabel 4. 2. Perbaikan Instrumen Kepribadian Big Five.....	68
Tabel 4. 3. Statistik Deskriptif Data Kuesioner Kepribadian Big Five.....	71
Tabel 4. 4. Frekuensi dan Presentase Jawaban Kuesioner Kepribadian Big Five (Skala Likert 1-5).....	72
Tabel 4. 5. Uji Validasi Konstruk Domain Kepribadian Big Five (Extraversion, Agreeableness, Conscientiousness, Neuroticism, dan Openess).....	74
Tabel 4. 6. Uji Realibilitas Cronbach's Alpha	76
Tabel 4. 7. Klasifikasi Kelas Nilai Cronbach's Alpha	77
Tabel 4. 8. Informasi Transformasi Data	80
Tabel 4. 9. Hasil Pengujian dengan K-Fold Cross Validation Skenario 1	101
Tabel 4. 10. Bobot Kepentingan Fitur (Feature Importance) Skenario 1.....	106
Tabel 4. 11. Hasil Pengujian dengan K-Fold Cross Validation Skenario 2	108
Tabel 4. 12. Bobot Kepentingan Fitur (Feature Importance) Skenario 2.....	113
Tabel 4. 13. Hasil Pengujian dengan K-Fold Cross Validation Skenario 3	114
Tabel 4. 14. Bobot Kepentingan Fitur (Feature Importance) Skenario 3.....	119

Tabel 4. 15. Hasil Pengujian dengan K-Fold Cross Validation Skenario 4.....	121
Tabel 4. 16. Bobot Kepentingan Fitur (Feature Importance) Skenario 4.....	126
Tabel 4. 17. Evaluasi Model Prediksi Kinerja Akademik.....	129



DAFTAR GAMBAR

Gambar 2.1. Prosedur Decision Tree	37
Gambar 2. 2. Majority Voting (Raschka & Mirjalili, 2017)	40
Gambar 2. 3. Metriks Evaluasi Confusion Matrix	44
Gambar 2. 4. ROC Curve (Géron, 2019)	47
Gambar 3. 1. Alur Penelitian.....	58
Gambar 4. 1. Grafik Penyebaran Kelas Target	78
Gambar 4. 2. Distribusi Data Akademik Berdasarkan Nilai Kinerja Akademik ..	83
Gambar 4. 3. Distribusi Data Demografi Berdasarkan Nilai Kinerja Akademik..	86
Gambar 4. 4. Distribusi Data Kepribadian Berdasarkan Nilai Kinerja Akademik	88
Gambar 4. 5. Matriks Korelasi Fitur Akademik dengan Fitur Kinerja Akademik	90
Gambar 4. 6. Matriks Korelasi Fitur Demografi dengan Fitur Kinerja Akademik	92
Gambar 4. 7. Matriks Korelasi Fitur Kepribadian dengan Fitur Kinerja Akademik	94
Gambar 4. 8 Nilai Korelasi Semua Fitur dengan Fitur Kinerja Akademik.....	96
Gambar 4. 9. Grafik Learning Curve k-fold = 7 Skenario 1	103
Gambar 4. 10. Confusion Matrix Skenario 1	104
Gambar 4. 11 ROC-AUC Skenario 1.....	105
Gambar 4. 12. Distribusi Kelas Sebelum SMOTE Skenario 2	107
Gambar 4. 13. Distribusi Kelas Setelah SMOTE Skenario 2	107
Gambar 4. 14. Grafik Learning Curve k-fold = 10 Skenario 2	110
Gambar 4. 15. Confusion Matrix Skenario 2	111

Gambar 4. 16 ROC-AUC Skenario 2.....	112
Gambar 4. 17. Grafik Learning Curve k-fold = 7 Skenario 3.....	115
Gambar 4. 18. Confusion Matrix Skenario 3.....	117
Gambar 4. 19. ROC-AUC Skenario 3.....	118
Gambar 4. 20. Distribusi Kelas Sebelum SMOTE Skenario 4.....	120
Gambar 4. 21. Distribusi Kelas Sesudah SMOTE Skenario 4.....	120
Gambar 4. 22. Grafik Learning Curve k-fold = 6 Skenario 4.....	123
Gambar 4. 23. Confusion Matrix Skenario 4.....	124
Gambar 4. 24. ROC-AUC Skenario 4.....	125
Gambar Lampiran 1. Form Penilaian Angket Psikolog 1.....	153
Gambar Lampiran 2. Form Catatan Psikolog 1.....	154
Gambar Lampiran 3. Form Penilaian Angket Psikolog 2.....	155
Gambar Lampiran 4. Form Catatan Psikolog 2.....	156

INTISARI

Masa sekolah menengah merupakan periode penting untuk perkembangan kinerja akademik dan sosial siswa. Educational data mining (EDM) menjadi salah satu metode strategis yang mampu mengeksplorasi pola dalam data pendidikan untuk memprediksi kinerja akademik berdasarkan berbagai faktor, termasuk kepribadian siswa. Namun, ketidakseimbangan data pendidikan masih menjadi masalah yang dapat menyebabkan bias pada model prediksi. Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor yang berkontribusi terhadap kinerja akademik matematika siswa sekolah menengah pertama, seperti faktor akademik, demografis, dan kepribadian model Big Five. Metode Random Forest dan teknik oversampling SMOTE digunakan untuk mengidentifikasi komponen yang berkontribusi terhadap kinerja akademik siswa, serta meningkatkan performa model prediksi. Penelitian ini menunjukkan bahwa faktor akademik menjadi faktor penting, sementara faktor sosial-ekonomi dan kepribadian kurang signifikan terhadap kinerja akademik. Selain itu, penerapan teknik SMOTE terbukti efektif dalam mengatasi ketidakseimbangan data, serta model Random Forest memiliki performa optimal dengan tuning yang tepat. Kombinasi antara Random Forest, hyperparameter tuning GridSearchCV dan SMOTE berhasil mengembangkan model dengan tingkat akurasi mencapai 99%.

Kata kunci: Big Five, Educational Data Mining, Kinerja Siswa, Random Forest, SMOTE

ABSTRACT

The secondary school period is a crucial time for the development of students' academic and social performance. Educational data mining (EDM) has emerged as a strategic method capable of exploring patterns in educational data to predict academic performance based on various factors, including students' personalities. However, the imbalance in educational data remains an issue that can lead to bias in predictive models. This study aims to identify the factors contributing to the academic performance in mathematics of junior high school students, such as academic, demographic, and Big Five personality factors. The Random Forest method and SMOTE oversampling technique are employed to identify components that contribute to students' academic performance and to enhance the performance of the predictive model. The research indicates that academic factors are significant, while socio-economic and personality factors are less significant in relation to academic performance. Additionally, the application of the SMOTE technique proves effective in addressing data imbalance, and the Random Forest model demonstrates optimal performance with appropriate tuning. The combination of Random Forest, hyperparameter tuning using GridSearchCV, and SMOTE successfully develops a model with an accuracy rate of 99%.

Keyword: Big Five, Educational Data Mining, Student Performance, Random Forest, SMOTE

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Educational Data Mining (EDM) merupakan suatu disiplin ilmu yang memfokuskan diri dalam pengembangan sistem pendidikan, termasuk sekolah menengah, universitas, dan sistem pembelajaran cerdas (Nachouki et al., 2023). EDM memanfaatkan teknik berbasis computer, seperti *data mining*, *machine learning*, dan teknik statistik (Nachouki et al., 2023), untuk mengidentifikasi pola dalam sebuah data pendidikan yang cukup kompleks, sehingga sulit untuk dianalisis secara manual (Ozyurt et al., 2023). Penggunaan EDM dapat membantu dalam memodelkan proses pembelajaran, meningkatkan kualitas pengambilan keputusan, serta membantu sekolah dalam mengidentifikasi informasi untuk pemahaman siswa ke arah yang lebih baik, yang pada gilirannya dapat meningkatkan hasil belajar dan kinerja akademik mereka (Ozyurt et al., 2023). Prediksi kinerja akademik menjadi salah satu isu penting yang banyak diteliti dalam bidang EDM, karena memiliki peran dalam meningkatkan proses pembelajaran dan pemahaman siswa (Ozyurt et al., 2023). Peningkatan minat penelitian pada topik ini dipicu oleh pertumbuhan data pembelajaran, yang memungkinkan lembaga pendidikan untuk mengidentifikasi siswa yang mengalami kesulitan, memperbaiki metode pengajaran, serta meningkatkan tingkat keberhasilan akademik siswa (Dol & Jawandhiya, 2024; Issah et al., 2023; Roslan & Chen, 2022). Pada dasarnya dengan memprediksi kinerja siswa, institusi pendidikan dapat mengatasi kelemahan dalam

sistem pendidikan dengan mengembangkan materi yang lebih responsif untuk meningkatkan prestasi siswa, serta memotivasi dan membangun kepercayaan diri siswa (Dol & Jawandhiya, 2022; Issah et al., 2023; Rodrigues et al., 2022). Oleh karena itu, penelitian prediksi kinerja akademik dalam EDM sangat penting untuk dilakukan, mengingat bahwa pendekatan strategis ini penting bagi lembaga pendidikan dalam mengidentifikasi siswa yang memiliki potensi rendah, mencegah terjadinya kekurangan dalam sistem pendidikan, serta membantu dalam mengembangkan materi yang responsif dan memberikan umpan balik yang signifikan bagi kemajuan belajar siswa (Issah et al., 2023; Rodrigues et al., 2022).

Meskipun penelitian serupa terkait prediksi kinerja akademik di tingkat perguruan tinggi atau kursus online telah berkembang pesat, studi di tingkat sekolah menengah masih berada pada tahap awal (Rodrigues et al., 2022), terutama dalam bidang matematika yang menunjukkan prestasi belajar matematika di Indonesia masih berada pada tingkat yang rendah (Wawan & Retnawati, 2022). Matematika menjadi tantangan bagi para siswa, karena para siswa sulit dalam memahami dan menerapkan konsep-konsep matematika (Wawan & Retnawati, 2022). Oleh karena itu, mengidentifikasi faktor-faktor yang mempengaruhi kinerja akademik siswa menjadi langkah penting, terutama selama masa sekolah menengah khususnya yang berusia antara 10 hingga 18 tahun, yang merupakan periode kritis dalam perkembangan akademik dan sosial (Costa et al., 2024). Masa sekolah menengah merupakan fase penting, di mana perubahan fisik, emosional, dan sosial yang dialami dapat berdampak pada pencapaian akademik, yang pada gilirannya juga berpengaruh langsung terhadap keputusan pendidikan, pilihan karier di masa depan,

serta aspek-aspek kehidupan jangka panjang mereka (Costa et al., 2024). Dengan demikian, penelitian di bidang ini dianggap penting karena dapat memperbaiki proses pembelajaran yang pada gilirannya dapat meningkatkan kualitas pendidikan. Berdasarkan hal tersebut, penelitian ini bertujuan untuk memprediksi kinerja akademik siswa Sekolah Menengah Pertama (SMP) dengan fokus pada mata pelajaran matematika, guna mengidentifikasi faktor-faktor yang mempengaruhi prestasi siswa.

Banyak penelitian telah mengidentifikasi sejumlah faktor yang berfungsi sebagai indikator utama dalam kinerja akademik. Secara umum faktor akademik siswa telah menunjukkan tingkat keakuratan yang signifikan, ini menunjukkan bahwa aspek akademik dianggap sebagai indikator utama dalam memprediksi kinerja siswa (Nachouki et al., 2023). Namun, penggabungan faktor akademik dengan variabel lain telah membuka potensi baru dalam menghasilkan prediksi yang lebih akurat. Ini didasarkan pada pernyataan (Costa et al., 2024; Shaninah & Mohd Noor, 2024) bahwa prestasi akademik siswa merupakan fenomena multidimensional yang dapat dipengaruhi oleh berbagai faktor, baik yang bersifat personal (kognitif, kepribadian) maupun kontekstual (keluarga, lingkungan). Faktor kepribadian, terutama yang diukur melalui dimensi Big Five, telah terbukti menjadi salah satu faktor utama yang mempengaruhi prestasi akademik pada anak-anak dan remaja (Shaninah & Mohd Noor, 2024). Sementara itu faktor akademik, seperti nilai yang diperoleh sebelumnya, berfungsi sebagai indikator kemampuan kognitif dan merupakan prediktor yang kuat untuk keberhasilan akademik di masa depan (Costa et al., 2024). Selain itu, faktor demografis seperti jenis kelamin dan latar

belakang orang tua juga merupakan variabel penting dalam memahami variasi pola prestasi di antara kelompok siswa (Costa et al., 2024).

Secara lebih rinci beberapa penelitian telah menemukan bahwa penggabungan variabel akademik, demografi, dan kepribadian mampu memberikan hasil prediksi yang lebih komprehensif. Penelitian yang dilakukan oleh (El-Keiey et al., 2022) mengusulkan pendekatan menggunakan faktor akademik, kepribadian *big five*, demografi dan skor *Intelligence Quotient* (IQ) untuk memprediksi kinerja akademik mahasiswa sarjana. Hasil penelitian tersebut menunjukkan bahwa faktor kepribadian *big five* ditemukan sebagai indikator paling signifikan dalam memprediksi kinerja akademik mahasiswa, di mana penghapusan faktor kepribadian menurunkan performa model. Kepribadian Big Five ini mengelompokkan kepribadian ke dalam lima faktor atau domain utama, yaitu *Neuroticism* (N), *Extraversion* (E), *Agreeableness* (A), *Conscientiousness* (C), dan *Openness* (O), di mana setiap variabel ini diyakini memiliki pengaruh terhadap kinerja akademik (Meyer et al., 2023; Rico-Juan et al., 2024; Shaninah & Mohd Noor, 2024). Pada dasarnya model *Big Five* dapat memberikan pemahaman yang lebih rinci tentang hubungan antara kepribadian dan prestasi siswa, terutama ketika mempertimbangkan spesifikasi *domain* dan ukuran prestasi siswa (Meyer et al., 2023; Shaninah & Mohd Noor, 2024). Selain faktor kepribadian, penting untuk mempertimbangkan faktor demografi seperti jenis kelamin, latar belakang keluarga, dan status sosial-ekonomi ketika memprediksi kinerja akademik (Costa et al., 2024; Roslan & Chen, 2023; Shaninah & Mohd Noor, 2024). Hal ini juga ditekankan oleh penelitian (Costa et al., 2024) yang mengungkapkan bahwa prestasi

akademik siswa sekolah menengah, dipengaruhi oleh berbagai faktor pribadi, meliputi karakteristik kepribadian, peran orang tua, status sosial-ekonomi, serta lingkungan di sekitar mereka. Hal ini menegaskan bahwa prestasi akademik tidak hanya dipengaruhi oleh kemampuan kognitif, tetapi juga oleh berbagai faktor non-kognitif yang lebih stabil dan dapat mendukung upaya belajar siswa dalam jangka panjang (Costa et al., 2024). Dengan demikian, penggabungan tiga kelompok variabel yang berbeda, seperti akademik, demografis, dan kepribadian dalam satu penelitian memberikan kerangka yang lebih komprehensif. Pendekatan ini memungkinkan peneliti untuk melihat bagaimana kemampuan akademik dasar, karakteristik individu (demografi), dan aspek kepribadian saling berinteraksi dalam memprediksi kinerja akademik siswa dengan lebih akurat. Berdasarkan hal tersebut, penelitian ini bertujuan untuk menerapkan faktor akademik, demografi, dan kepribadian model *Big Five* dalam memprediksi kinerja akademik siswa.

Para peneliti telah mengeksplorasi berbagai faktor yang dianggap mempengaruhi kinerja akademik dengan menggunakan berbagai teknik, salah satunya teknik *data mining* (Roslan & Chen, 2022). Metode klasifikasi menjadi teknik *data mining* yang paling umum digunakan dalam penelitian EDM, dengan metode yang umum digunakan dalam teknik ini adalah Decision Tree, Naïve Bayes, Random Forest (Roslan & Chen, 2022). Algoritma Decision Tree, Naïve Bayes dan Random Forest ditemukan efektif dalam memprediksi kinerja siswa dengan tingkat akurasi yang cukup baik (El-Keiey et al., 2022; Khairy et al., 2024; Nachouki et al., 2023; Roslan & Chen, 2023; Santoso et al., 2024). Naïve Bayes memiliki beberapa keunggulan, termasuk efisiensi dalam komputasi, kemampuannya untuk

menangani data besar dengan dimensi tinggi, serta toleran terhadap data yang hilang dan fitur yang tidak relevan (Pajila et al., 2023). Namun, Naïve Bayes juga memiliki kelemahan, yaitu mengasumsikan bahwa fitur-fitur yang ada sifatnya independen atau menganggap perubahan pada satu fitur tidak akan mempengaruhi fitur lainnya yang mungkin tidak selalu benar dalam situasi nyata, serta metode ini juga sensitif terhadap interaksi antar fitur, sehingga hal ini dapat mengurangi tingkat akurasi prediksi ketika fitur-fitur tersebut sebenarnya saling terikat (Pajila et al., 2023; Santoso et al., 2024). Berbeda dengan Naïve Bayes, Decision Tree tidak memerlukan asumsi bahwa fitur-fitur bersifat independen. Decision tree secara aktif memilih fitur yang paling informatif untuk memisahkan data di setiap node (Mienye & Jere, 2024). Hal ini memungkinkan Decision Tree untuk mendeteksi pola-pola kompleks yang mungkin ada di antara fitur-fitur. Ketika terdapat interaksi antar fitur, Decision Tree dapat mengidentifikasinya melalui pemilihan fitur yang berulang selama proses pembentukan pohon (Mienye & Jere, 2024). Oleh karena itu, Decision Tree menawarkan fleksibilitas dan efektivitas yang lebih baik dalam menangani dataset dengan korelasi antar fitur. Namun, metode ini juga memiliki kelemahan, yaitu rentan terhadap *overfitting*, terutama ketika berhadapan dengan dataset yang memiliki banyak fitur atau ketika pohon yang dibentuk terlalu dalam, karena metode ini sensitif terhadap *outlier* dan noise, yang dapat mengurangi akurasi prediksi dan tidak mempresentasikan pola data yang sebenarnya (Mienye & Jere, 2024). Oleh karena itu, Random Forest dapat dianggap sebagai solusi yang efektif dalam mengatasi atau setidaknya mengurangi risiko *overfitting* (Mienye & Jere, 2024; Santoso et al., 2024). Random Forest mengatasi masalah *overfitting*

melalui dua pendekatan utama, yaitu *bootstrap sampling* yang menghasilkan beberapa set data pelatihan yang berbeda untuk menurunkan varians, serta teknik *feature randomization* yang memilih fitur secara acak untuk setiap pohon (Mienye & Jere, 2024). Pada dasarnya, untuk mengatasi keterbatasan yang ada pada satu pohon keputusan tunggal, Random Forest membangun beberapa pohon keputusan untuk mengurangi korelasi antar pohon dan meningkatkan keragaman, selain itu metode ini menerapkan mekanisme *majority voting* antara pohon keputusan, sehingga dapat menghasilkan prediksi yang lebih tepat serta meningkatkan akurasi model (Santoso et al., 2024). Sehingga, mengingat bahwa prediksi kinerja akademik melibatkan sejumlah fitur yang mungkin saling berhubungan, metode Naïve Bayes dianggap kurang tepat karena tidak efektif dalam menangani interaksi antar fitur. Selain itu, dengan banyaknya fitur yang terdapat dalam dataset, Decision Tree juga kurang ideal karena cenderung mengalami *overfitting*. Oleh karena itu, penggunaan Random Forest dianggap sebagai pendekatan yang paling tepat untuk mengembangkan model prediksi kinerja akademik. Random Forest dianggap menjadi pilihan tepat dalam membangun model prediksi, karena memiliki jumlah *hyperparameter* yang lebih banyak dibandingkan dengan metode lainnya, yang berkontribusi secara positif terhadap kinerjanya (Pecuchova & Drlik, 2023). Banyaknya *hyperparameter* dalam Random Forest, ini menunjukkan betapa pentingnya pengaturan yang tepat untuk pengembangan model, karena setiap parameter memiliki dampak signifikan terhadap kinerja prediksi (Pecuchova & Drlik, 2023). Oleh karena itu, guna mendapatkan parameter yang optimal, Random Forest dapat menerapkan proses *hyperparameter tuning* seperti *grid search*, untuk

menemukan kombinasi parameter yang optimal (Arden & Safitri, 2022; Pecuchova & Drlik, 2023; Rimal et al., 2024). Proses *tuning* ini dilakukan untuk mencapai keseimbangan antara *overfitting* dan *underfitting*, yang pada gilirannya dapat meningkatkan akurasi serta kemampuan generalisasi model (Rimal et al., 2024). Tanpa *tuning* yang memadai, Random Forest berisiko tidak mencapai kinerja terbaiknya, karena *hyperparameter* seperti jumlah pohon dan kedalaman pohon berpengaruh secara signifikan terhadap hasil prediksi (Arden & Safitri, 2022).

Meskipun penelitian dalam bidang EDM ini sudah menunjukkan kemajuan yang cukup signifikan, salah satu penelitian telah mengidentifikasi bahwa ketidakseimbangan data masih menjadi tantangan yang sering muncul dalam konteks data pendidikan (Dol & Jawandhiya, 2022). Meski Random Forest memiliki keakuratan yang cukup baik dalam membangun model prediksi, model ini juga tetap memiliki kekurangan ketika dihadapkan pada data tidak seimbang, yang dapat mengakibatkan model mengalami bias dan menghasilkan kesalahan atau kinerja yang buruk (Khairy et al., 2024; Nachouki et al., 2023; Santoso et al., 2024). Hal ini juga sejalan dengan temuan (Dol & Jawandhiya, 2022) yang mengidentifikasi bahwa ketidakseimbangan data harus diatasi karena dapat mengakibatkan bias terhadap kinerja model. Pada dasarnya ketika jumlah data dalam satu kelas jauh lebih banyak dibandingkan dengan kelas lainnya, model cenderung memberikan perhatian lebih pada kelas mayoritas dan mengabaikan kelas minoritas yang seringkali memiliki informasi yang kurang lengkap, sehingga hal tersebut menyebabkan performa prediksi menjadi buruk terutama dalam mengidentifikasi siswa yang berisiko gagal (Abdul Bujang et al., 2023). Penelitian

yang dilakukan oleh (Dol & Jawandhiya, 2022) juga menyoroti pentingnya penggunaan dataset yang komprehensif, serta perlunya penanganan ketidakseimbangan data, seperti menerapkan salah satu teknik sampling untuk memastikan keakuratan dalam analisis. Metode *oversampling* seperti *Synthetic Minority Over-Sampling Technique* (SMOTE) telah terbukti efektif dalam menangani ketidakseimbangan data dalam konteks prediksi nilai siswa (Abdul Bujang et al., 2023). Salah satu penelitian telah menunjukkan bahwa teknik SMOTE berhasil membangun model analisis kepribadian pelanggan dengan tingkat akurasi, presisi, dan *recall* tertinggi, terutama ketika digunakan bersama dengan model random forest (Ghaniaviyanto Ramadhan & Adiwijaya, 2022). Penggunaan teknik SMOTE secara garis besar dapat menyeimbangkan dataset yang tidak seimbang, sehingga berpotensi meningkatkan akurasi model prediksi tanpa mengorbankan informasi penting yang terkandung dalam data tersebut (Pecuchova & Drlik, 2023). Namun, penerapan SMOTE juga dapat mengakibatkan *overfitting* karena data sintesis yang dihasilkan sering kali terlalu mirip dengan data asli, di mana SMOTE menciptakan titik sintesis diantara data asli dan tetangganya, sehingga tidak menambah informasi baru dan dapat menurunkan kemampuan generalisasi model (Abdul Bujang et al., 2023). Penurunan performa juga dapat terjadi karena proses *oversampling* memaksa jumlah data kelas minoritas untuk seimbang dengan mayoritas tanpa mempertimbangkan distribusi alami, sehingga berpotensi menimbulkan noise dan bias (Ghaniaviyanto Ramadhan & Adiwijaya, 2022). Akibatnya, model dapat menjadi kurang sensitif terhadap karakteristik sebenarnya dari kelas minoritas, yang memunculkan bias dan meningkatkan potensi

kesalahan prediksi. Dengan kata lain, model cenderung terlalu umum (*over-generalization*), karena data sintesis yang dihasilkan belum tentu mencerminkan distribusi asli dari kelas minoritas. Meskipun begitu, secara keseluruhan, teknik SMOTE mampu memberikan peningkatan yang signifikan terhadap akurasi algoritma *supervised*. Penelitian lain juga menyimpulkan bahwa algoritma Random Forest menunjukkan performa terbaik dalam memprediksi kinerja mahasiswa di perguruan tinggi setelah penerapan teknik SMOTE, berdasarkan metrik *precision*, *recall*, dan *F-Measure* (Hamoud et al., 2022). Oleh karena itu, dalam penelitian ini, teknik SMOTE akan diterapkan untuk mengatasi ketidakseimbangan data dengan meningkatkan jumlah sampel dalam kelas minoritas.

Berdasarkan uraian latar belakang yang telah disampaikan, dapat disimpulkan bahwa penelitian dalam bidang EDM yang berkaitan dengan prediksi kinerja akademik siswa, khususnya tingkat sekolah menengah, memiliki peranan yang penting dalam peningkatan mutu pendidikan. Oleh karena itu, penelitian ini bertujuan untuk mengidentifikasi berbagai faktor yang mempengaruhi prestasi akademik siswa di Sekolah Menengah Pertama (SMP), dengan penekanan pada mata pelajaran matematika. Penelitian ini akan mempertimbangkan faktor-faktor akademik, demografi, serta karakteristik kepribadian berdasarkan model Big Five dalam memprediksi kinerja akademik siswa, yang diyakini memiliki pengaruh signifikan terhadap kinerja siswa, sehingga pada gilirannya dapat meningkatkan efektivitas pengembangan model prediksi kinerja akademik siswa. Mengingat kinerja akademik siswa dipengaruhi oleh berbagai faktor pribadi termasuk karakteristik kepribadian, peran orang tua, status sosial-ekonomi, serta lingkungan

sekitar mereka, maka pengumpulan serta penggunaan data yang relevan mengenai kepribadian dan kinerja siswa menjadi sangat penting. Sehingga, dapat disimpulkan bahwa dataset yang diperlukan dalam penelitian ini mencakup informasi sensitif dan privasi, seperti karakteristik atau perilaku siswa yang hanya dapat diketahui oleh individu tersebut, data demografis, dan informasi akademik. Dengan demikian, dataset *private* dianggap lebih tepat karena sifatnya yang lebih personal dan kontekstual, sedangkan dataset publik cenderung bersifat umum. Oleh karena itu, penelitian ini bertujuan untuk mengumpulkan data yang diperlukan melalui lembaga pendidikan SMP Negeri 2 Kuningan dengan mengakses data pendidikan (DAPODIK) melalui operator, serta menyebarkan kuesioner untuk mengukur kepribadian Big Five siswa. Penyebaran kuesioner dipilih karena memberikan kontrol yang lebih besar terhadap data, sehingga memungkinkan peneliti untuk mengatur format dan standar pengisian sesuai dengan kebutuhan, sehingga menghasilkan data yang lebih relevan dan bersih (Jatnika et al., 2024).

Selain itu, metode klasifikasi Random Forest, yang dianggap sebagai pendekatan paling efektif untuk membangun model prediksi akan digunakan untuk mengelompokkan siswa ke dalam tiga kategori, yaitu cukup baik, baik, dan sangat baik berdasarkan kinerja akademiknya. Mengingat banyaknya *hyperparameter* yang perlu disesuaikan dalam Random Forest, penentuan pengaturan yang tepat menjadi tantangan. Oleh karena itu, proses *hyperparameter tuning* akan dilakukan menggunakan *grid search* untuk mengidentifikasi kombinasi parameter yang paling optimal. Masalah ketidakseimbangan data yang mungkin muncul akan diatasi dengan menerapkan teknik *oversampling* seperti SMOTE, sehingga

ketidakseimbangan data dapat teratasi, yang pada gilirannya dapat meningkatkan performa model atau tingkat akurasi. Dengan demikian, pendekatan ini diharapkan dapat memberikan kontribusi penting dalam pemahaman terkait faktor-faktor yang mempengaruhi kinerja akademik, serta menghasilkan model yang komprehensif dan dapat diandalkan dalam memprediksi kinerja siswa, sehingga dapat membantu institusi pendidikan dalam mengidentifikasi siswa yang berpotensi mengalami kegagalan, serta membantu meningkatkan strategi pembelajaran yang lebih efektif.

1.2. Rumusan Masalah

Dengan mempertimbangkan konteks masalah yang telah diuraikan pada latar belakang masalah, maka perumusan masalah dalam penelitian ini dapat dirinci sebagai berikut :

- a. Bagaimana hubungan antara variabel-variabel seperti faktor akademik, faktor demografi, dan faktor kepribadian Big Five terhadap nilai akhir siswa dalam mata pelajaran Matematika?
- b. Bagaimana pengaruh penerapan teknik SMOTE terhadap performa model Random Forest dalam memprediksi nilai akhir siswa pada dataset yang tidak seimbang?
- c. Bagaimana kinerja model Random Forest secara keseluruhan dalam memprediksi nilai akhir siswa berdasarkan metrik evaluasi yang sudah ditentukan?

1.3. Batasan Masalah

Dalam rangka menjaga fokus penelitian, sangatlah penting untuk menetapkan batasan-batasan yang tegas guna mengarahkan fokus pada penelitian ini. Berikut ini merupakan batasan masalah yang telah ditetapkan untuk penelitian ini :

- a. Data siswa diperoleh melalui data pendidikan (DAPODIK) SMPN 2 Kuningan
- b. Fokus pada hasil klasifikasi kinerja akademik siswa dalam mata pelajaran matematika
- c. Data yang digunakan terbatas pada kinerja mata pelajaran matematika siswa kelas 8 dan 9 pada saat kelas 7 di semester 2
- d. Data siswa terdiri dari 377 siswa angkatan 2022 (kelas 9) dan 416 siswa angkatan 2023 (kelas 8)
- e. Data siswa yang diperoleh diklasifikasikan dalam tiga kategori, yaitu cukup (75-82), baik (83-90), dan sangat baik (91-100), berdasarkan kinerja akademik mata pelajaran matematika, dengan proses pelabelan melibatkan diskusi bersama guru matematika kelas 7
- f. Data dependent pada penelitian ini ialah kinerja nilai akhir matematika, dengan kriteria cukup (75-82), baik (83-90), dan sangat baik (91-100)
- g. Data independent pada penelitian ini terdiri dari faktor akademik, demografi, dan kepribadian Big Five. Faktor akademik terdiri dari nilai harian 1, nilai harian 2, nilai harian 3, nilai PAS. Faktor demografi terdiri dari jenis kelamin, jenis tinggal, jumlah saudara kandung, anak ke, status pernikahan orang tua, pendidikan ayah, pendidikan ibu, pekerjaan ayah, pekerjaan ibu, penghasilan

ayah, dan penghasilan ibu. Faktor kepribadian siswa diukur melalui model Big Five

- h. Data kepribadian siswa diperoleh melalui penyebaran kuisioner menggunakan instrumen *Big Five Inventory* (BFI), yang terdiri dari 44 item yang dikembangkan untuk mengukur dimensi *openness*, *conscientiousness*, *extraversion*, *agreeableness*, dan *neuroticism*. Perhitungan skala likert terdiri dari 1-5, dimana 1=Sangat Tidak Sesuai, 2=Tidak Sesuai, 3=Netral, 4=Sesuai, dan 5=Sangat Sesuai
- i. Penyusunan kuesioner merujuk pada instrumen kepribadian *Big Five Inventory* (BFI) yang dikembangkan oleh (John et al., 1991; John et al., 2008)
- j. Validitas dan reliabilitas instrumen kepribadian Big Five diukur melalui validasi isi oleh psikolog, serta validasi konstruksi melalui analisis korelasi dan uji reliabilitas dengan *Cronbach's Alpha* yang akan diimplementasikan menggunakan *Python*
- k. Melakukan analisis hubungan antara kinerja akademik dengan faktor demografi dan kepribadian model *Big Five*.
- l. Model random forest akan digunakan untuk melakukan prediksi kinerja akademik pada mata pelajaran matematika dengan penentuan hyperparameter menggunakan metode grid search
- m. Pembagian dataset dilakukan dengan rasio 70% untuk data latih dan 30% untuk data uji
- n. Penanganan dataset tidak seimbang dilakukan dengan menerapkan teknik SMOTE

- o. Evaluasi performa model prediksi kinerja akademik siswa dilakukan berdasarkan metrik evaluasi, seperti *confusion matrix*, *accuracy*, *precision*, *recall*, *F1-Score*, *ROC-AUC*, dan *Cross Validation*

1.4. Tujuan Penelitian

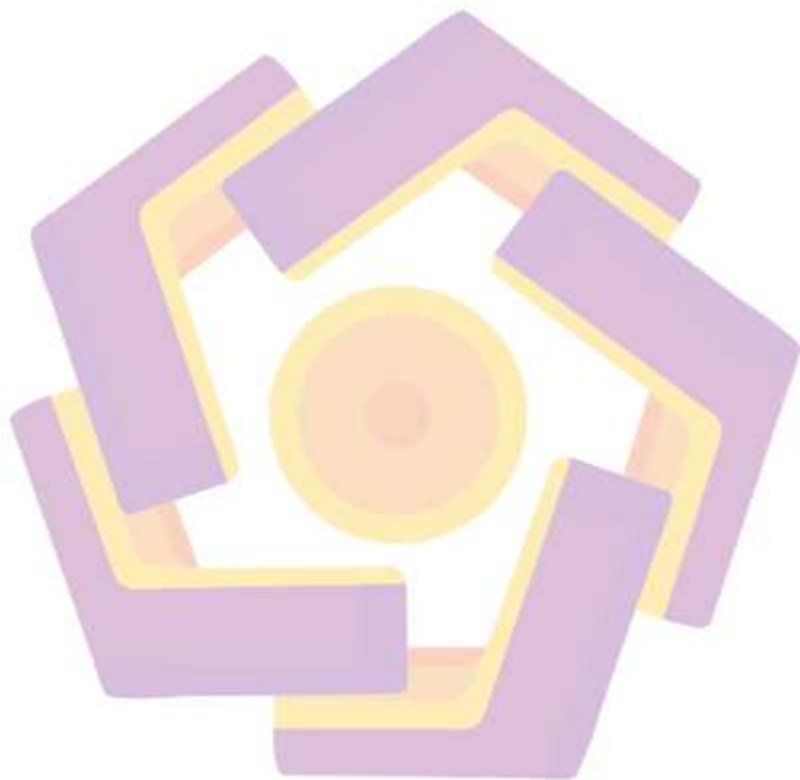
Tujuan dari penelitian ini adalah untuk menyelidiki atau mengeksplorasi hubungan antara variabel yang digunakan, seperti fitur akademik, demografi dan fitur kepribadian model Big Five terhadap nilai akhir atau kinerja akademik siswa. Penelitian ini juga bertujuan untuk mengevaluasi dampak penggunaan teknik SMOTE dalam mengatasi ketidakseimbangan dataset terhadap performa model Random Forest dalam melakukan prediksi kinerja akademik siswa. Selain itu, penelitian ini juga akan mengevaluasi bagaimana performa model Random Forest dalam melakukan prediksi kinerja akademik menggunakan berbagai metrik evaluasi.

1.5. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini ialah sebagai berikut :

- a. Memvalidasi atau memperbaiki pemahaman yang sudah ada mengenai bagaimana faktor-faktor tertentu dapat mempengaruhi kinerja akademik siswa.
- b. Menjadi landasan pengembangan strategi pendidikan untuk meningkatkan prestasi akademik siswa yang lebih efektif, yang dapat disesuaikan dengan karakteristik demografi dan kepribadian siswa.
- c. Memberikan kontribusi yang berharga bagi peneliti dan praktisi dalam menemukan serta menerapkan metode analisis prediksi kinerja akademik yang

paling efektif, sehingga dapat meningkatkan tingkat akurasi dan kehandalan dalam prediksi yang dihasilkan.



BAB II

TINJAUAN PUSTAKA

2.1. Tinjauan Pustaka

Tinjauan pustaka dilakukan untuk menemukan informasi yang relevan dengan topik dan untuk menghindari duplikasi data, serta untuk menegaskan bahwa kajian yang dilakukan oleh penulis dapat memberikan kontribusi penting terhadap perkembangan ilmu pengetahuan. Beberapa artikel penelitian terdahulu yang mencakup data dan metode yang digunakan penulis sebagai referensi dalam menyusun penelitian ini telah disertakan dibawah ini.

Educational Data Mining (EDM) memanfaatkan berbagai metode analisis data untuk memprediksi kinerja siswa berdasarkan aktivitas belajar, dengan penekanan terhadap topik prediksi kinerja akademik (Ozyurt et al., 2023). Penilaian kinerja akademik dilakukan untuk meningkatkan kualitas pendidikan (Issah et al., 2023) dan menciptakan pengalaman belajar yang optimal (Rodrigues et al., 2022), yang pada gilirannya dapat berkontribusi pada peningkatan tingkat keberhasilan akademik siswa (Dol & Jawandhiya, 2024). Namun, penelitian mengenai prediksi kinerja akademik di tingkat menengah masih berada pada tahap awal (Rodrigues et al., 2022), adapun prestasi pada jenjang ini memiliki dampak yang signifikan terhadap keputusan pendidikan dan pilihan karier di masa depan, karena masa ini merupakan periode penting dalam perkembangan individu, di mana perubahan fisik, emosional, dan sosial dapat mempengaruhi hasil akademik (Costa et al., 2024). Khususnya di Indonesia, prestasi matematika siswa masih tergolong rendah

(Wawan & Retnawati, 2022). Dalam penelitian ini fokus utama berada pada prediksi kinerja akademik siswa dalam mata pelajaran matematika di tingkat Sekolah Menengah Pertama (SMP), khususnya di SMPN 2 Kuningan, yang kondisinya dianggap relevan dengan topik penelitian ini.

Berbagai penelitian telah menunjukkan bahwa berbagai faktor, termasuk faktor akademik, demografi, dan kepribadian, dapat mempengaruhi prediksi kinerja akademik, di mana faktor akademik berperan sebagai indikator utama yang mempengaruhi kinerja akademik. Sebuah studi berhasil mengidentifikasi fitur nilai rata-rata kumulatif (GPA) dan nilai dari sekolah menengah dianggap sebagai indikator utama yang mempengaruhi keberhasilan akademik secara signifikan (Nachouki et al., 2023). Penelitian lain menunjukkan bahwa selain faktor akademik, faktor kepribadian juga dapat digunakan untuk memprediksi kinerja siswa. Sebuah studi mengungkapkan bahwa karakteristik model *Big Five* memiliki kemampuan untuk memprediksi kinerja akademik mahasiswa sarjana dan memberikan dampak yang signifikan terhadap prestasi akademik mereka (El-Keiey et al., 2022). Hasil penelitian tersebut menunjukkan dimensi *conscientiousness* dari model *Big Five* dan fitur akademik memiliki hubungan yang paling erat dengan kinerja akademik. Penelitian tersebut juga mengungkapkan bahwa lima dimensi kepribadian *Big Five* dengan fitur IQ, akademik, serta demografi, menghasilkan kinerja yang lebih optimal pada semua metrik evaluasi (akurasi, presisi, recall dan f1-measure) jika dibandingkan dengan analisis faktor-faktor tersebut secara individual. Penelitian lain juga menunjukkan bahwa kepribadian *Big Five* memiliki pengaruh terhadap kinerja akademik siswa melalui berbagai mekanisme. Sebagai contoh, siswa yang

memiliki skor tinggi dalam *conscientiousness* cenderung lebih aktif terlibat dalam tugas-tugas rumah, yang pada gilirannya dapat meningkatkan prestasi akademik mereka (Meyer et al., 2023). Penelitian lain juga menunjukkan bahwa tiga variabel kepribadian yang paling berhubungan dengan pencapaian akademik adalah *conscientiousness*, *openness*, dan *agreeableness* (Rico-Juan et al., 2024). Penelitian terbaru dengan menggunakan metode statistik, seperti *Partial Least Squares Structural Equation Modeling* (PLS SEM), telah berhasil menunjukkan bahwa semua dimensi pada model Big Five yang meliputi *openness*, *conscientiousness*, *extraversion*, *agreeableness*, dan *neuroticism*, memiliki pengaruh yang signifikan terhadap kinerja akademik mahasiswa (Shaninah & Mohd Noor, 2024). Selain itu penelitian (Shaninah & Mohd Noor, 2024) juga telah menyoroti signifikansi untuk mempertimbangkan variabel demografi, termasuk jenis kelamin dan latar belakang sosial-ekonomi, yang dapat berinteraksi dengan karakteristik kepribadian dan berdampak pada kinerja akademik siswa. Penelitian yang dilakukan oleh (Roslan & Chen, 2023) juga berhasil menunjukkan bahwa fitur-fitur pada faktor demografi seperti status sosial-ekonomi dan jenis kelamin memiliki dampak terhadap kinerja siswa sekolah menengah pada mata pelajaran bahasa inggris dan matematika. Temuan-temuan tersebut menunjukkan adanya keterkaitan antara beberapa faktor terhadap kinerja akademik, seperti faktor kepribadian model *Big Five* dan demografi. Dalam penelitian ini, penerapan faktor akademik, demografi, kepribadian model Big Five akan dilakukan guna memberikan kontribusi dalam melakukan analisis yang lebih dalam mengenai hubungan antara variabel-variabel tersebut dalam memprediksi kinerja akademik.

Meskipun penelitian yang dilakukan oleh (Shaninah & Mohd Noor, 2024) telah memanfaatkan berbagai fitur, termasuk aspek demografi dan kepribadian berdasarkan model Big Five, penelitian ini juga akan mengadopsi pendekatan yang serupa. Namun, penelitian tersebut menekankan perlunya pengembangan metode baru yang lebih mutakhir untuk memprediksi kinerja akademik. Dalam konteks ini, teknik data mining dengan metode klasifikasi menjadi sangat relevan, mengingat temuan dari (Roslan & Chen, 2022) telah menunjukkan bahwa metode klasifikasi merupakan pendekatan yang paling umum dan mudah dipahami, serta memiliki tingkat akurasi prediksi yang cukup baik. Selain itu, penelitian tersebut juga mengidentifikasi bahwa algoritma Naïve Bayes, Decision Tree dan Random Forest menjadi yang paling umum digunakan untuk melakukan prediksi akademik.

Penelitian yang dilakukan oleh (El-Keiey et al., 2022) berfokus pada pengembangan model prediksi kinerja akademik mahasiswa sarjana ke dalam tiga kategori (rendah, sedang, dan tinggi) dengan memanfaatkan algoritma Decision Tree (DT), Naïve Bayes (NB), dan Random Forest (RF). Hasil penelitian tersebut menunjukkan bahwa fitur nilai GPA, dimensi kepribadian *conscientiousness*, dan IQ memiliki dampak signifikan terhadap kinerja akademik. Akurasi model yang dihasilkan masing-masing mencapai 86% untuk NB, 83% untuk RF, dan 90% untuk DT. Selain itu, studi ini juga menekankan bahwa penghilangan fitur kepribadian menyebabkan penurunan signifikan dalam akurasi, presisi, recall, dan f1-score, yang semakin menegaskan pentingnya karakteristik kepribadian dalam memprediksi kinerja akademik. Penelitian lain yang dilakukan oleh (Santoso et al., 2024) menunjukkan bahwa metode Naïve Bayes (NB) dan Random Forest (RF)

dapat digunakan untuk mengembangkan model prediksi waktu kelulusan mahasiswa sarjana dengan mengelompokkan hasil ke dalam dua kategori, yaitu sukses dan tidak sukses. Hasil penelitian menunjukkan bahwa algoritma RF memiliki tingkat akurasi yang lebih baik dibandingkan dengan NB, dengan akurasi masing-masing sebesar 76% (RF) dan 64% (NB). Selain itu, penelitian ini juga mengidentifikasi bahwa IPK (GPA) sebagai fitur yang memiliki pengaruh signifikan terhadap waktu kelulusan mahasiswa.

Penelitian lain yang dilakukan oleh (Roslan & Chen, 2023) mengungkapkan bahwa algoritma Decision Tree (DT) dan Naïve Bayes (NB) berhasil menciptakan model prediksi kinerja untuk nilai ujian akhir siswa dalam mata pelajaran Bahasa Inggris dan Matematika, yang mengelompokkan ke dalam kategori kinerja rendah, sedang, dan tinggi. Temuan dari penelitian ini menunjukkan bahwa DT dan NB memperoleh tingkat akurasi masing-masing 87.1% dan 71% untuk Bahasa Inggris, sementara keduanya mencapai akurasi yang sama pada mata pelajaran Matematika, yaitu sebesar 83.9%. Penelitian ini menekankan bahwa penggunaan fitur nilai akademik dan fitur demografi, khususnya latar belakang orang tua siswa, memiliki dampak yang signifikan terhadap hasil prediksi. Penelitian lain juga menunjukkan bahwa Naïve Bayes, Decision Tree dan Random Forest efektif dalam memprediksi kelulusan mahasiswa pada mata kuliah statistik dengan menggunakan sebuah dataset yang terdiri dari 830 *instance* (Khairy et al., 2024). Hasil penelitian tersebut menunjukkan bahwa Random Forest dan Decision Tree memiliki performa terbaik dalam mengklasifikasikan kinerja mahasiswa, dengan Random Forest tetap lebih unggul dibandingkan Decision Tree. Hasil menunjukkan bahwa model Decision

Tree mengalami penurunan dari 99% menjadi 96% dalam *f-measure* saat proses validasi, sementara Random Forest menunjukkan konsistensi dalam semua aspek, dengan tingkat akurasi stabil di tingkat 99% (Khairy et al., 2024). Penelitian lain yang dilakukan oleh (Nachouki et al., 2023) mengungkapkan bahwa Random Forest berhasil mencapai akurasi di tingkat 90.33% dalam memprediksi nilai suatu mata kuliah, dengan fitur GPA dan skor sekolah menengah sebagai variabel utama.

Berdasarkan temuan dari berbagai penelitian menunjukkan bahwa metode Naïve Bayes, Decision Tree dan Random Forest cukup efektif dalam membangun model prediksi kinerja akademik dengan hasil yang memuaskan. Namun, mengingat bahwa prediksi kinerja akademik melibatkan berbagai fitur yang mungkin saling terkait, metode Naïve Bayes dianggap kurang sesuai karena tidak efektif dalam menangani interaksi antar fitur (Pajila et al., 2023; Santoso et al., 2024). Selain itu, dengan banyaknya fitur dalam dataset, Decision Tree cenderung mengalami *overfitting* dan kurang optimal (Mienye & Jere, 2024). Sementara itu, Random Forest dapat dianggap sebagai solusi yang efektif untuk mengatasi atau setidaknya mengurangi risiko *overfitting* (Mienye & Jere, 2024; Santoso et al., 2024), serta diketahui memiliki lebih banyak *hyperparameter* dibandingkan metode lainnya, yang dapat berdampak positif terhadap peningkatan kinerja model (Pecuchova & Drlik, 2023). Berdasarkan hal tersebut, dalam penelitian ini, Random Forest akan digunakan, karena telah menunjukkan tingkat akurasi yang cukup baik dan konsisten dalam mengembangkan model prediksi kinerja akademik, sebagaimana ditunjukkan oleh beberapa peneliti sejenis. Mengingat bahwa Random Forest memiliki jumlah *hyperparameter* yang lebih banyak dibandingkan

dengan metode lainnya, yang dapat mempengaruhi performa model (Pecuchova & Drlik, 2023), maka dalam penelitian ini penerapan teknik *hyperparameter tuning*, seperti *grid search* (Arden & Safitri, 2022; Pecuchova & Drlik, 2023; Rimal et al., 2024) akan dilakukan guna mengidentifikasi kombinasi parameter yang paling efektif, sehingga pada gilirannya dapat meningkatkan performa model prediksi.

Penelitian (Nachouki et al., 2023) juga menyoroti bahwa meskipun performa Random Forest konsisten, metode ini menghadapi tantangan ketika berhadapan dengan data yang tidak seimbang, khususnya dalam pengelompokan data dalam kategori yang tidak proporsional. Ketidakseimbangan dalam data ini perlu ditangani, karena dapat menyebabkan rendahnya akurasi pada kelas minoritas dan menghasilkan bias, yang pada gilirannya berdampak negatif terhadap kinerja model (Abdul Bujang et al., 2023). Temuan ini sejalan dengan hasil penelitian oleh (Dol & Jawandhiya, 2022), yang juga menyoroti permasalahan yang timbul akibat ketidakseimbangan dataset dalam EDM, dengan menekankan bahwa ketidakseimbangan data harus dibenahi karena dapat berdampak pada kinerja model. Metode *oversampling* seperti SMOTE (*Synthetic Minority Over-sampling Technique*), bisa menjadi salah satu solusi untuk mengatasi masalah ketidakseimbangan data dengan menambah jumlah *instance* dari kelas minoritas dengan menciptakan instance baru (Abdul Bujang et al., 2023). Meskipun penerapan SMOTE dalam model Random Forest tidak menunjukkan peningkatan akurasi, penelitian mengindikasikan adanya peningkatan pada presisi dan *recall* dari 72% dan 58% menjadi 79% dan 70%, yang menunjukkan efektivitasnya dalam meningkatkan kinerja analisis kepribadian pelanggan (Ghaniaviyanto Ramadhan &

Adiwijaya, 2022). Penelitian lain juga menyoroti peran penting SMOTE dalam meningkatkan akurasi algoritma *supervised*, khususnya Random Forest yang mencatat nilai presisi, recall, dan f1-measure tertinggi sebelum dilakukan seleksi fitur (Hamoud et al., 2022). Tercatat bahwa setelah proses seleksi fitur, akurasi mengalami penurunan menjadi 0.816 dari 0.83 akibat penghapusan beberapa fitur penting. Selain itu, penerapan SMOTE pada Random Forest menunjukkan peningkatan akurasi antara 1-3% dalam memprediksi mahasiswa yang berisiko gagal (Pecuchova & Drlik, 2023). Model tersebut juga berhasil mengidentifikasi variabel yang mempengaruhi hasil prediksi, seperti nilai ujian, tugas, dan ujian akhir semester. Namun, perlu diperhatikan bahwa SMOTE memiliki kelemahan yang mencakup risiko overfitting, noise, dan bias, karena SMOTE menghasilkan data sintesis yang terlalu mirip dengan data asli serta memaksa keseimbangan kelas tanpa memperhatikan distribusi alami, sehingga dapat mengurangi kemampuan generalisasi dan membuat model kurang peka terhadap karakteristik kelas minoritas (Abdul Bujang et al., 2023; Ghaniaviyanto Ramadhan & Adiwijaya, 2022). Berdasarkan temuan-temuan yang sudah dijelaskan, dalam tujuan untuk menangani ketidakseimbangan data pada model prediksi kinerja siswa pendekatan menggunakan teknik SMOTE akan dilakukan agar tercipta titik data tambahan dalam dataset pelatihan guna mencapai keseimbangan kelas data dan performa model menjadi lebih baik. Serta, evaluasi menggunakan berbagai metrik perlu dilakukan guna menggeneralisasi performa model.

2.2. Keaslian Penelitian

Tabel 2. 1. Matriks Literature Review dan Posisi Penelitian

Prediksi Kinerja Akademik Siswa Sekolah Menengah Pertama dalam Pelajaran Matematika dengan Metode Random Forest

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
1	Student's Performance Prediction based on Personality Traits and Intelligence Quotient using Machine Learning	Samar El-Keiey, Dina ElMenshawy, Ehab Hassancin ; International Journal of Advanced Computer Science and Application ; 2022	Penelitian ini bertujuan untuk melakukan prediksi prestasi akademik mahasiswa sarjana suatu universitas menggunakan metode machine learning dengan mempertimbangkan sifat kepribadian (Big Five), intelligence quotient (IQ) dan nilai akademik.	Penelitian ini menyimpulkan bahwa penggunaan berbagai algoritma machine learning (k-Nearest Neighbor (KNN), Support Vector Machine (SVM), Decision Tree, Naive Bayes, dan Random Forest) dengan Decision Tree menjadi model dengan akurasi terbaik (90%) dan fitur-fitur yang relevan dapat meningkatkan prediksi kinerja akademik mahasiswa. Penelitian ini menyoroti bahwa gender, jumlah saudara, sifat kepribadian dan IQ dapat mempengaruhi kinerja akademik mahasiswa. Namun, dibandingkan dengan	Algoritma Random Forest (83%) memiliki tingkat akurasi dibawah Decision Tree (90%) , hal ini menunjukkan bahwa masih ada ruang untuk mengeksplorasi terkait performa dari Random Forest. Model yang dibangun pada penelitian ini, hanya menggunakan beberapa faktor demografi, sehingga masih ada ruang untuk mengeksplorasi faktor demografi lainnya untuk mengetahui apakah faktor-faktor tersebut dapat mempengaruhi kinerja akademik.	Penelitian ini menggunakan beberapa faktor demografi dan menunjukkan bahwa algoritma Random Forest masih memiliki potensi untuk dieksplorasi lebih lanjut terkait performa modelnya. Penelitian yang akan dilakukan berupaya untuk memprediksi kinerja akademik siswa di sekolah menengah pertama menggunakan Random Forest, dengan menerapkan <i>hyperparameter tuning</i> menggunakan <i>grid search</i> , dalam upaya untuk meningkatkan performanya. Penelitian juga akan mempertimbangkan faktor-faktor tambahan selain faktor yang telah dipertimbangkan sebelumnya, seperti menggunakan faktor demografi sosial-ekonomi (status pernikahan orang tua, jenis tinggal, anak ke, pendidikan orang tua, pekerjaan orang tua, dan pendapatan orang tua).

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
				faktor IQ, faktor sifat kepribadian menjadi faktor yang paling mempengaruhi prediksi kinerja akademik.		
2	Student course grade prediction using the random forest algorithm: Analysis of predictors' importance	Mirna Nachouki, Elfadil A. Mohamed, Riyadh Mehdi, Mahmoud Abou Naaj ; Trends in Neuroscience and Education ; 2023	Penelitian ini bertujuan untuk mengusulkan model yang menggunakan algoritma Random Forest untuk memprediksi nilai matakuliah mahasiswa berdasarkan tujuh prediktor input (kategori kelas, kehadiran mahasiswa, jenis kelamin, tipe sekolah menengah, nilai sekolah menengah, penyampaian materi, dan nilai IPK) sehingga model tersebut dapat mengidentifikasi mahasiswa yang berisiko putus sekolah	Penelitian ini menemukan bahwa algoritma Random Forest dapat memprediksi nilai matakuliah mahasiswa dengan tingkat akurasi sebesar 90,33%. Hasil ini menunjukkan bahwa model ini efektif dalam memprediksi keberhasilan akademik mahasiswa. Model ini juga menunjukkan bahwa nilai IPK dan nilai sekolah menengah, menjadi faktor paling penting atau memiliki hubungan yang paling signifikan dengan nilai matakuliah.	Penelitian ini mengelompokkan data ke dalam kategori-kategori pada sebuah kelas, namun pembagian tersebut tidak seimbang sehingga dapat mengakibatkan bias terhadap performa model. Kemudian, para peneliti merekomendasikan agar penelitian di masa depan dapat mengeksplorasi algoritma-algoritma lain dan mengeksplorasi juga faktor-faktor lain, seperti demografis dan karakteristik pribadi untuk menentukan pengaruhnya terhadap kinerja akademik.	Penelitian ini menunjukkan bahwa meskipun Random Forest telah berhasil membangun model dengan performa yang baik, namun terdapat permasalahan pada ketidakseimbangan data. Penelitian yang akan dilakukan berupaya untuk mengatasi masalah ketidakseimbangan data dengan menerapkan metode oversampling, seperti teknik SMOTE. Serta, menerapkan fitur-fitur lain yang masih relevan dengan kinerja akademik, seperti faktor demografi dan kepribadian model Big Five.
3	Predicting Students Performance in English and	Muhammad Haziq Bin Roslan & Chewn Jen Chen ; Education and	Penelitian ini memiliki tujuan untuk melakukan prediksi kinerja siswa di sekolah menengah	Hasil penelitian menunjukkan bahwa algoritma yang digunakan berhasil	Penelitian ini memiliki keterbatasan di mana sample yang dimiliki ukurannya relatif kecil	Penelitian ini menunjukkan bahwa empat algoritma yang digunakan dapat digunakan untuk membangun model prediksi kinerja akademik

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	Mathematics Using Data Mining Techniques	Information Technologies ; 2023	dalam pelajaran Bahasa Inggris dan Matematika menggunakan metode data mining, yakni dengan menerapkan teknik klasifikasi seperti algoritma Decision Tree, Neural Networks, SVM, dan Naive Bayes. Model dibangun dengan tujuan untuk mengungkapkan karakteristik siswa yang memiliki kinerja rendah, sedang dan tinggi, serta dapat memberikan kinerja yang baik untuk kedua mata pelajaran.	memprediksi kinerja siswa dalam pelajaran bahasa Inggris dan Matematika. di mana Decision Tree memiliki kinerja yang paling baik pada mata pelajaran Bahasa Inggris, dan Naive Bayes memiliki kinerja terbaik untuk mata pelajaran Matematika. Penelitian ini juga berhasil mengungkapkan bahwa faktor nilai akademik sebelumnya merupakan faktor paling berpengaruh dalam memprediksi kinerja siswa dalam kedua mata pelajaran tersebut. Selain itu, atribut demografi seperti tingkat pendidikan orang tua dan jenis kelamin juga memiliki peran pada kinerja akademik meskipun tidak sebesar faktor kinerja akademik sebelumnya.	(159 siswa), sehingga hasilnya tidak dapat digeneralisasi secara luas. Kemudian, penelitian ini menyarankan untuk melakukan penelitian yang lebih mendalam dalam mengeksplorasi teknik data mining yang lain dan faktor-faktor lain guna mengungkap informasi yang belum terungkap	siswa, dengan hasil Decision Tree dan Naive Bayes menjadi yang terbaik. Penelitian yang akan dilakukan memiliki fokus untuk memprediksi kinerja akademik siswa sekolah menengah pertama khususnya pada mata pelajaran Matematika dengan memanfaatkan algoritma lain, seperti, Random Forest dengan <i>hyperparameter tuning (grid search)</i> untuk meningkatkan performa model. Selain mempertimbangkan faktor-faktor yang telah terbukti mempengaruhi kinerja model (demografi dan akademik), penelitian juga akan mempertimbangkan penambahan faktor lain seperti sifat kepribadian (Big Five).

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
4	Prediction of student exam performance using data mining classification algorithms	Daila Khairy, Nouf Alharbi, Mohamed A Amasha, Marwa F Areeed, Salem Alkhalaf, Rania A Abougalala ; Education and Information Technologies ; 2024	Tujuan penelitian ini adalah untuk melakukan prediksi kinerja hasil ujian mahasiswa sarjana menggunakan algoritma klasifikasi data mining dengan membandingkan kinerja dari lima algoritma machine learning, yaitu Random Forest, Decision Tree, Naive Bayes, Neural Networks, dan K-Nearest Neighbours. Penelitian ini berusaha untuk mengidentifikasi algoritma yang paling efektif untuk memprediksi kinerja mahasiswa.	Model dibangun dengan menerapkan variabel tahun akademik, ujian tengah semester, ujian praktikum, ujian tertulis, total nilai akhir dan kinerja nilai akhir (lima kelas). Berdasarkan hasil penelitian ini, dari kelima algoritma yang digunakan, Random Forest dan Decision Tree menunjukkan kinerja terbaik, dengan tingkat akurasi maksimum mencapai 98,7%. Meskipun begitu, akurasi Decision Tree sangat sesuai dengan set validasi, serta nilai F-Measure nya menurun dari 99% menjadi 96%. Dengan demikian, model Random Forest dianggap sebagai model yang paling stabil dan efektif dalam memprediksi kinerja mahasiswa. Kemudian, enam variabel yang digunakan dalam penelitian ini	Penelitian ini hanya menggunakan fitur akademik dalam membangun model prediksi kinerja. Fitur-fitur lain yang mungkin relevan tidak disertakan, yang dapat membatasi kemampuan model untuk mengungkap faktor-faktor yang dapat mempengaruhi kinerja akademik. Meskipun penelitian menunjukkan performa yang cukup baik, temuan yang diperoleh menunjukkan adanya ketidakseimbangan dalam pembagian kelas pada dataset, di mana kelas F mendominasi dibandingkan dengan kelas lainnya. Oleh karena itu, penanganan terhadap ketidakseimbangan data menjadi sangat penting agar kinerja model dapat digeneralisasikan.	Penelitian ini menyoroti penggunaan faktor akademik memiliki pengaruh signifikan, namun penggunaan fitur lain yang relevan dengan kinerja akademik masih perlu dilakukan. Serta, penanganan terhadap ketidakseimbangan kelas dalam dataset sangat diperlukan untuk menggeneralisasikan kinerja model. Penelitian yang akan dilakukan bertujuan untuk mengeksplorasi fitur-fitur lain selain akademik, seperti sifat kepribadian (Big Five) dan demografi pada siswa dengan tetap memperhatikan faktor akademik. Serta, dalam upaya meningkatkan performa Random Forest, <i>hyperparameter tuning (grid search)</i> akan diterapkan. Teknik SMOTE akan diterapkan pada model prediksi untuk menangani ketidakseimbangan data, sehingga dapat menghasilkan model yang lebih akurat.

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
				(tahun akademik, ujian tengah semester, ujian praktikum, ujian tulis, total nilai akhir, dan nilai akhir) terbukti penting dan signifikan dalam memprediksi kinerja mahasiswa.		
5	Predicting Time to Graduation of Open University Students: An Educational Data Mining Study	Agus Santoso, Heri Retnawati, Ezi Apino, Ibnu Rafi, Munaya Nikma Rosyada ; Open Education Studies ; 2024	Penelitian ini bertujuan untuk membangun model dengan menggunakan teknik data mining untuk memprediksi waktu kelulusan mahasiswa di Indonesia, serta mengidentifikasi faktor-faktor yang dapat menentukan kesuksesan akademik mahasiswa berdasarkan waktu kelulusan atau waktu yang dibutuhkan untuk menyelesaikan studi	Hasil penelitian menunjukkan bahwa model prediksi yang dibangun menggunakan beberapa algoritma seperti Random Forest, Neural Networks, Logistic Regression, Naive Bayes, dan K-Nearest Neighbors dapat secara efektif memprediksi waktu kelulusan mahasiswa. Algoritma Random Forest dan Neural Networks ditemukan sebagai algoritma dengan tingkat akurasi klasifikasi (CA) dan nilai Area Under The Curve (AUC) tertinggi dalam memprediksi	Penelitian ini tidak mempertimbangkan faktor-faktor lain yang mungkin mempengaruhi kinerja prediksi. Penelitian selanjutnya dapat difokuskan pada pengembangan dan peningkatan salah satu algoritma yang digunakan. Serta, selain menambahkan faktor-faktor lain yang dapat mempengaruhi kinerja prediksi, penelitian selanjutnya juga dapat melakukan analisis lebih lanjut terhadap faktor-faktor yang telah diidentifikasi sebagai faktor yang secara signifikan berpengaruh	Penelitian ini hanya mempertimbangkan faktor akademik dan demografi. Serta, perlunya penanganan ketidakseimbangan data. Penelitian yang akan dilakukan menambahkan fitur lain, seperti sifat kepribadian (Big Five) sambil memperhatikan faktor demografi dan akademik yang memiliki pengaruh dengan fokus menggunakan algoritma Random Forest yang berdasarkan penelitian ini telah terbukti memiliki performa terbaik diantara algoritma lain, dengan menerapkan <i>hyperparameter tuning (grid search)</i> dalam upaya meningkatkan performa model. Serta, penelitian ini akan memperhatikan atau menyelesaikan permasalahan data

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
				waktu kelulusan mahasiswa, dengan nilai CA mencapai 76% dan AUC mencapai 79%. Meskipun begitu, Random Forest lebih unggul 0,2 - 0,3 % dari Neural Networks. Penelitian ini juga berhasil mengidentifikasi atribut yang mempengaruhi kinerja prediksi, seperti IPK, usia, status pekerjaan, dan upah minimum dengan variabel IPK menjadi faktor utama yang mempengaruhi.	dalam kinerja prediksi. Penelitian ini juga menyoroti perlunya memperhatikan kelas pada dataset agar seimbang atau menyelesaikan permasalahan terkait data tidak seimbang agar hasil prediksi tidak bias.	tidak seimbang dengan menerapkan teknik SMOTE.
6	The impact of big five personality trait in predicting student academic performance	Fathi Said Emhemed Shaninah & Mohd Halim Mohd Noor ; Journal of Applied Research in Higher Education ; 2024	Penelitian ini bertujuan untuk mengembangkan model prediktif menggunakan metode statistik, seperti Partial Least Squares Structural Equation Modeling (PLS-SEM) untuk memprediksi kinerja akademik mahasiswa dengan menggabungkan faktor	Hasil penelitian menunjukkan bahwa model prediktif yang diusulkan, mampu memprediksi kinerja akademik dengan tepat. Penelitian ini menemukan bahwa faktor kepribadian dapat menjadi prediktor yang lebih kuat untuk kinerja akademik dibandingkan	Penelitian ini hanya dilakukan di institusi pendidikan tingkat tinggi (universitas tingkat sarjana), sehingga diharapkan penelitian selanjutnya dapat memperluas penelitian di tingkat pendidikan yang lainnya. Serta, penelitian meyoroti untuk mengusulkan metode lain	Penelitian ini memiliki fokus pada pendidikan di tingkat tinggi, serta menggunakan model prediktif Partial Least Squares Structural Equation Modeling (PLS-SEM) Penelitian yang akan dilakukan mengeksplorasi pendidikan di tingkat lain, seperti sekolah menengah pertama yang masih kurang diteliti. Penelitian ini juga akan berusaha membangun model prediksi kinerja akademik dengan

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			kepribadian (Big Five) dan demografis.	dengan motivasi saja, terutama domain kepribadian big five conscientiousness yang memiliki pengaruh signifikan terhadap kinerja akademik. Kemudian, faktor demografi seperti jenis kelamin dan latar belakang sosial-ekonomi juga berinteraksi dengan faktor kepribadian untuk mempengaruhi kinerja akademik mahasiswa	yang lebih modern untuk melakukan prediksi kinerja akademik	pendekatan baru menggunakan teknik data mining dengan menerapkan algoritma Random Forest, serta tetap memperhatikan faktor-faktor yang telah terbukti memiliki hubungan signifikan terhadap kinerja akademik siswa

2.3. Landasan Teori

2.3.1 Educational Data Mining

Menurut (Chakraborty et al., 2022) data mining merupakan suatu metode analisis yang digunakan untuk mengidentifikasi pola dan hubungan yang belum diketahui dalam kumpulan data yang besar. Proses ini melibatkan penerapan model statistik, algoritma dan metodologi *machine learning* untuk mengekstraksi informasi yang signifikan. Menurut (Ha et al., 2011) juga menyebutkan bahwa data mining berfungsi untuk mengubah data mentah menjadi informasi yang bernilai dengan cara mengenali pola-pola tertentu. Tahapan ini melibatkan beberapa tahapan, seperti membersihkan, mengintegrasikan, melakukan seleksi, mentransformasi, menambang, mengevaluasi dan presentasi data.

Educational Data Mining (EDM) merupakan bidang yang berfokus pada pengembangan teknik untuk menganalisis data yang berasal dari konteks pendidikan (Pena-Ayala, 2013). EDM termasuk dalam ranah ilmu komputer dan bertujuan untuk mengekstrak, serta memanfaatkan data guna mendapatkan informasi berharga dalam konteks pendidikan melalui penerapan teknik data *mining* (Sweta, 2024). Metode data *mining* dalam konteks pendidikan melibatkan berbagai teknik yang berasal dari *machine learning*, statistik, dan sistem basis data (Pena-Ayala, 2013). Teknik klasifikasi merupakan salah satu teknik yang dapat diterapkan, dengan fungsi mengelompokkan data ke dalam kategori yang telah ditetapkan. Proses ini mencakup analisis karakteristik objek baru dan penempatannya ke dalam kelas yang sudah ada. Teknik EDM sendiri sering digunakan untuk memperdalam pemahaman mengenai siswa dan konteks

pembelajaran, serta untuk mengidentifikasi informasi yang tidak terlihat dalam data. Melalui analisis data mentah yang diperoleh dari sistem pendidikan, EDM berfungsi sebagai alat yang sangat berguna bagi para pemangku kepentingan dalam upaya meningkatkan kinerja dan memahami perilaku siswa dalam proses belajar mengajar (Pena-Ayala, 2013).

2.3.2 Kinerja Akademik

Kinerja akademik merupakan salah satu bidang yang dalam prosesnya menerapkan teknik *educational data mining* dengan tujuan untuk menganalisis dan memprediksi kinerja atau nilai siswa, kesuksesan atau kegagalan, serta promosi siswa dalam proses pembelajaran mereka (Sweta, 2024). Penilaian kinerja akademik siswa didasarkan pada berbagai faktor, seperti data demografi, status sosial-ekonomi, kepribadian dan prestasi akademik mereka (Pena-Ayala, 2013). Secara keseluruhan, teknik ini digunakan untuk menilai dan memprediksi kinerja siswa berdasarkan data akademik yang tersedia pada data pendidikan, dengan mempertimbangkan juga aktivitas belajar dan umpan balik siswa, hal ini dilakukan dengan tujuan untuk mendapatkan umpan balik dari pelajar dan manajemen pendidikan.

2.3.3 Kepribadian Big Five

Model kepribadian *Big Five* pertama kali ditemukan oleh dua peneliti Angkatan udara AS, Ernest Tupes dan Raymond Christal, pada tahun 1961 (McCrae & Jr Costa, 2006). Hasil penelitian mereka mengungkapkan bahwa dari 35 skala penilaian yang dikembangkan oleh Cattell, hanya lima faktor yang

konsisten di berbagai sampel. Pada tahun 1980-an, Costa dan McCrae berperan dalam penemuan kembali model *Big Five* (Five Factor Model, FFM) dan membuktikan bahwa model ini merupakan model yang komprehensif untuk struktur sifat kepribadian, dengan mencakup hampir semua sifat yang diidentifikasi dalam bahasa umum dan teori ilmiah (McCrae & Jr Costa, 2006). Costa dan McCrae juga melakukan uji validitas model ini melalui berbagai instrument dan pengamatan, dan hasil menunjukkan bahwa lima faktor ini dapat diukur dengan akurasi yang dapat diterima.

Model *Big Five* merupakan kerangka yang digunakan untuk mengukur sifat kepribadian manusia dengan menggunakan lima dimensi, seperti *Neuroticism* (N), *Extraversion* (E), *Openness* (O), *Agreeableness* (A), dan *Conscientiousness* (C) (McCrae & Jr Costa, 2006). Setiap dimensi yang digunakan dalam model *Big Five* untuk mengukur kepribadian seseorang akan dijelaskan di bawah ini :

- 1) *Neuroticism* : Karakteristik ini mencakup kecenderungan individu merasakan emosi negative seperti kecemaasan, kemarahan atau depresi. Individu dengan tingkat *neurotisme* tinggi cenderung lebih mudah merasa cemas, tegang, dan khawatir.
- 2) *Extraversion* : Karakteristik ini mencakup kecenderungan individu bersikap sosial, aktif, dan bersemangat. Individu dengan tingkat *extraversion* tinggi cenderung bersikap ramah, energik, dan senang berinteraksi dengan orang lain.
- 3) *Openness* : Karakteristik ini mencakup kecenderungan individu menerima pengalaman baru, memiliki daya imajinasi yang tinggi, dan

memiliki minat yang beragam. Individu dengan tingkat keterbukaan tinggi cenderung lebih kreatif, ingin tahu, dan terbuka terhadap ide-ide baru.

- 4) *Agreeableness* : Karakteristik ini mencakup kecenderungan individu bersikap ramah, kooperatif, dan penuh empati. Individu dengan tingkat *agreeableness* tinggi cenderung lebih mudah bekerja sama, penuh perhatian, dan peduli terhadap kesejahteraan orang lain.
- 5) *Conscientiousness* : Karakteristik ini mencakup kecenderungan individu bersikap disiplin, teratur, dan bertanggung jawab. Individu dengan tingkat *conscientiousness* tinggi cenderung lebih teliti, terorganisir, dan memiliki motivasi yang kuat untuk mencapai tujuan.

Data karakteristik pribadi dapat dikumpulkan melalui pengisian kuesioner survei menggunakan *Big Five Inventory* (BFI) (John et al., 1991; John et al., 2008), hal ini dilakukan untuk mengukur kepribadian individu berdasarkan lima faktor (Widiger, 2015). BFI atau *Big Five Inventory* merupakan alat ukur kepribadian yang diciptakan oleh John, Donahue, dan Kentle pada tahun 1991. BFI digunakan sebagai metode pengukuran atau *instrument* yang didasarkan pada pernyataan singkat lima faktor utama kepribadian yang terdiri dari 44 item yang dinilai menggunakan skala lima poin dari 1 (sangat tidak sesuai) hingga 5 (sangat sesuai) (Widiger, 2015). BFI sering digunakan untuk menganalisis karakteristik kepribadian pada anak-anak yang berusia 10 tahun ke atas serta untuk meneliti perkembangan kepribadian dari usia 10 hingga 65 tahun (Widiger, 2015). Alat ini juga digunakan dalam berbagai konteks, seperti bisnis, industri, dan penelitian pendidikan. Lima faktor yang diterapkan merupakan prinsip dasar yang konsisten

sepanjang masa dewasa dan berperan penting dalam membentuk kehidupan individu, di mana sebagian besar karakteristik kepribadian dapat dipahami melalui dimensi dalam model ini (McCrae & Jr Costa, 2006).

2.3.4 Klasifikasi

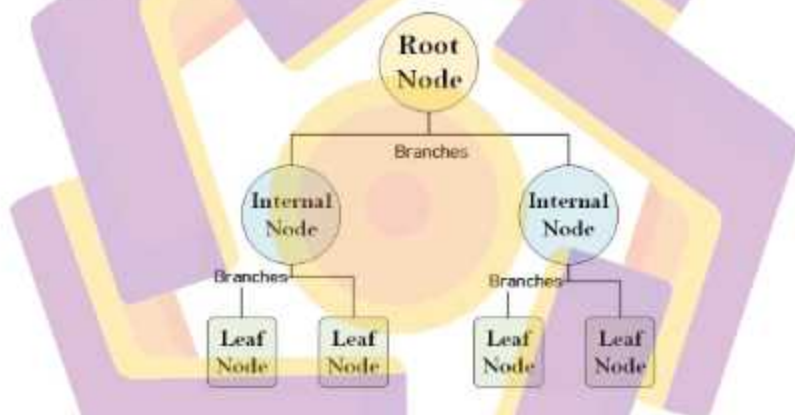
Klasifikasi dalam *supervised learning* merupakan proses untuk memprediksi label kelas yang telah ditentukan sebelumnya dari daftar kemungkinan (Muller & Guido, 2016). Pada dasarnya konsep dari teknik klasifikasi adalah mengelompokkan setiap instance ke dalam kategori tertentu berdasarkan label yang sudah ditentukan sebelumnya (Géron, 2019). Teknik ini bertujuan untuk memprediksi label kelas dari data input (Muller & Guido, 2016). Data pelatihan pada *supervised learning* yang diberikan pada algoritma mencakup solusi yang diinginkan atau disebut label. Pada algoritma dengan teknik klasifikasi, algoritma harus belajar bagaimana mengklasifikasikan data baru berdasarkan label yang sudah ada, artinya algoritma dilatih menggunakan dataset yang telah diberi label (Géron, 2019).

Menurut (Muller & Guido, 2016) klasifikasi sering kali terbagi menjadi klasifikasi biner dan klasifikasi multikelas. Klasifikasi biner merupakan proses membedakan antara dua kelas, seperti menjawab pertanyaan ya/tidak atau membuat satu kelas menjadi kelas positif dan kelas lainnya menjadi kelas negatif. Sedangkan klasifikasi multikelas melibatkan lebih dari dua kelas dimana hasil dari klasifikasi ini lebih kompleks karena melibatkan lebih banyak kelas dan hasilnya lebih sulit dipahami. Sebagai contoh, memprediksi bahasa yang digunakan dalam situs web

dari teks di situs web, dimana kelas-kelasnya adalah daftar bahasa yang telah ditentukan sebelumnya.

2.3.5 Random Forest

Algoritma Random Forest merupakan teknik *ensemble* yang menggabungkan beberapa Decision Tree (Raschka & Mirjalili, 2017). Menurut (Mienye & Jere, 2024) prosedur pembelajaran Decision Tree terdiri dari serangkaian langkah di mana data dipecah menjadi subset yang seragam, seperti yang dapat dilihat pada Gambar 2.1.



Gambar 2.1. Prosedur Decision Tree

Root node (node akar) merupakan titik awal dalam struktur pohon yang menggambarkan keseluruhan dataset. Algoritma akan mengidentifikasi fitur serta ambang batas yang menghasilkan pembagian paling optimal berdasarkan kriteria tertentu. Proses ini berlangsung secara berulang, di mana setiap subset data akan dibagi lebih lanjut di node anak (Internal Node). Pembagian ini akan terus berlanjut hingga mencapai kriteria penghentian, seperti ketika node menjadi murni atau

ketika kedalaman pohon yang telah ditentukan tercapai. Node terakhir (node leaf) menggambarkan hasil atau label kelas.

Keputusan untuk pembagian di setiap node didasarkan pada rumus matematis seperti *Information Gain* (IG). IG merupakan rumus matematis yang didasarkan pada prinsip *Entropy* dalam teori informasi (Mienye & Jere, 2024). *Entropy* berfungsi untuk mengukur derajat ketidakpastian atau acak dalam suatu kumpulan data. IG mengevaluasi seberapa efisien atribut dalam membagi data pelatihan menjadi kelompok yang seragam. *Entropy* (E) dari suatu set (himpunan) S dihitung dengan menggunakan rumus sebagai berikut (Mienye & Jere, 2024) :

$$E(S) = \sum_{i=1}^n p_i \log_2(p_i) \quad (1)$$

Dengan keterangan n merupakan jumlah kelas unik dalam himpunan data, dan p_i menunjukkan proporsi sampel dalam himpunan yang tergolong dalam kelas i . Maka, persamaan *Information Gain* (IG) untuk pemisahan pada dataset S berdasarkan atribut A dapat dihitung dengan cara berikut (Mienye & Jere, 2024) :

$$IG = - \sum_{v \in \text{values}(A)} \frac{|S_v|}{|S|} E(S_v) \quad (2)$$

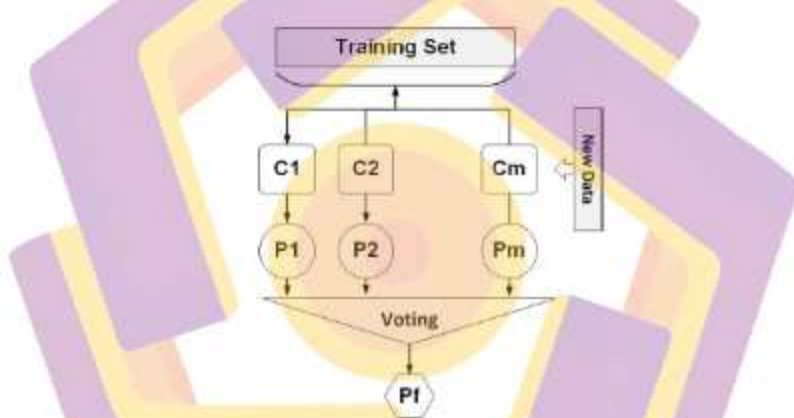
Dengan keterangan bahwa $\text{Values}(A)$ merupakan berbagai nilai yang mungkin dimiliki oleh atribut A , dan S_v merupakan subset dari S di mana atribut A memiliki nilai v . Secara garis besar, persamaan ke-2 ini menghitung perubahan *entropy* dari set awal (S) ke set S_v yang terbentuk setelah proses pembagian. Nilai IG yang lebih tinggi mencerminkan atribut yang lebih efisien dalam membagi data, karena mampu menghasilkan subset yang lebih homogen atau seragam.

Random Forest merupakan algoritma *ensemble* yang dikenal dengan istilah *bootstrap aggregation* atau *bagging* (Brownlee, 2016). Sebelum membahas konsep

bagging, penting untuk memahami konsep *bootstrap*, yang merupakan teknik statistik yang digunakan untuk memperkirakan parameter dari suatu kumpulan data berdasarkan sampel, seperti estimasi rata-rata atau standar deviasi. Proses *bootstrap* memastikan bahwa setiap pohon keputusan dalam random forest dibangun dari kumpulan data yang sedikit bervariasi, sementara pemilihan fitur di setiap node memungkinkan setiap pemisahan untuk menggunakan subset fitur yang berbeda (Muller & Guido, 2016). Metode *bagging* itu sendiri mengintegrasikan prediksi dari berbagai algoritma *machine learning* untuk meningkatkan akurasi dibandingkan dengan model tunggal (Brownlee, 2016). Teknik ini berfungsi untuk mengurangi varians, terutama pada algoritma yang memiliki varians tinggi seperti pohon keputusan, yang sangat dipengaruhi oleh data pelatihan. Implementasi *bagging* mencakup pembuatan sub-sampel acak dari dataset dengan penggantian (*replacement*), sehingga elemen yang sama dapat terpilih lebih dari satu kali. Selanjutnya, model pohon keputusan dilatih pada setiap sub-sampel dan rata-rata prediksi dari semua model dihitung untuk menghasilkan prediksi akhir.

Konsep dasar dari random forest melibatkan penggabungan sejumlah pohon keputusan yang memiliki kedalaman dan varians yang tinggi, dengan tujuan untuk meningkatkan akurasi prediksi dan mengurangi risiko *overfitting* (Muller & Guido, 2016). Dalam proses prediksi, random forest memanfaatkan sejumlah pohon keputusan yang memanfaatkan elemen keacakan selama proses pembangunan (Muller & Guido, 2016). Keacakan dalam proses ini dicapai dengan memilih titik data dan fitur secara acak pada saat pembuatan setiap pohon. Sebagai contoh, jika model menghasilkan 10 pohon, setiap pohon tersebut dibangun secara terpisah atau

independent tanpa ketergantungan satu sama lain. Algoritma ini menetapkan kelas akhir dengan mengintegrasikan hasil dari berbagai pohon melalui metode *majority voting* (Gambar 2.2), di mana label kelas yang memperoleh suara terbanyak (modus) dari pengklasifikasian dipilih, yaitu memperoleh lebih dari 50% suara (Raschka & Mirjalili, 2017). Artinya setiap pohon memberikan prediksi probabilitas untuk setiap label, dan kelas yang memiliki probabilitas tertinggi akan menjadi prediksi akhir (Muller & Guido, 2016).



Gambar 2. 2. Majority Voting (Raschka & Mirjalili, 2017)

Secara sederhana, algoritma random forest dapat diringkas dalam empat langkah, yaitu sebagai berikut (Raschka & Mirjalili, 2017) :

- i. Ambil sampel acak sebanyak n dari kumpulan data latih dengan penggantian (*replacement*), yang dikenal sebagai bootstrap
- ii. Buatlah pohon keputusan dengan menggunakan sampel bootstrap. Pada setiap simpul :
 - i. Pilih fitur d secara acak tanpa penggantian

- ii. Bagi simpul berdasarkan fitur yang memberikan pemisahan terbaik sesuai dengan fungsi tujuan, seperti memaksimalkan informasi yang diperoleh
- iii. Lakukan langkah-langkah ini (1-2) sebanyak N kali
- iv. Gabungkan prediksi dari setiap pohon untuk menentukan label kelas dengan mayoritas suara atau suara terbanyak

2.3.6 Hyperparameter

Dalam *machine learning*, terdapat dua kategori parameter, yaitu parameter yang diperoleh dari data pelatihan dan dioptimalkan secara terpisah, serta parameter penyesuaian (tuning) atau *hyperparameter* yang terdapat dalam suatu model, seperti parameter kedalaman pada pohon keputusan (Raschka & Mirjalili, 2017). Random Forest sendiri memiliki sejumlah parameter penting yang dapat disesuaikan untuk meningkatkan efektivitas model (Muller & Guido, 2016). Parameter tersebut mencakup *n_estimators* yang menentukan jumlah pohon dalam hutan. Selain itu, terdapat *max_features* yang mengatur jumlah fitur yang dipertimbangkan pada setiap *split* untuk mengendalikan keragaman pohon, serta *random_state* yang menjamin bahwa hasil dapat diulang dengan menetapkan kondisi acak. Parameter lain, seperti *max_depth* yang mengatur kedalaman maksimum setiap pohon, *min_samples_split* yang mengatur jumlah minimum sampel untuk memecah sebuah *node*, dan *min_samples_leaf* yang mengatur jumlah minimum sampel pada setiap daun pohon.

2.3.7 Grid Search

Memahami arti dari setiap parameter dan menetapkan nilai yang sesuai untuk mencapai performa generalisasi yang optimal merupakan tugas yang kompleks. *Grid Search* adalah metode optimasi *hyperparameter* yang digunakan untuk menentukan kombinasi nilai *hyperparameter* yang paling optimal. Metode *grid search* pada dasarnya melibatkan pengujian semua kombinasi parameter yang relevan guna menemukan parameter yang paling efektif untuk meningkatkan kinerja model (Muller & Guido, 2016). Pendekatannya pun tergolong sederhana, karena melibatkan pencarian menyeluruh melalui metode *brute-force*, di mana pengguna menetapkan daftar nilai untuk berbagai *hyperparameter*, dan komputer akan menguji setiap kombinasi nilai untuk mengidentifikasi yang paling optimal (Raschka & Mirjalili, 2017). Namun, metode ini tidak dapat secara otomatis memfokuskan perhatian pada kombinasi yang telah terbukti efektif. Karena untuk mendapatkan hasil terbaik, metode ini umumnya dimulai dengan ruang pencarian yang luas dan langkah yang besar, kemudian secara bertahap mempersempit ruang dan langkah tersebut berdasarkan hasil dari kombinasi *hyperparameter* yang paling berhasil (Chowdhury et al., 2022)

2.3.8 Teknik SMOTE

SMOTE (*Synthetic Minority Over-sampling Technique*) merupakan metode yang dikembangkan untuk mengatasi ketidakseimbangan kelas dalam dataset dengan cara mengubah bias algoritma untuk kelas minoritas (He & Ma, 2013). Metode ini diperkenalkan dan dikembangkan oleh (Chawla et al., 2002) dengan tujuan untuk meningkatkan akurasi klasifikasi untuk kelas minoritas dengan

menciptakan contoh sintetis dari data yang ada. Berdasarkan pemaparan yang dilakukan oleh (Chawla et al., 2002), untuk setiap contoh minoritas, teknik SMOTE akan memilih salah satu dari k tetangga terdekatnya.

Pada dasarnya metode ini menggunakan pendekatan berbasis *k-nearest neighbors* dengan cara menghitung jarak *Euclidean* dan memilih tetangga terdekat dari contoh minoritas yang dipilih (Chawla et al., 2002). Proses pembuatan data sintetis dilakukan dengan mengidentifikasi kepadatan contoh dalam kelas minoritas, sehingga SMOTE dapat menemukan contoh nyata yang memberikan informasi penting mengenai label kelas minoritas serta *outlier* yang berpotensi mempengaruhi hasil secara negatif (Azhar et al., 2023). Secara garis besar, data sintesis untuk kelas minoritas dihasilkan dengan cara memilih sampel dari kelas tersebut dan menghasilkan sampel baru di antara pilihan yang ada serta tetangga terdekat, tanpa menghapus sampel yang sudah ada, sehingga memastikan bahwa informasi tetap utuh dan tidak ada yang hilang. Persamaan terkait prosedur teknik SMOTE dapat dilihat dalam bentuk berikut (Azhar et al., 2023) :

$$D_{new} = D_i + (D_j - D_i) \times w \quad (3)$$

Keterangan :

- D_{new} = Data sintesis baru
- D_i = Contoh kelas minoritas
- D_j = Salah satu dari k -tetangga terdekat dari D_i
- w = Angka acak antara 0 dan 1

2.3.9 Evaluasi Metode

Evaluasi metode klasifikasi dilakukan dengan memisahkan data menjadi set pelatihan dan set pengujian, lalu mengevaluasi kinerja model pada set pengujian (Muller & Guido, 2016). Evaluasi model klasifikasi dilakukan dengan menggunakan berbagai metrik evaluasi sesuai dengan tujuan akhir dari model. Salah satu metode yang efektif untuk mengevaluasi performa pengklasifikasi adalah dengan memeriksa *confusion matrix*. *Confusion matrix* umumnya menghitung seberapa sering sampel dari kelas A salah diklasifikasikan sebagai kelas B (Géron, 2019). *Confusion matrix* dapat membantu dalam mengevaluasi akurasi model dalam melakukan prediksi, karena matrix ini memberikan representasi yang komprehensif terkait jumlah prediksi yang benar dan salah untuk setiap kelas (Muller & Guido, 2016). *Confusion matrix* dapat digambarkan sebagai sebuah tabel persegi yang mencatat jumlah *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) pada prediksi *classifier*, seperti yang terlihat dalam Gambar 2.3 berikut ini (Raschka & Mirjalili, 2017) :

		Predicted Class	
		Positive	Negative
Actual Class	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

Gambar 2. 3. Metriks Evaluasi Confusion Matrix

Confusion matrix memberikan banyak informasi, namun terkadang lebih mudah menggunakan matrix yang lebih ringkas (Géron, 2019). Ada beberapa cara untuk merangkum hasil dari *confusion matrix*, dengan yang paling umum adalah *accuracy*, *precision*, *recall*, dan *F1-Score* (Muller & Guido, 2016). *Accuracy* adalah

jumlah prediksi yang benar (TP dan TN) dibagi dengan jumlah semua sampel, yang dapat dinyatakan sebagai berikut :

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

Precision digunakan sebagai metrik kinerja dengan tujuan untuk membatasi jumlah positif palsu. Metrik ini menghitung berapa banyak sampel yang diprediksi positif ternyata positif, yang dapat dinyatakan sebagai berikut:

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

Recall digunakan untuk mengukur berapa banyak sampel positif yang ditangkap oleh prediksi positif. Recall digunakan ketika perlu mengidentifikasi semua sampel positif, yaitu ketika menghindari negatif palsu, yang dapat dijelaskan sebagai berikut :

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

Meskipun precision dan recall sangat penting sebagai tolak ukur evaluasi, hanya mempertimbangkan salah satu dari keduanya tidak akan memberikan gambaran keseluruhan. Sebagai alternatif, metrik F1-Score yang menggabungkan rata-rata harmonik dari precision dan recall, yang dapat dinyatakan sebagai berikut:

$$F1 - Score = 2 \times \frac{precision \times recall}{precision + recall} \quad (7)$$

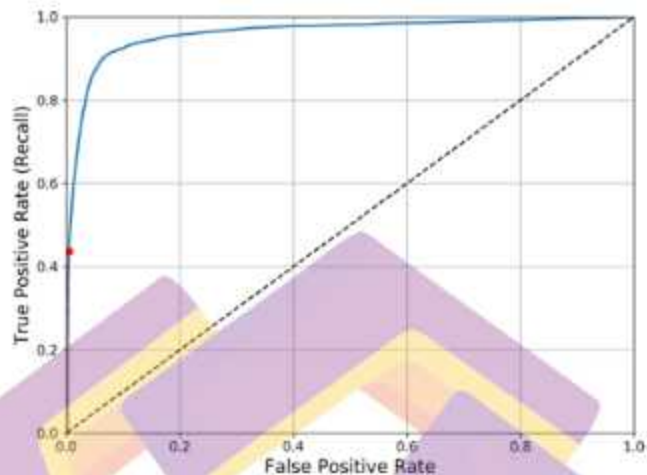
2.3.10 ROC-AUC

Receiver Operating Characteristic (ROC) Curve merupakan matrik evaluasi yang menggambarkan kinerja klasifikator (Hackeling, 2014). Secara garis besar, ROC Curve berbeda dengan akurasi, di mana matriks ini tidak dipengaruhi oleh proporsi kelas yang tidak seimbang dalam dataset, serta kurva ini lebih

menggambarkan kinerja pengklasifikasi untuk berbagai nilai ambang diskriminasi (Hackeling, 2014). ROC Curve menggambarkan recall dari pengklasifikasian terhadap tingkat kesalahan positifnya. Artinya kurva ini menggambarkan tingkat positif asli terhadap tingkat positif palsu (recall) (Géron, 2019; Hackeling, 2014). Tingkat positif palsu merupakan rasio dari *instances* negative yang diklasifikasikan sebagai positif, di mana tingkatan ini dihitung dengan satu dikurangi tingkat negative asli, yang merupakan rasio dari *instances* negative yang diklasifikasikan dengan benar (Géron, 2019). Kesalahan positif akan didefinisikan sebagai jumlah positif palsu dibagi dengan total jumlah negative. Perhitungan ini dilakukan dengan menerapkan rumus berikut :

$$F = \frac{FP \text{ (False Positive)}}{TN \text{ (True Negative)} + FP \text{ (False Positive)}} \quad (8)$$

AUC (Area Under ROC Curve) merupakan salah satu alat ukur yang mengukur luas area dibawah kurva (Géron, 2019). Pengklasifikasian yang sempurna akan memiliki nilai ROC-AUC sama dengan 1, sedangkan pengklasifikasian yang sepenuhnya acak akan memiliki nilai ROC-AUC sama dengan 0.5 (Géron, 2019; Hackeling, 2014). Aspek yang perlu diperhatikan dalam kurva ini adalah bahwa semakin tinggi tingkat True Positive (TP), semakin banyak pula False Positive (FP) yang dihasilkan oleh pengklasifikasi (Géron, 2019). ROC-AUC Curve dapat digambarkan (Gambar 2.4) sebagai garis putus-putus yang menunjukkan sebuah pengklasifikasi yang memprediksi sepenuhnya acak, sementara kurva yang padat atau berada jauh dari garis putus-putus menggambarkan pengklasifikasi yang menunjukkan kinerja lebih baik (menuju sudut kiri atas) (Géron, 2019; Hackeling, 2014).

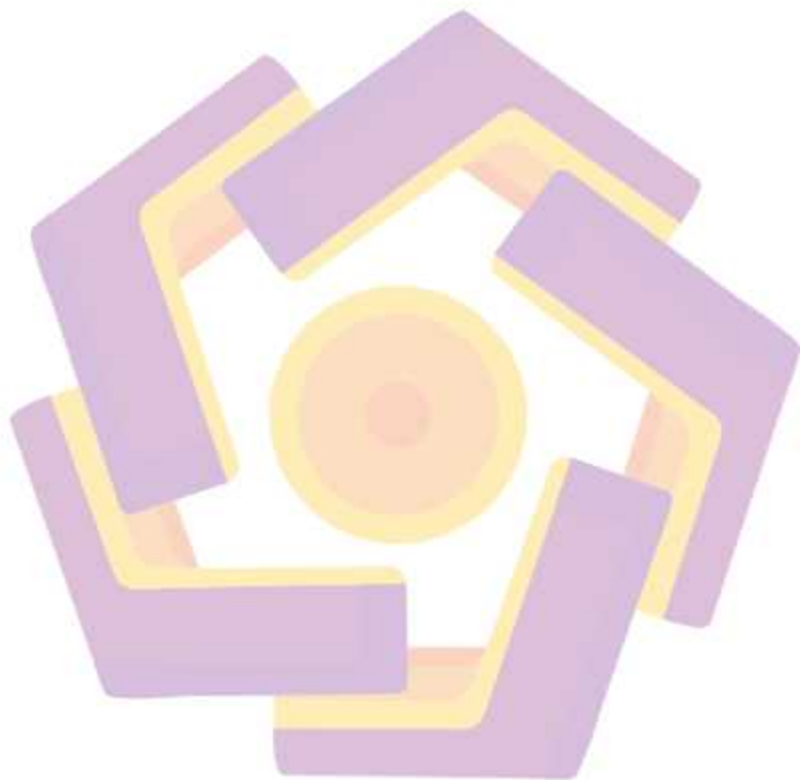


Gambar 2. 4. ROC Curve (Géron, 2019)

2.3.11 Cross Validation

Cross Validation merupakan salah satu metode yang dapat digunakan untuk mengatasi permasalahan *overfitting* yang seringkali muncul, dimana model *machine learning* terlalu menyesuaikan diri dengan setiap data dalam dataset, sehingga menyebabkan penyimpanan informasi menjadi tidak relevan dan timbul noise (Rajvanshi et al., 2024). Cross Validation merupakan metodologi dasar dalam *machine learning* yang membagi data menjadi k-blok untuk pengujian dan pelatihan (Rajvanshi et al., 2024). Pada metode ini data akan dipartisi menjadi k-blok dan menggunakan k-blok tersebut untuk evaluasi serta pengembangan pengetahuan. K-Folds Validation berarti membagi set pelatihan menjadi k-folds, kemudian melakukan prediksi dan evaluasi pada setiap folds menggunakan model yang dilatih pada setiap folds yang tersisa (Géron, 2019). Metode ini memastikan tidak ada tumpang tindih antara set pengujian selama proses pengambilan sampel,

di mana setiap folds merujuk pada jumlah subet yang dihasilkan (Rajvanshi et al., 2024).



BAB III

METODE PENELITIAN

3.1. Jenis, Sifat, dan Pendekatan Penelitian

Penelitian ini melibatkan eksperimen terhadap kumpulan data yang telah diperoleh dengan memanfaatkan algoritma Random Forest, serta menerapkan teknik SMOTE dalam mengembangkan model prediksi, guna mengidentifikasi fitur-fitur yang memiliki pengaruh terhadap kinerja akademik. Performa model prediksi yang dikembangkan akan di evaluasi menggunakan jumlah dataset yang serupa, guna mengidentifikasi metode yang paling akurat dalam memprediksi kinerja akademik siswa. Penelitian ini juga mengadopsi pendekatan kuantitatif, dimana analisis dilakukan secara sistematis, terencana, dan terstruktur dari awal hingga akhir penelitian. Penelitian ini juga bersifat deskriptif, karena secara umum menggambarkan objek yang diteliti dan menjelaskan hasil evaluasi pada dataset untuk menentukan model dengan kinerja yang optimal.

3.2. Metode Pengumpulan Data

Pengumpulan data dilakukan untuk menganalisis proses penilaian yang diterapkan oleh guru matematika di SMPN 2 Kuningan selama kegiatan pembelajaran berlangsung. Proses penelitian diawali dengan melakukan wawancara terhadap guru matematika kelas tujuh, guna mendapatkan pemahaman mengenai metode serta kriteria penilaian yang diterapkan dalam proses pembelajaran. Data yang digunakan berfokus pada 793 siswa-siswi Angkatan 2022 dan 2023. Data mencakup informasi demografi serta nilai akademik dalam mata

pelajaran matematika yang diperoleh dari DAPODIK (Data Pokok Pendidikan) melalui operator sekolah yang bertanggung jawab dalam pengelolaan data administratif dan akademik siswa. Pemilahan data dilakukan untuk memilih data yang sesuai dengan kebutuhan penelitian, berikut merupakan informasi terkait data yang dibutuhkan dari DAPODIK sekolah :

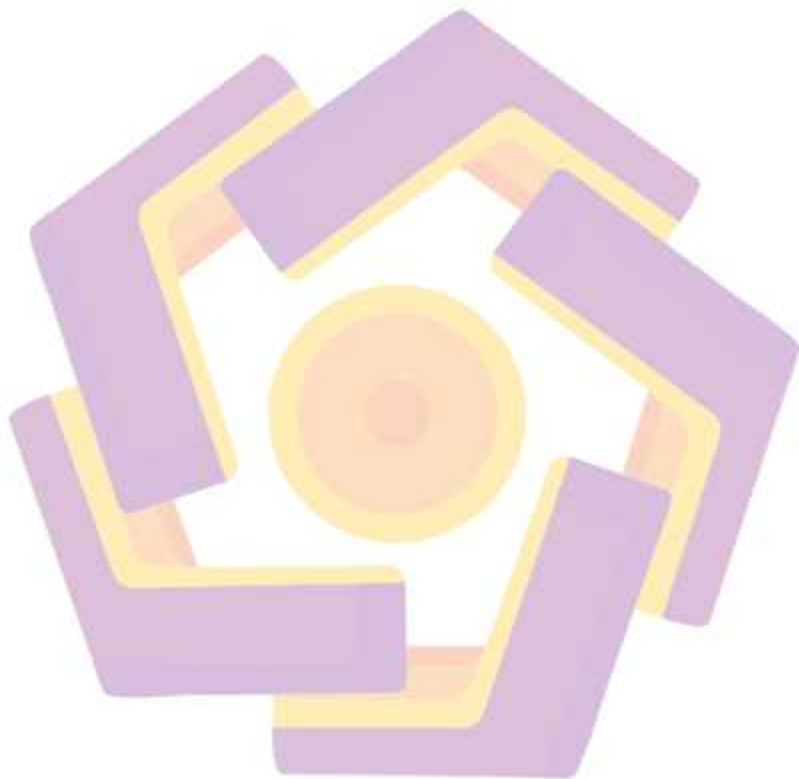
Tabel 3. 1. Informasi Data Demografi dan Akademik Siswa

Kategori	Fitur	Type
Demografi	Nama	String
	Jenis Kelamin	Kategorikal (Biner)
	Jenis Tinggal	Kategorikal (Multi)
	Pendidikan Ayah	Kategorikal (Multi)
	Pekerjaan Ayah	Kategorikal (Multi)
	Penghasilan Ayah	Kategorikal (Multi)
	Pendidikan Ibu	Kategorikal (Multi)
	Pekerjaan Ibu	Kategorikal (Multi)
	Penghasilan Ibu	Kategorikal (Multi)
	Anak Ke	Numerik
	Jumlah Saudara Kandung	Numerik
	Status Pernikahan Orang Tua	Kategorikal (Multi)
Akademik	Nilai Harian 1	Numerik
	Nilai Harian 2	Numerik
	Nilai Harian 3	Numerik
	Nilai PAS	Numerik
	Nilai Akhir	Numerik

Pengumpulan informasi kepribadian siswa dilakukan dengan menyebarkan kuesioner melalui *google form*. Kuesioner yang digunakan dalam penelitian ini mengacu pada instrumen *Big Five Inventory (BFI)* yang dikembangkan oleh (John et al., 1991; John et al., 2008), terdiri dari 44 item yang digunakan untuk mengukur

lima dimensi kepribadian pada model Big Five (*Neuroticism, Extraversion, Agreeableness, Conscientiousness, dan Openness*). Penggunaan kuesioner dilakukan karena instrument dengan 44 item telah berulang kali divalidasi oleh komunitas peneliti (Rico-Juan et al., 2024). Meskipun begitu, uji validitas dan reliabilitas akan tetap dilakukan dengan mengimplementasikan *python*. Kevalidan instrumen kepribadian Big Five yang dilakukan melalui uji validitas dan reliabilitas akan melibatkan beberapa metode (Budiastuti & Bandur, 2018; Garson, 2013). Proses validasi isi dilakukan sebelum instrumen disebarakan dengan melibatkan dua psikolog pendidikan yang memiliki keahlian dalam menganalisis karakteristik siswa/i, yaitu Ibu Vivi Ade Cerliana, S.Psi., M.Psi., Psikolog dan Ibu Hikmatul Fatiyah, S.Psi., M.Psi., Psikolog., dengan tujuan untuk memastikan instrumen yang digunakan dan disebarakan memiliki relevansi dan cakupan teori Big Five. Validasi konstruk menggunakan metode analisis korelasi dilakukan setelah proses penyebaran selesai, untuk menilai sejauh mana data dapat mengukur dimensi Big Five secara efektif. Selanjutnya, pengujian reliabilitas juga akan dilakukan untuk memastikan konsistensi pengukuran setiap dimensi, menggunakan metode *Cronbach's Alpha (CA)*. Metode korelasi dan CA akan diimplepentasikan menggunakan *python*, dengan kriteria valid, nilai korelasi (r) harus di atas 0,3 serta nilai signifikasi atau p_value harus dibawah 0,05 dan CA harus di atas 0,6 (Budiastuti & Bandur, 2018; Garson, 2013). Kemudian, perhitungan nilai rata-rata (John et al., 1991; John et al., 2008) untuk setiap dimensi dalam domain kepribadian Big Five akan dilakukan setelah validitas data kuesioner terkonfirmasi. Penggunaan BFI sendiri diharapkan dapat memberikan pemahaman yang komprehensif

mengenai karakteristik kepribadian siswa yang berkaitan dengan penelitian ini. Berikut ini merupakan kuisioner yang digunakan untuk menilai karakteristik siswa berdasarkan model kepribadian Big Five, dengan pengukuran dilakukan pada skala 1 hingga 5 :



Tabel 3. 2. Kuesioner Kepribadian Model Big Five berdasarkan Big Five Inventory (BFI)

Item	Pernyataan (Saya adalah seseorang yang ...)	Skala				
		1	2	3	4	5
		Sangat Tidak Sesuai	Tidak Sesuai	Netral	Sesuai	Sangat Sesuai
Extraversion						
E1	banyak bicara					
E2-R	pendiam					
E3	penuh energi					
E4-R	penuh semangat					
E5	cenderung pendiam					
E6	memiliki kepribadian yang tegas					
E7-R	kadang-kadang pemalu dan kurang terbuka					
E8	suka bersosialisasi					
Agreeableness						
A1-R	cenderung mencari kesalahan orang lain					
A2	suka membantu dan tidak egois terhadap orang lain					
A3-R	suka bertengkar dengan orang lain					
A4	pemaaf					
A5	sering kali dapat dipercaya					
A6-R	dapat bersikap dingin dan acuh tak acuh					
A7	pengertian dan baik kepada hampir semua orang					
A8-R	kadang kasar kepada orang lain					
A9	suka bekerja sama dengan orang lain					
Conscientiousness						
C1	bekerja dengan teliti					
C2-R	kadang ceroboh					

C3	pekerja yang dapat diandalkan					
C4-R	cenderung tidak teratur					
C5-R	cenderung malas					
C6	gigih sampai tugas selesai					
C7	melakukan segala sesuatu dengan efisien					
C8	menyusun rencana dan melaksanakannya					
C9-R	mudah kehilangan fokus					
Neuroticism						
N1	murung					
N2-R	menangani stress dengan baik					
N3	mudah merasa tegang					
N4	sering khawatir					
N5-R	stabil secara emosional, tidak mudah tersinggung					
N6	terkadang suasana hatinya tidak stabil					
N7-R	tetap tenang dalam situasi tegang					
N8	mudah gugup					
Openness						
O1	mehasilkan ide-ide baru					
O2	penasaran dengan berbagai hal					
O3	pemikir yang mendalam					
O4	memiliki imajinasi yang aktif					
O5	pemih ide					
O6	memiliki apresiasi terhadap seni dan keindahan					
O7-R	yang lebih suka pekerjaan yang rutin					
O8	suka merenung untuk mengeksplorasi ide-ide					
O9-R	kurang tertarik pada seni					
O10	berwawasan luas dalam seni, musik, atau sastra					

Data demografi, akademik, dan kepribadian (nilai rata-rata domain) akan diintegrasikan menjadi file dengan format CSV untuk memungkinkan dilakukannya pemodelan prediksi kinerja akademik menggunakan *python*. Data yang telah digabungkan akan diklasifikasikan menjadi tiga kategori, yaitu cukup (C), baik (B), dan sangat baik (SB). Pengkategorian ini didasarkan pada kriteria yang ditetapkan oleh guru matematika selaku pendidik, di mana penilaian untuk kriteria nilai harian 1, 2, dan 3 sendiri sudah mencakup aspek keaktifan siswa di dalam kelas, kehadiran, kedisiplinan, serta ketekunan dalam mengumpulkan tugas. Nilai akhir dihitung dengan memperhatikan nilai harian (1,2, dan 3) dan nilai PAS (ujian akhir). Berdasarkan hasil wawancara dengan tenaga pendidik yang mengajar mata pelajaran Matematika, standar penilaian di sekolah yang diteliti dilakukan secara kuantitatif (sumatif). Secara umum, para guru mengelompokkan siswa ke dalam tiga kategori berdasarkan rentang nilai tertentu. Nilai akhir yang diperoleh, membuat siswa-siswi akan diberi label cukup (75-82), baik (83-90), dan sangat baik (91-100). Hasil wawancara dengan guru matematika juga menunjukkan bahwa standar nilai 75-100 ditetapkan melalui musyawarah guru mata pelajaran dengan mempertimbangkan tiga aspek utama, yaitu karakteristik peserta didik, kompleksitas mata pelajaran, dan kondisi sekolah. Sehingga, berdasarkan ketiga aspek yang dimiliki oleh sekolah yang diteliti, nilai 75 ditetapkan menjadi nilai terendah yang bisa diperoleh oleh siswa. Ini juga sejalan dengan Kurikulum Merdeka yang telah mengganti Kriteria Ketuntasan Minimal (KKM) menjadi Kriteria Ketercapaian Tujuan Pembelajaran (KKTP). KKTP sendiri menekankan penilaian deskriptif berbasis proses belajar untuk mendiagnosis penguasaan

komptensi siswa, sehingga fokus tidak hanya pada angka semata. Oleh karena itu, nilai akhir ini diperoleh dari berbagai pertimbangan beberapa nilai yang sudah diakumulasikan dalam nilai harian, termasuk nilai akademik, keaktifan di kelas, dan ketepatan dalam pengumpulan tugas. Proses pelabelan dan validasi data terkait kinerja akademik akan dilakukan melalui diskusi bersama guru matematika kelas 7, yaitu Ibu Yoyoh Sadiyah, S.Pd. dan Bapak Ghifar Nur Hamzah, S.Pd., yang secara langsung mengamati, berinteraksi, dan memberikan penilaian kepada siswa/i. Data yang berhasil dilabeli akan dianalisa lebih lanjut, sehingga dapat dipilih fitur-fitur yang memiliki pengaruh signifikan terhadap kinerja siswa, yang selanjutnya akan digunakan untuk pemodelan prediksi menggunakan metode yang diajukan.

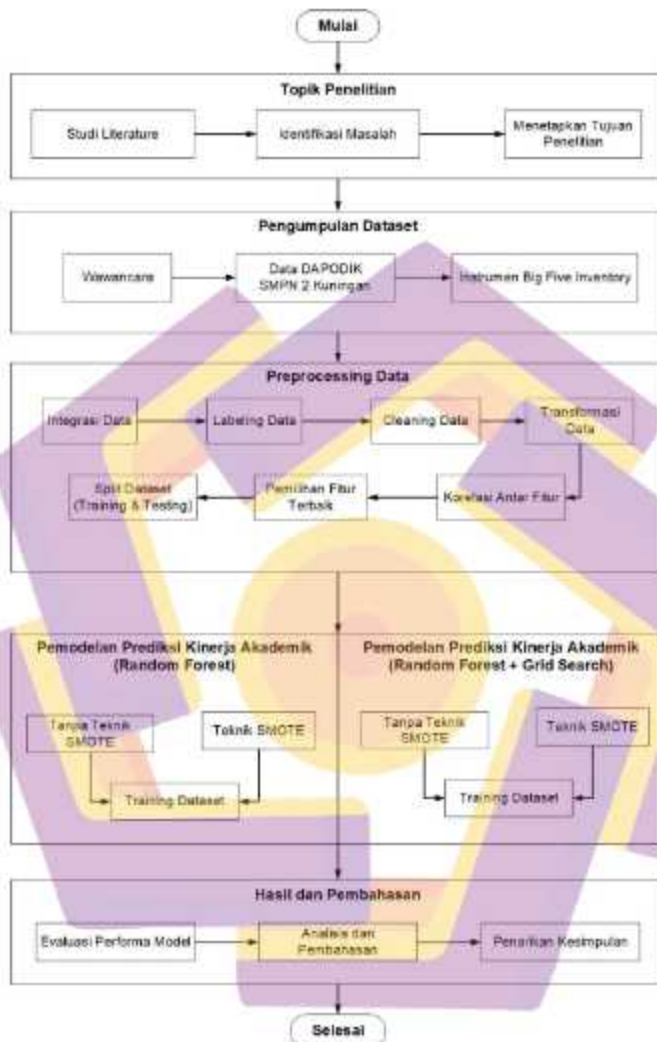
3.3. Metode Analisis Data

Metode analisis data yang diterapkan bertujuan untuk memprediksi kinerja akademik siswa SMP dalam pelajaran matematika, dengan fokus utama pada pemahaman siswa terhadap konsep yang diajarkan. Kinerja siswa dikategorikan menjadi tiga kelas, yaitu cukup (75-82), baik (83-91), dan sangat baik (91-100) pada saat kenaikan kelas. Proses pembangunan model prediksi akan dilakukan dengan mengimplementasikan python, serta dataset yang digunakan memiliki format data Comma-separated Values (CSV). Model prediksi ini akan dikembangkan menggunakan metode Random Forest, yang menggabungkan berbagai faktor seperti akademik, demografi, dan karakteristik kepribadian model Big Five. Analisis ini bertujuan untuk mengidentifikasi fitur-fitur yang memiliki pengaruh terhadap kinerja akademik, dengan mengeksplorasi keterkaitan antara kinerja

akademik dan aspek demografi serta antara kinerja akademik dan aspek kepribadian. Analisis ini juga bertujuan untuk membandingkan performa model Random Forest dalam berbagai skenario, termasuk model yang menggunakan pengaturan hyperparameter secara default dan model yang dioptimalkan melalui grid search. Selain itu, akan dilakukan perbandingan kinerja model Random Forest yang menerapkan teknik SMOTE dan yang tidak. Skenario-skenario ini akan dievaluasi untuk memperoleh model dengan tingkat keakuratan prediksi dan akurasi terbaik dalam memprediksi kinerja akademik siswa. Bagian ini memuat penjelasan secara lengkap dan terinci tentang metode dan alat yang digunakan untuk analisis data.



3.4. Alur Penelitian



Gambar 3. 1. Alur Penelitian

Penelitian ini terdiri dari lima tahapan utama sebagaimana tertera pada Gambar 3.1. Tahapan penelitian tersebut adalah sebagai berikut :

a. Topik Penelitian

Studi literatur ini bertujuan untuk menganalisis penelitian-penelitian sebelumnya mengenai topik *Educational Data Mining* (EDM) guna memahami perkembangan serta peluang yang dapat dieksplorasi lebih lanjut. Identifikasi masalah dilakukan untuk menetapkan topik penelitian serta kontribusi yang akan dilakukan. Penulis memutuskan untuk melakukan penelitian terkait prediksi kinerja akademik siswa di tingkat sekolah menengah pertama, suatu bidang yang masih minim diteliti. Penelitian ini juga akan menerapkan fitur-fitur yang relevan, seperti aspek akademik, demografi, dan kepribadian model Big Five, yang menurut beberapa penelitian memiliki hubungan signifikan dengan kinerja akademik siswa namun belum banyak diteliti lebih lanjut hubungannya. Penelitian ini juga akan berkontribusi dalam menangani tantangan ketidakseimbangan data dengan menerapkan teknik SMOTE pada model Random Forest, yang diyakini memiliki ketahanan terhadap *overfitting* dan telah terbukti efektif dalam memprediksi kinerja akademik. Penelitian ini juga akan berkontribusi dalam meningkatkan performa model Random Forest dengan menerapkan *hyperparameter tuning*, seperti *grid search*.

b. Pengumpulan Dataset

Pengumpulan data dilakukan secara berjenjang, seperti melakukan wawancara dengan guru matematika sekolah menengah pertama. Langkah ini bertujuan untuk memperoleh pemahaman menyeluruh tentang konteks dan dinamika kelas dalam lingkungan akademik. Sebanyak 793 sampel data siswa/siswi dari angkatan 2022 dan 2023 SMPN 2 Kuningan, berhasil dikumpulkan dari

DAPODIK melalui operator yang bertugas. Untuk memastikan kualitas dan akurasi data yang diperoleh sesuai dengan tujuan penelitian, penilaian kepribadian menggunakan model *Big Five* dilakukan dengan menyebarkan kuesioner *Big Five Inventory* (BFI) secara bertahap kepada siswa/siswi, dengan memanfaatkan *Google Form*. Selain itu, uji validitas isi melalui para ahli (psikolog) serta uji validitas konstruk menggunakan analisis korelasi dan uji reliabilitas menggunakan *cronbach's alpha* juga dilakukan untuk memastikan kevalidan instrumen kepribadian *Big Five* yang digunakan

c. Preprocessing Data

Dalam penelitian ini, terdapat beberapa langkah yang harus dilakukan untuk mempersiapkan data. Langkah pertama ialah melakukan integrasi terhadap data kepribadian yang tervalidasi dan data siswa (demografi dan akademik) yang diperoleh dari DAPODIK agar menjadi satu dataset tunggal. Langkah selanjutnya adalah melakukan klasifikasi data ke dalam tiga kategori, yaitu cukup atau C (75-82), baik atau B (83-91) dan sangat baik atau SB (92-100). Proses klasifikasi atau *labeling* data dilakukan berdasarkan penilaian yang dilakukan langsung oleh guru matematika selaku seseorang yang mengamati dan berinteraksi saat proses pembelajaran berlangsung. Langkah berikutnya adalah melakukan pembersihan data dengan tujuan untuk menangani data yang tidak lengkap melalui berbagai proses, seperti menghapus atribut yang tidak relevan, mengisi data yang hilang, menghapus data yang tidak konsisten, serta mencari dan menghapus data duplikat (Gori et al., 2024). Selanjutnya, melakukan transformasi data pada beberapa atribut dalam dataset, di mana data yang bersifat kategorikal (biner dan multi) akan diubah

menjadi data numerik melalui metode *label encoding* (Gori et al., 2024; Santoso et al., 2024). Label encoding dilakukan karena model *machine learning* seperti Random Forest yang dirancang untuk melakukan perhitungan matematis dan statistic, hanya dapat memproses data dalam bentuk numerik. Pada dasarnya, label encoding merupakan teknik pengolahan data yang mengkonversi data kategorikal menjadi data numerik dengan cara memberikan label angka (0-9) untuk masing-masing kategori (Gori et al., 2024). Langkah selanjutnya ialah membangun matriks korelasi untuk melakukan analisis terhadap hubungan antar fitur yang ada pada dataset. Berdasarkan hasil analisis matriks korelasi terhadap fitur yang ada, di mana fitur-fitur yang menunjukkan nilai korelasi tertinggi akan dipilih untuk diterapkan dalam pemodelan prediksi yang akan dilakukan. Terakhir, data yang telah dibersihkan akan dibagi menjadi dua kategori, yaitu data latih dan data uji dengan proporsi 70:30 (Ghaniaviyanto Ramadhan & Adiwijaya, 2022). Data latih digunakan untuk melatih model dalam mengenali pola, sedangkan data uji digunakan untuk mengevaluasi kinerja model secara terpisah. Proses ini sangat penting untuk meningkatkan kinerja model dan memastikan kemampuannya dalam menggeneralisasi terhadap data yang belum pernah dilihat sebelumnya.

d. Pemodelan Random Forest

Model prediksi kinerja akademik siswa dalam penelitian ini akan menggunakan metode Random Forest dengan beberapa skenario. Skenario pertama akan menggunakan dataset asli yang tidak dimodifikasi, yang menunjukkan ketidakseimbangan kelas. Pada skenario kedua, akan diterapkan teknik SMOTE untuk menyeimbangkan dataset tersebut. Skenario ketiga dan keempat akan

melibatkan penerapan *hyperparameter tuning* melalui *grid search* pada model yang menggunakan teknik SMOTE dan yang tidak, dengan tujuan untuk meningkatkan akurasi serta performa model Random Forest. Parameter yang akan diterapkan dalam proses ini ialah *random_state*, *n_estimators*, *max_depth*, *max_features*, *min_samples_leaf* dan *min_samples_split*. Dalam penelitian ini, penerapan teknik SMOTE sendiri akan diterapkan dalam seluruh dataset sebelum data di split (Ghaniaviyanto Ramadhan & Adiwijaya, 2022; Hamoud et al., 2022). Dalam proses ini, data dari kelas minoritas akan ditambahkan untuk mendekati keseimbangan jumlah dengan kelas mayoritas. Sehingga, dengan menerapkan SMOTE pada seluruh dataset, distribusi kelas menjadi lebih seimbang serta model tidak cenderung bias terhadap kelas mayoritas. Skenario-skenario tersebut dilakukan dengan menggunakan dataset yang sama, serta performanya akan diuji secara terpisah, di mana hasil evaluasi akan dibandingkan untuk menentukan pendekatan yang paling efektif dalam menangani ketidakseimbangan kelas dan meningkatkan kinerja model prediksi.

e. Hasil dan Pembahasan

Pada tahap ini, masing-masing skenario akan dievaluasi performanya menggunakan berbagai metrik, seperti *confusion matrix* termasuk akurasi, presisi, *recall*, *f1-measure* serta *ROC-AUC*. Selain itu, pendekatan menggunakan *cross-validation* akan dilakukan pada data pelatihan sebelum pemodelan guna mencegah terjadinya *overfitting*, serta mengukur kestabilan data (Nachouki et al., 2023; Santoso et al., 2024). Dalam konteks ini, *k_fold* antara 2 hingga 10 diterapkan untuk membandingkan kinerja model dan mendapatkan hasil yang optimal. Cross

validation membagi data pelatihan menjadi beberapa fold, di mana setiap fold secara bergantian berfungsi sebagai data uji, sementara sisanya digunakan untuk proses pelatihan (Rico-Juan et al., 2024). Hasil akhir diperoleh dari rata-rata seluruh *fold* untuk memberikan estimasi akurasi yang lebih stabil. Sementara itu, nilai $k_fold=1$ tidak dapat diterapkan karena *cross-validation* membutuhkan minimal dua bagian/*fold* agar data dapat dibagi menjadi data latih dan data uji, jika nilainya 1 maka seluruh data hanya bisa dijadikan satu bagian, sehingga proses validasi tidak dapat dilakukan (Rico-Juan et al., 2024). Hasil evaluasi dari setiap model akan dianalisis dan dibandingkan untuk mengidentifikasi model yang memiliki efektivitas paling baik dalam memprediksi kinerja siswa. Penelitian ini juga akan menganalisis terkait hubungan antara fitur-fitur yang digunakan dalam prediksi kinerja siswa, sehingga dapat diidentifikasi fitur-fitur yang memiliki korelasi yang signifikan dengan kinerja akademik siswa. Temuan dari analisis ini akan digunakan untuk menarik kesimpulan mengenai efektivitas dan keandalan model dalam konteks prediksi kinerja akademik siswa.

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

4.1. Pembuatan Dataset

Secara umum, data yang digunakan dalam penelitian ini diperoleh dari basis data sebuah lembaga pendidikan, serta melalui penyebaran kuesioner. Selain itu, sebelum dan setelah penyebaran, instrument kuesioner akan menjalani serangkaian pengujian guna memastikan validitasnya, sehingga dapat mengukur kepribadian secara akurat sesuai dengan kapasitasnya.

4.1.1 Pengumpulan Data

Dalam penelitian ini, dataset private digunakan karena dianggap lebih relevan mengingat sifatnya yang personel dan kontekstual, berbeda dengan dataset publik yang cenderung bersifat umum. Namun demikian, perlu disoroti bahwa kondisi dan karakteristik khas yang terdapat pada data yang digunakan tidak dapat digeneralisasikan secara umum. Oleh karena itu, dalam melakukan perbandingan, diperlukan ketelitian dalam menganalisis setiap aspek yang digunakan dalam penelitian ini. Data diperoleh melalui DAPODIK (Data Pokok Pendidikan) sekolah serta melalui penyebaran kuesioner. Secara keseluruhan, sebanyak 793 data akademik dan demografis berhasil dikumpulkan dari DAPODIK sekolah dengan atribut yang relevan. Namun, setelah melakukan penyebaran kuesioner terkait kepribadian model Big Five, dari total 793 siswa angkatan 2022 dan 2023, hanya 763 sampel yang berhasil dikumpulkan. Hasil dari pengumpulan data ini, diketahui

bahwa dataset terdiri dari beberapa atribut dengan nilai yang bervariasi sebagai berikut :

Tabel 4. 1. Informasi Atribut dan Value Data Akademik, Demografis, dan Kepribadian Big Five

Kategori	Fitur	Value
Demografis	Nama	Text
	Jenis Kelamin	Percempuan (P)
		Laki-laki (L)
	Jenis Tinggal	Bersama orang tua
		Wali
		Asrama Lainnya
	Pendidikan Ayah Pendidikan Ibu	Tidak Sekolah
		SD/Sederajat
		SMP/Sederajat
		SMA/Sederajat
		Diploma
		S1
		S2
	S3	
	Pekerjaan Ayah Pekerjaan Ibu	Putus Sekolah
		Tidak Bekerja
		Sudah Meninggal
		Karyawan Swasta
		PNS/TNI/Polri
		Wiraswasta/Wirausaha
Pedagang Lainnya		
Penghasilan Ayah Penghasilan Ibu	Tidak Berpenghasilan	
	Kurang dari Rp.500000	
	Rp.500000 s.d Rp.999999	
	Rp.1000000 s.d Rp.1999999	
	Rp.2000000 s.d Rp.4999999 Rp.5000000 s.d Rp. 20000000	
Anak Ke	0-9	
Jumlah Saudara Kandung	0-9	
Status Pernikahan Orang Tua	Menikah	
	Orang tua tidak tinggal bersama lagi (bercerai)	
	Salah satu orang tua sudah tiada (meninggal)	
Akademik	Nilai Harian 1	75-100
	Nilai Harian 2	75-100
	Nilai Harian 3	75-100
	Nilai PAS	75-100

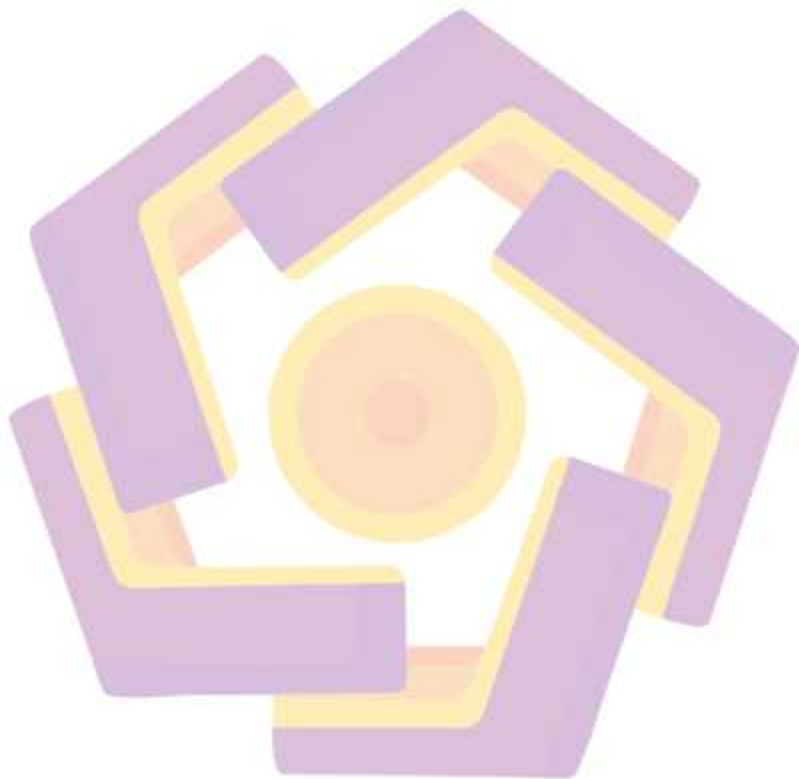
	Nilai Akhir (Dependent)	Cukup (C)
		Baik (B)
		Baik Sekali (SB)
Kepribadian Big Five	Extraversion	1-5
	Agreeableness	1-5
	Conscientiousness	1-5
	Neuroticism	1-5
	Openness	1-5

4.1.2 Uji Validitas Isi Kuesioner

Uji validitas isi instrument kepribadian ini dilakukan oleh seorang psikolog yang merupakan pakar dalam penilaian kepribadian. Proses validasi dilakukan melalui pengisian angket validasi kepribadian Big Five oleh dua psikolog yang memiliki fokus pada kepribadian anak-anak, khususnya di dunia pendidikan, serta angket validasi ini dapat ditemukan pada halaman lampiran. Angket ini bertujuan untuk menilai sejauh mana setiap komponen isi pernyataan dengan masing-masing domain kepribadian mampu mencerminkan maksud dan tujuan yang diinginkan. Penilaian mencakup kesesuaian isi pernyataan dengan masing-masing domain, ketepatan penggunaan tata bahasa, serta konstruksi dari setiap bagian pernyataan.

Hasil dari uji validitas isi ini menunjukkan bahwa sejumlah pernyataan dinilai cenderung membuat responden takut untuk bersikap transparan terhadap apa yang mereka alami atau rasakan, ini akibat pernyataan yang terlalu singkat. Selain itu, terdapat pula pernyataan yang mengandung istilah dengan makna ganda, bahkan terskesan tidak jelas atau ambigu, sehingga makna dan tujuan dari kalimat tersebut tidak dapat dipahami dengan baik. Sehingga, para ahli menekankan pentingnya perbaikan dalam penyusunan pernyataan agar lebih komunikatif bagi siswa, sehingga responden dapat mengisi kuesioner dengan lebih efektif. Secara keseluruhan, kedua psikolog menyimpulkan bahwa instrument yang diajukan dapat

digunakan, namun perlu dilakukan revisi pada beberapa aspek. Berdasarkan hasil penilaian yang diberikan oleh kedua psikolog, maka perbaikan dilakukan guna menghasilkan instrument yang lebih baik, di mana hasil perbaikan ini telah disajikan dalam tabel 4.2 .



Tabel 4. 2. Perbaikan Instrumen Kepribadian Big Five

Item	Pernyataan (Saya adalah seseorang yang ...)	Pernyataan Revisi (Saya adalah seseorang yang ...)
Extraversion		
E1	banyak bicara	gemar berbicara dan tidak mudah kehabisan bahan pembicaraan
E2-R	pendiam	tertutup atau tidak terlalu ekspresif
E3	penuh energi	memiliki banyak energi
E4-R	penuh semangat	memiliki semangat tinggi
E5	cenderung pendiam	cenderung pendiam
E6	memiliki kepribadian yang tegas	memiliki kepribadian yang tegas dan percaya diri
E7-R	kadang-kadang pemalu dan kurang terbuka	terkadang pemalu dan canggung saat berinteraksi dengan orang lain
E8	suka bersosialisasi	suka bergaul atau bersosialisasi dengan orang lain
Agreeableness		
A1-R	cenderung mencari kesalahan orang lain	cenderung mencari kesalahan-kesalahan orang lain
A2	suka membantu dan tidak egois terhadap orang lain	suka membantu dan tidak hanya mementingkan diri sendiri
A3-R	suka bertengkar dengan orang lain	suka memulai perdebatan dengan orang lain
A4	pemaaf	pemaaf
A5	sering kali dapat dipercaya	dapat dipercaya
A6-R	dapat bersikap dingin dan acuh tak acuh	bisa bersikap dingin dan cuek
A7	pengertian dan baik kepada hampir semua orang	suka memberi perhatian dan baik kepada hampir semua orang
A8-R	kadang kasar kepada orang lain	terkadang bersikap kasar atau menyebalkan kepada orang lain
A9	suka bekerja sama dengan orang lain	suka bekerja sama dengan orang lain
Conscientiousness		
C1	bekerja dengan teliti	melakukan tugas dengan teliti
C2-R	kadang ceroboh	terkadang bertindak ceroboh
C3	pekerja yang dapat diandalkan	bisa diandalkan dalam pekerjaan atau pengerjaan tugas
C4-R	cenderung tidak teratur	cenderung tidak teratur atau kesulitan mengatur segala sesuatunya dengan baik

C5-R	cenderung malas	cenderung bersikap malas
C6	gigih sampai tugas selesai	gigih mengerjakan tugas sampai selesai
C7	melakukan segala sesuatu dengan efisien	melakukan segala sesuatu dengan efisien
C8	menyusun rencana dan melaksanakannya	suka membuat rencana-rencana dan kemudian melaksanakannya
C9-R	mudah kehilangan fokus	mudah kehilangan fokus atau mudah terganggu
Neuroticism		
N1	murung	mudah tertekan atau murung
N2-R	menangani stress dengan baik	cenderung santai dan bisa menangani stress dengan baik
N3	mudah merasa tegang	mudah merasa tegang atau cemas
N4	sering khawatir	mudah merasa khawatir
N5-R	stabil secara emosional, tidak mudah tersinggung	tidak mudah tersinggung atau tidak mudah marah
N6	terkadang suasana hatinya tidak stabil	suasana hati atau mood nya mudah berubah
N7-R	tetap tenang dalam situasi tegang	bisa tetap tenang dalam situasi-situasi mengganggan
N8	mudah gugup	mudah gugup
Openness		
O1	mehasilkan ide-ide baru	dapat mehasikan ide-ide baru
O2	penasaran dengan berbagai hal	suka mencari tahu tentang berbagai hal atau memiliki rasa penasaran yang tinggi
O3	pemikir yang mendalam	memikirkan sesuatu secara teliti dan mendalam atau penuh pertimbangan
O4	memiliki imajinasi yang aktif	suka berimajinasi
O5	penuh ide	kreatif atau suka berinovasi
O6	memiliki apresiasi terhadap seni dan keindahan	suka menikmati karya seni dan keindahan
O7-R	yang lebih suka pekerjaan yang rutin	lebih suka melakukan kegiatan yang sama setiap hari atau teratur daripada mencoba hal baru
O8	suka merenung untuk mengeksplorasi ide-ide	suka merenung dan memikirkan ide-ide
O9-R	kurang tertarik pada seni	kurang tertarik pada seni
O10	berwawasan luas dalam seni, musik, atau sastra	memiliki keahlian dalam seni, musik, ataupun sastra

4.1.3 Uji Validitas Konstruk Kuesioner

Uji validitas konstruk dilakukan dengan menggunakan metode analisis korelasi setelah kuesioner disebarikan. Tujuan dari proses validasi ini adalah untuk mengevaluasi sejauh mana data dapat secara efektif mengukur dimensi big five. Namun, sebelum melakukan uji validitas konstruk, penting untuk menyoroti bahwa sejumlah pernyataan atau item perlu diubah nilainya atau dilakukan invers (John et al., 1991; John et al., 2008). Ini berarti beberapa pernyataan memiliki nilai yang bertentangan dengan domain yang dimaksud, sehingga harus dilakukan invers. Dalam konteks ini, item-item yang perlu mengalami pembalikan nilai (invers) sebelum dilakukan analisis lebih lanjut meliputi E2-R, E4-R, dan E7-R dalam domain extraversion, A1-R, A3-R, A6-R, dan A8-R dalam domain agreeableness, C2-R, C4-R, C5-R, dan C9-R dalam domain conscientiousness, N2-R, N5-R, dan N7-R dalam domain neuroticism, O7-R, dan O9-R dalam domain openness.

Analisis statistik deskriptif dilakukan terlebih dahulu untuk mengetahui penyebaran data responden. Berdasarkan hasil analisis statistik deskriptif (Tabel 4.3) dan data penyebaran jawaban responden (Tabel 4.4) mengenai lima dimensi kepribadian big five, menunjukkan bahwa setiap dimensi memperoleh nilai antara 1 hingga 5, yang mencerminkan distribusi yang merata pada skala yang telah ditetapkan. Variasi yang jelas dalam respon responden terlihat dari nilai standar deviasi yang berada dalam rentang 0,8 hingga 1. Selain itu, kesamaan antara nilai modus dan median pada dimensi neuroticism, extraversion, conscientiousness, dan openness menunjukkan bahwa sebagian besar responden memiliki distribusi yang cukup merata di sekitar nilai tengah atau netral. Sementara itu, nilai median 4 pada

dimensi agreeableness menunjukkan bahwa mayoritas peserta cenderung memiliki sifat kooperatif dan kemampuan sosial yang cukup baik. Secara keseluruhan, responden umumnya memperoleh skor pada tingkat sedang, dengan sedikit kecenderungan menuju skor yang lebih tinggi atau lebih rendah.

Tabel 4. 3. Statistik Deskriptif Data Kuesioner Kepribadian Big Five

Dimensi	Mean	Standar Deviasi	Min	Max	Median	Modus
Neuroticism	3,206	1,002	1	5	3	3
Extraversion	3,172	0,916	1	5	3	3
Agreeableness	3,554	0,867	1	5	4	3
Conscientiousness	3,152	0,862	1	5	3	3
Openness	3,337	0,909	1	5	3	3

Tabel 4. 4. Frekuensi dan Presentase Jawaban Kuesioner Kepribadian Big Five (Skala Likert 1-5)

Item	Nilai									
	1		2		3		4		5	
	Frekuensi	Presentase (%)	Frekuensi	Presentase (%)	Frekuensi	Presentase (%)	Frekuensi	Presentase (%)	Frekuensi	Presentase (%)
E1	37	4,8	394	51,6	219	28,7	90	11,8	23	3,0
E2-R	31	4,1	140	18,3	367	48,1	166	21,8	59	7,7
E3	11	1,4	77	10,1	350	45,9	215	28,2	110	14,4
E4-R	6	0,8	48	6,3	319	41,8	264	34,6	126	16,5
E5	34	4,5	105	13,8	307	40,2	194	25,4	123	16,1
E6	14	1,8	101	13,2	386	50,6	205	26,9	57	7,5
E7-R	129	16,9	265	34,7	266	34,9	79	10,4	24	3,1
E8	8	1,0	48	6,3	325	42,6	255	33,4	127	16,6
A1-R	2	0,3	29	3,8	134	17,6	270	35,4	328	43,0
A2	4	0,5	24	3,1	268	35,1	339	44,4	128	16,8
A3-R	12	1,6	40	5,2	171	22,4	265	34,7	275	36,0
A4	10	1,3	15	2,0	228	29,9	266	34,9	244	32,0
A5	6	0,8	15	2,0	374	48,5	283	37,1	112	14,7
A6-R	33	4,3	434	56,9	240	31,5	47	6,2	9	1,2
A7	11	1,4	50	6,6	391	51,2	218	28,6	93	12,2
A8-R	54	7,1	145	19,0	270	35,4	220	28,8	74	9,7
A9	15	2,0	49	6,4	331	43,4	272	35,6	96	12,6
C1	5	0,7	45	5,9	406	53,2	228	29,9	79	10,4
C2-R	61	8,0	274	35,9	329	43,1	91	11,9	8	1,0
C3	10	1,3	73	9,6	449	58,8	175	22,9	56	7,3
C4-R	25	3,3	132	17,3	350	45,9	204	26,7	52	6,8
C5-R	30	3,9	125	16,4	358	46,9	172	22,5	78	10,2
C6	7	0,9	41	5,4	382	50,1	238	31,2	95	12,5
C7	8	1,0	65	8,5	444	58,2	180	23,6	66	8,7
C8	14	1,8	91	11,9	343	45,0	244	32,0	71	9,3
C9-R	114	14,9	271	35,5	236	30,9	117	15,3	25	3,3

N1	72	9,4	206	27,0	266	34,9	145	19,0	74	9,7
N2-R	82	10,7	248	32,5	333	43,6	82	10,7	18	2,4
N3	30	3,9	101	13,2	239	31,3	237	31,1	156	20,4
N4	15	2,0	39	5,1	232	30,4	316	41,4	161	21,1
N5-R	65	8,5	153	20,1	381	49,9	102	13,4	62	8,1
N6	24	3,1	83	10,9	276	36,2	173	22,7	207	27,1
N7-R	62	8,1	144	18,9	375	49,1	146	19,1	36	4,7
N8	22	2,9	103	13,5	301	39,4	225	29,5	112	14,7
O1	5	0,7	86	11,3	449	58,8	174	22,8	49	6,4
O2	6	0,8	43	5,6	248	32,5	292	38,3	174	22,8
O3	3	0,4	55	7,2	381	49,9	223	29,2	101	13,2
O4	16	2,1	61	8,0	307	40,2	224	29,4	155	20,3
O5	7	0,9	91	11,9	449	58,8	264	34,5	52	6,8
O6	3	0,4	42	5,5	323	42,3	231	30,3	164	21,5
O7-R	25	3,3	335	43,9	189	24,8	118	15,5	96	12,6
O8	15	2,0	102	13,4	392	51,4	181	23,7	73	9,6
O9-R	20	2,6	93	12,2	295	38,7	236	30,9	119	15,6
O10	66	8,7	219	28,7	319	41,8	98	12,8	61	8,0

Hasil uji validitas konstruk pada instrument kepribadian big five, ini menunjukkan bahwa semua item dalam instrument ini memiliki korelasi yang signifikan dengan total skor dari dimensi yang diukur. Seluruh nilai korelasi (r) berada di atas 0,3 (Budiastuti & Bandur, 2018; Garson, 2013), yang menunjukkan bahwa setiap item memiliki hubungan yang cukup kuat dengan konstruksinya. Hal ini dijasikan dalam Tabel 4.5, di mana pada dimensi extraversion (E) yang terdiri dari 8 item, nilai korelasi yang diperoleh berkisar antara 0,55 hingga 0,67. Dimensi agreeableness (A) yang terdiri dari 9 item, menunjukkan nilai korelasi antara 0,52 hingga 0,57. Dimensi conscientiousness (C) yang juga terdiri dari 9 item, memiliki nilai korelasi antara 0,56 hingga 0,66. Dimensi neuroticism (N) yang terdiri dari 9 item, memiliki nilai korelasi antara 0,49 hingga 0,75. Terakhir, dimensi openness (O) yang terdiri dari 10 item, menunjukkan nilai korelasi antara 0,52 hingga 0,70. Temuan ini menunjukkan bahwa setiap item dapat merepresentasikan aspek kepribadian dengan cukup akurat.

Tabel 4. 5. Uji Validasi Konstruk Domain Kepribadian Big Five (Extraversion, Agreeableness, Conscientiousness, Neuroticism, dan Openess)

Item	Nilai Korelasi (r)	p_value
E1	0,61	0,00
E2-R	0,58	0,00
E3	0,64	0,00
E4-R	0,57	0,00
E5	0,63	0,00
E6	0,55	0,00
E7-R	0,57	0,00
E8	0,67	0,00
A1-R	0,56	0,00
A2	0,54	0,00
A3-R	0,54	0,00
A4	0,57	0,00

A5	0,53	0,00
A6-R	0,52	0,00
A7	0,53	0,00
A8-R	0,52	0,00
A9	0,52	0,00
C1	0,63	0,00
C2-R	0,61	0,00
C3	0,56	0,00
C4-R	0,59	0,00
C5-R	0,62	0,00
C6	0,66	0,00
C7	0,62	0,00
C8	0,57	0,00
C9-R	0,56	0,00
N1	0,69	0,00
N2-R	0,42	0,00
N3	0,75	0,00
N4	0,68	0,00
N5-R	0,49	0,00
N6	0,59	0,00
N7-R	0,54	0,00
N8	0,64	0,00
O1	0,64	0,00
O2	0,52	0,00
O3	0,54	0,00
O4	0,60	0,00
O5	0,70	0,00
O6	0,70	0,00
O7-R	0,54	0,00
O8	0,62	0,00
O9-R	0,53	0,00
O10	0,60	0,00

Secara keseluruhan, item dengan nilai korelasi tertinggi adalah N3 (0,75), yang menunjukkan bahwa pernyataan tersebut memiliki tingkat keterkaitan yang sangat kuat dengan total skornya. Sebaliknya, item dengan nilai korelasi terendah ialah N2-R (0,42). Selain itu, semua item memiliki nilai $p_value = 0,00$, ini menunjukkan bahwa hubungan antara setiap item dan total skor sangat signifikan ($p < 0,05$). Oleh karena itu, berdasarkan uji signifikansi menunjukkan tidak ada item yang perlu dihapus. Sehingga, ini menunjukkan bahwa semua item dalam setiap dimensi kepribadian big five dianggap valid, di mana semua pernyataan dalam setiap dimensi telah teruji dan terbukti signifikan.

4.1.4 Uji Reliabilitas

Pengujian reliabilitas dalam penelitian ini dilakukan menggunakan metode Cronbach's Alpha, yang disajikan dalam Tabel 4.6. Temuan menunjukkan bahwa nilai Cronbach's Alpha yang diperoleh oleh keseluruhan dimensi mencapai nilai 0,81. Pada tingkat domain, nilai Cronbach's Alpha untuk setiap dimensi menunjukkan variasi dalam reliabilitasnya. Dimensi Neuroticism (0,75), Extraversion (0,75), Conscientiousness (0,78), dan Openness (0,79) menunjukkan tingkat reliabilitas yang baik, yang mengindikasikan bahwa item-item dalam domain tersebut cukup konsisten dalam mengukur karakteristik yang dimaksud. Sementara itu, domain Agreeableness (0,69) menunjukkan tingkat reliabilitas yang lebih rendah dibandingkan dengan domain lainnya, meskipun ini masih menunjukkan bahwa domain tergolong dalam kategori yang cukup baik.

Tabel 4. 6. Uji Realibilitas Cronbach's Alpha

Domain	Cronbach's Alpha
Neuroticism	0,75
Extraversion	0,75
Agreeableness	0,69
Conscientiousness	0,78
Openness	0,79
Keseluruhan	0,81

Secara keseluruhan, hasil analisis reliabilitas ini menunjukkan bahwa instrument yang digunakan untuk mengukur kepribadian *big five* memiliki konsistensi internal yang cukup baik. Hal ini didasari pada rentang nilai *cronbach's alpha* yang disandarkan pada *judgement threshold* yang sudah banyak disepakati para ahli seperti yang ditunjukkan dalam Tabel 4.7 (Budiastuti & Bandur, 2018). Secara garis besar, nilai *cronbach's alpha* 0,6 dianggap dapat

diterima untuk tujuan eksplorasi, nilai 0,7 dianggap memadai untuk konfirmasi, dan 0,8 dianggap baik untuk tujuan konfirmasi (Garson, 2013).

Tabel 4. 7. Klasifikasi Kelas Nilai Cronbach's Alpha

1	Reliabilitas Sempurna (<i>Perfect Reliability</i>)
0,9	Reliabilitas Sangat Baik (<i>Excellent Reliability</i>)
> 0,8	Reliabilitas Baik (<i>Good Reliability</i>)
> 0,7	Reliabilitas Dapat Diterima (<i>Acceptable Reliability</i>)
0	Tidak memiliki reliabilitas (<i>No Reliability</i>)

Sehingga, berdasarkan hasil perhitungan *cronbach's alpha* dalam penelitian ini, instrument yang digunakan masuk dalam kelas *acceptable reliability* dalam mengukur dimensi kepribadian big five, karena rata-rata nilai reliabilitas berada di angka 0,7. Sehingga, temuan ini mengindikasikan bahwa instrument pengukuran kepribadian Big Five memiliki tingkat reliabilitas yang baik dan dapat diterima.

4.1.5 Preprocessing Data

Tahapan preprocessing dilakukan untuk mempersiapkan data sebelum memasuki fase pemodelan. Proses ini dilakukan menggunakan pendekatan Python, di mana dataset yang disimpan dalam format CSV. Sebanyak 763 sampel data diintegrasikan, yang mencakup faktor demografi, faktor akademik, serta nilai rata-rata dari masing-masing dimensi kepribadian Big Five. Selanjutnya, data tersebut dikategorikan ke dalam tiga kategori berdasarkan penilaian langsung dari guru matematika yang mengamati proses pembelajaran, yaitu cukup (C) dengan rentang nilai 75-82, baik (B) dengan rentang nilai 83-91, dan sangat baik (SB) dengan rentang nilai 92-100. Hasil klasifikasi ini (Gambar 4.1) menunjukkan bahwa

kategori C terdiri dari 155 data (20,31%), kategori B terdiri dari 425 data (55,7%), kategori SB terdiri dari 183 data (23,98%),



Gambar 4. 1. Grafik Penyebaran Kelas Target

Selanjutnya, pembersihan data, atau yang disebut sebagai *cleaning data*, dilakukan untuk mengatasi masalah pada data melalui berbagai langkah, seperti menghapus atribut yang tidak relevan, mengisi data yang hilang, menghapus data yang tidak konsisten, serta mencari dan menghapus data duplikat (Gori et al., 2024). Tahapan ini telah menemukan bahwa terdapat satu fitur yang dihapus, yaitu fitur nama karena dianggap tidak relevan.

Tahap berikutnya adalah melakukan transformasi data pada beberapa atribut yang bersifat kategorikal, baik biner maupun multi. Atribut tersebut akan diubah

menjadi format numerik menggunakan metode label encoding (Gori et al., 2024; Santoso et al., 2024). Label encoding merupakan teknik yang digunakan untuk mengkonversi data kategorikal menjadi data numerik dengan memberikan label angka 0-9 untuk setiap kategori. Hasil dari transformasi ini telah disajikan dalam Tabel 4.8, di mana berdasarkan informasi tersebut dapat diketahui bahwa variabel dependen dalam penelitian ini adalah nilai akhir matematika, yang terdiri dari tiga kelas dengan atribut cukup = 0, baik = 1, dan sangat baik = 2.



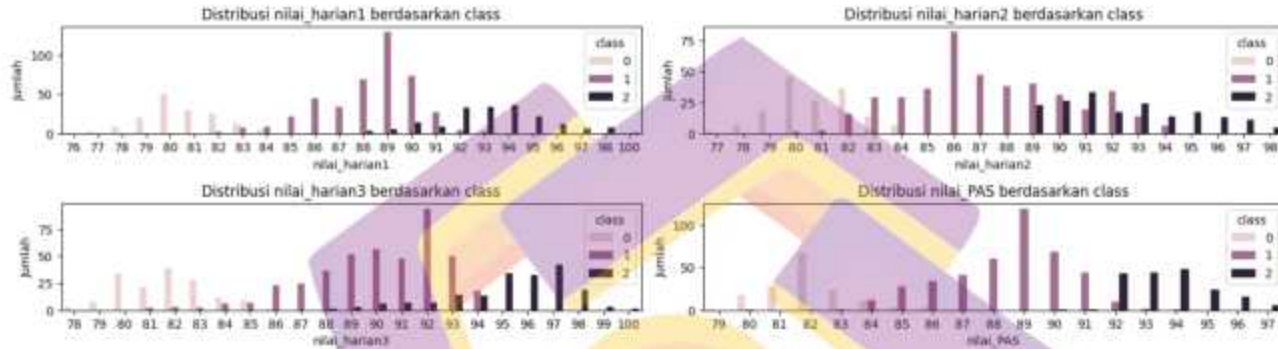
Tabel 4. 8. Informasi Transformasi Data

Kategori	Fitur	Value	Transformasi	Keterangan
	Jenis Kelamin	Perempuan (P)	0	Data Nominal
		Laki-laki (L)	1	
	Jenis Tinggal	Lainnya	0	Data Ordinal
		Asrama	1	
		Wali	2	
Pendidikan Ayah Pendidikan Ibu		Bersama Orang Tua	3	Data Ordinal
		Tidak Sekolah	0	
		Putus Sekolah	1	
		SD/Sederajat	2	
		SMP/Sederajat	3	
		SMA/Sederajat	4	
		Diploma	5	
		S1	6	
		S2	7	
S3	8			
Pekerjaan Ayah Pekerjaan Ibu		Tidak Bekerja	0	Data Ordinal
		Sudah Meninggal	1	
		Lainnya	2	
		Pedagang	3	
		Wiraswasta/Wirausaha	4	
		Karyawan Swasta	5	
PNS/TNI/Polri	6			
Penghasilan Ayah Penghasilan Ibu		Tidak Berpenghasilan	0	Data Ordinal
		Kurang dari Rp.500000	1	
		Rp.500000 s.d Rp.999999	2	
		Rp.1000000 s.d Rp.1999999	3	
		Rp.2000000 s.d Rp.4999999	4	
Rp.5000000 s.d Rp. 20000000	5			

	Anak Kc	0-9	-	Data Nominal	
	Jumlah Saudara Kandung	0-9	-	Data Nominal	
	Status Pernikahan Orang Tua	Salah satu orang tua sudah tiada (meninggal)	0	1	Data Ordinal
Orang tua tidak tinggal bersama lagi (bercerai)		1			
Menikah		2			
Akademik	Nilai Harian 1	75-100	-	Data Nominal	
	Nilai Harian 2	75-100	-		
	Nilai Harian 3	75-100	-		
	Nilai PAS	75-100	-		
	Nilai Akhir (Dependent)	Cukup (C)	0	1	Data Ordinal
		Baik (B)	1		
Baik Sekali (SB)		2			
Kepribadian Big Five	Extraversion	1-5	-	Data Nonimal	
	Agreeableness	1-5	-		
	Conscientiousness	1-5	-		
	Neuroticism	1-5	-		
	Openess	1-5	-		

Setelah dataset dinyatakan lengkap dan bersih, yang berarti tidak ada data yang bernilai null, duplikat, serta telah melalui proses transformasi, sehingga tahapan berikutnya adalah menampilkan distribusi setiap fitur terhadap kelas yang ada. Distribusi ini dapat dilihat pada Gambar 4.2 yang menunjukkan sebaran data akademik di setiap kelas. Gambar tersebut menampilkan distribusi nilai untuk setiap fitur akademik, yaitu nilai_harian1, nilai_harian2, nilai_harian3, dan nilai_PAS terhadap tiga kelas yang tersedia. Secara keseluruhan, terlihat adanya pola peningkatan nilai seiring dengan naiknya kelas, yang paling mencolok terlihat pada fitur nilai_PAS dan nilai_harian3. Hal ini menunjukkan bahwa fitur-fitur tersebut memiliki potensi kontribusi yang cukup signifikan dalam membedakan masing-masing kelas karena distribusi nilainya yang cukup berbeda.



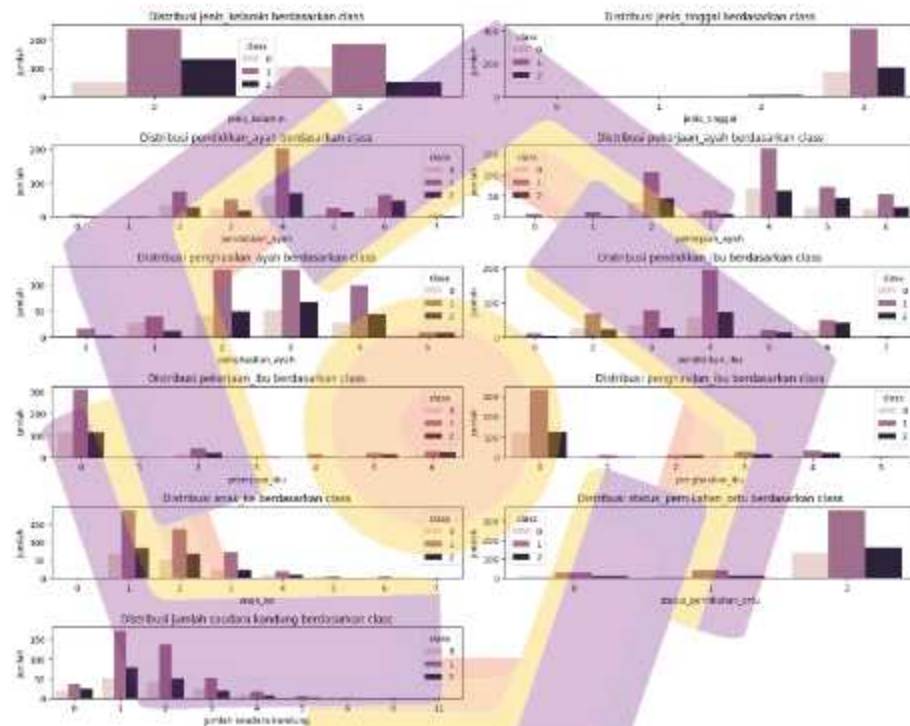


Gambar 4. 2. Distribusi Data Akademik Berdasarkan Nilai Kinerja Akademik

Selanjutnya, distribusi data demografi untuk setiap kelas telah ditunjukkan dalam Gambar 4.3. Berdasarkan analisis distribusi data demografi responden, dapat disimpulkan bahwa mayoritas responden adalah laki-laki (1) yang Sebagian besar tinggal bersama orang tua (3). Dari segi pendidikan, baik ayah maupun ibu responden umumnya berada pada tingkat SMP (3) hingga SMA (4) sederajat. Kondisi ini mencerminkan bahwa latar belakang pendidikan keluarga cenderung berada pada jenjang menengah, meskipun masih terdapat sebagian kecil dengan pendidikan rendah (tidak bersekolah atau hanya SD) serta yang memiliki pendidikan lebih tinggi hingga perguruan tinggi. Dari perspektif ekonomi, pekerjaan ayah pada umumnya para responden memiliki jawaban sebagai pedagang, wiraswasta, atau karyawan swasta (kode 3-5), sedangkan mayoritas pekerjaan ibu tercatat tidak bekerja (0) atau hanya memiliki profesi dalam kategori tertentu dengan jumlah yang kecil. Hal ini sejalan dengan kondisi penghasilan, di mana penghasilan ayah paling banyak berada di kategori Rp.1.000.000 – Rp.4.999.999, sementara penghasilan ibu didominasi oleh kategori tidak berpenghasilan atau kurang dari Rp.500.000. Dengan demikian, peran ekonomi keluarga sebagian besar masih bergantung pada ayah. Selain itu, secara garis besar sebagian besar responden merupakan anak pertama hingga anak ketiga, dengan jumlah saudara kandung antara 1 hingga 3 orang. Status pernikahan orang tua juga mayoritas memiliki status menikah (2), yang menunjukkan kondisi keluarga yang cukup stabil. Kemudian, jika dilihat dari distribusi kelas, hasil belajar (0=cukup, 1=baik, dan 2=sangat baik), tampak bahwa kelas 0 dan kelas 1 lebih mendominasi dibandingkan kelas 2 pada hampir semua variabel demografis. Dengan kata lain,

responden dengan hasil belajar dalam kategori cukup (0) dan baik (1) lebih banyak ditemukan, sedangkan kategori sangat baik (2) jumlahnya lebih sedikit. Secara keseluruhan, distribusi ini menunjukkan bahwa responden umumnya berasal dari keluarga yang memiliki dukungan orang tua yang utuh, pendidikan orang tua pada tingkat menengah, dengan ayah sebagai penyokong utama ekonomi keluarga, serta jumlah anak yang tergolong sedikit. Sehingga, pada dasarnya pola demografi ini memberikan wawasan penting mengenai latar belakang social ekonomi keluarga yang dapat mempengaruhi prestasi siswa, serta berpotensi memberikan kontribusi dalam proses klasifikasi.

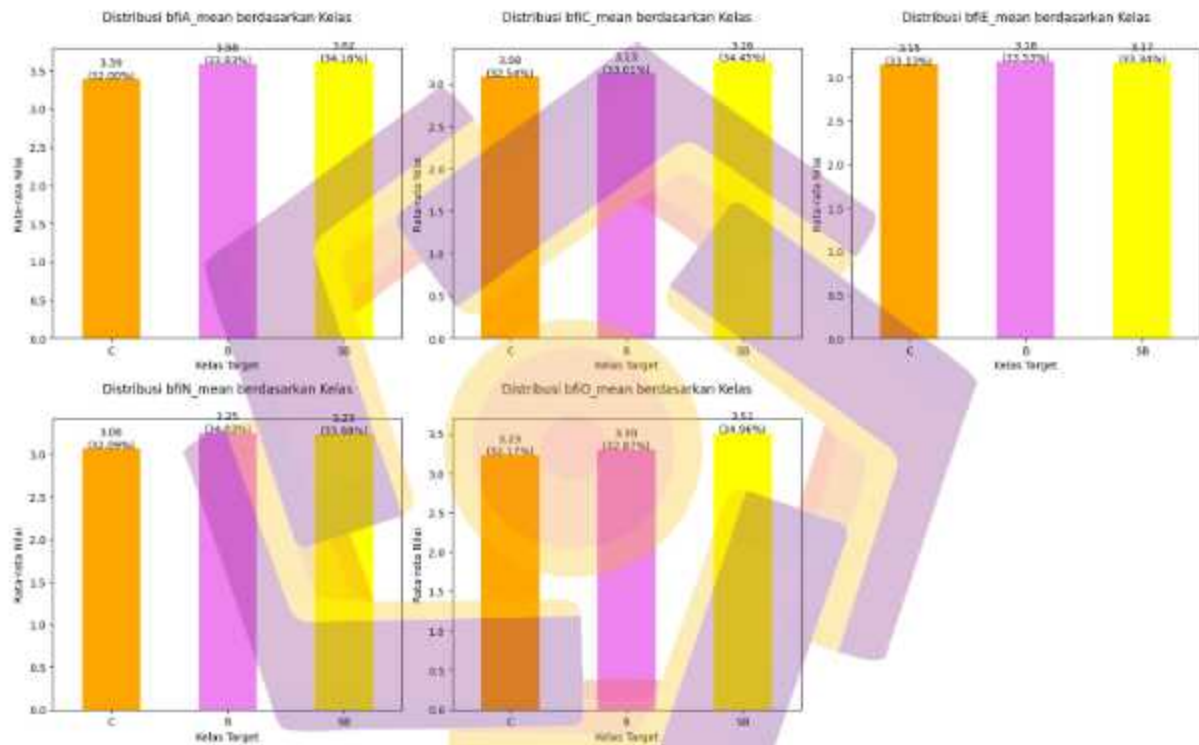




Gambar 4. 3. Distribusi Data Demografi Berdasarkan Nilai Kinerja Akademik

Gambar 4.4 menggambarkan distribusi data kepribadian dalam 5 dimensi untuk setiap kelas. Setiap grafik menampilkan nilai rata-rata dari dimensi kepribadian Big Five terkait dengan kelas target, yaitu cukup atau C (0), baik atau B (1) dan sangat baik atau SB (2). Secara keseluruhan, kelas SB menunjukkan nilai rata-rata kepribadian yang lebih tinggi, terutama pada dimensi *openness*, *agreeableness* dan *conscientiousness* yang memperlihatkan perbedaan paling signifikan antar kelas. Dimensi *neuroticism* juga sedikit lebih tinggi pada kelas SB, sedangkan kelas C memiliki rata-rata terendah di hampir semua dimensi. Sementara itu, *extraversion* menunjukkan variasi yang paling kecil. Oleh karena itu, *openness*, *agreeableness* dan *conscientiousness* berpotensi menjadi fitur yang paling membedakan antar kelas.





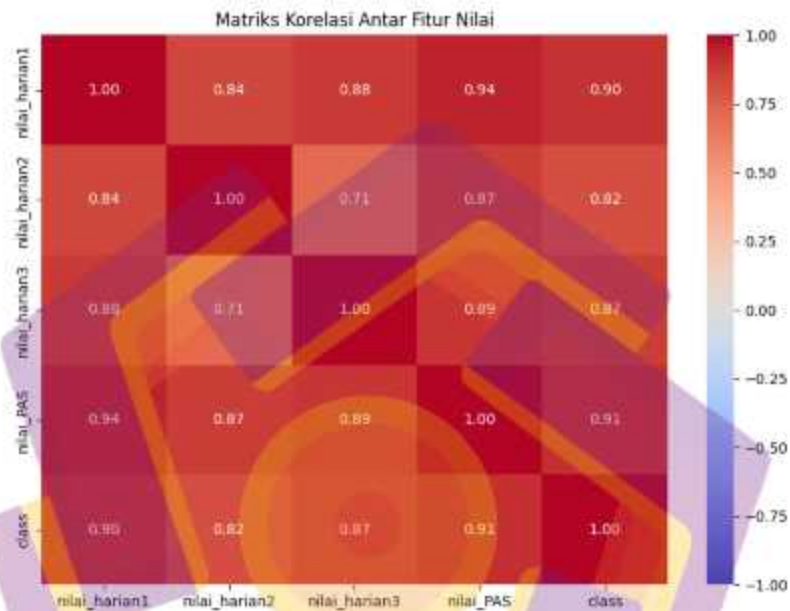
Gambar 4. 4. Distribusi Data Kepribadian Berdasarkan Nilai Kinerja Akademik

4.2. Analisis Hubungan Faktor

Menyusun matriks korelasi dan ketergantungan antar fitur bertujuan untuk menilai sejauh mana suatu fitur dapat mempengaruhi fitur lainnya. Korelasi yang signifikan terjadi ketika dua fitur memiliki nilai yang jauh dari nol, baik dalam arah positif maupun negatif (El-Keiey et al., 2022). Sebaliknya, jika nilai-nilai tersebut mendekati nol, maka hubungan antara kedua fitur tersebut dianggap lemah. Korelasi positif memiliki arti bahwa peningkatan nilai pada satu fitur diikuti oleh peningkatan nilai pada fitur lain yang berkorelasi. Sementara itu, korelasi negatif terjadi ketika kenaikan nilai pada satu fitur menyebabkan penurunan nilai pada fitur lain yang berkorelasi.

Analisis hubungan dilakukan pada fitur-fitur yang ada dalam faktor akademik, demografi, dan kepribadian terhadap kinerja akademik siswa. Hasil analisis hubungan antar fitur akademik dengan fitur kinerja akademik atau nilai akhir (class), dapat dilihat pada Gambar 4.5. Gambar tersebut menunjukkan bahwa seluruh fitur yang terdiri dari nilai_PAS, nilai_harian1, nilai_harian2, dan nilai_harian3 teridentifikasi memiliki korelasi positif dengan nilai yang cukup tinggi ($>0,1$). Artinya keempat fitur tersebut memiliki pengaruh signifikan terhadap hasil kinerja akademik siswa, di mana fitur dengan tingkat korelasi tertinggi ialah nilai_PAS, diikuti oleh nilai_harian2, nilai_harian1, dan nilai_harian1. Hal ini menunjukkan bahwa siswa yang mendapatkan nilai tinggi dalam salah satu aspek penilaian cenderung juga memperoleh nilai tinggi pada aspek lainnya. Sebaliknya, jika nilai akademik siswa menurun, maka kinerja akademik mereka secara keseluruhan juga akan mengalami penurunan. Oleh karena itu, dapat disimpulkan

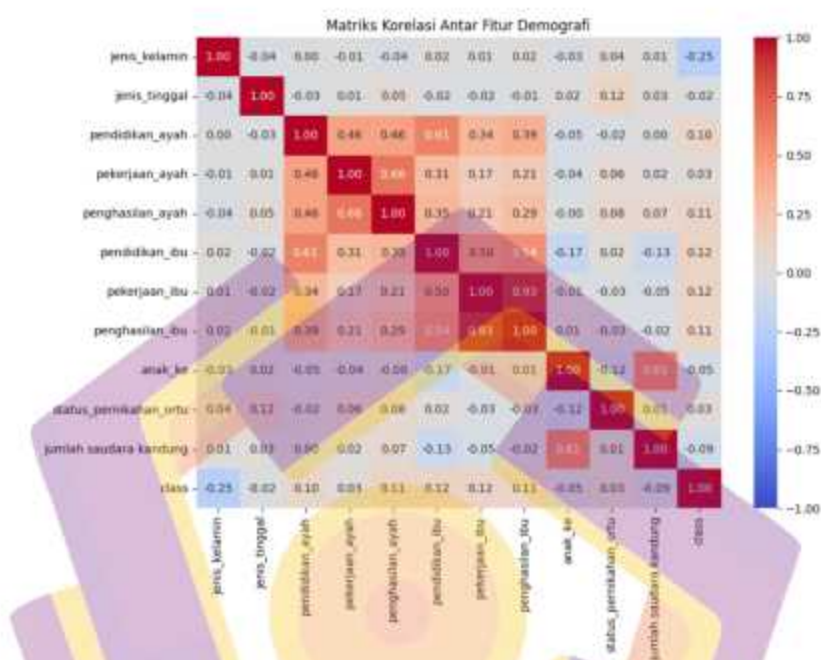
bahwa fitur akademik ini memiliki korelasi yang sangat kuat terhadap kinerja akademik siswa.



Gambar 4. 5. Matriks Korelasi Fitur Akademik dengan Fitur Kinerja Akademik

Analisis korelasi juga dilakukan antar fitur demografi dan fitur kinerja akademik (class) seperti yang ditunjukkan dalam Gambar 4.6, dengan fitur demografi terdiri dari jenis_kelamin, jenis_tinggal, pendidikan_ayah, pekerjaan_ayah, penghasilan_ayah, pendidikan_ibu, pekerjaan_ibu, penghasilan_ibu, anak_ke, status_pernikahan_ortu, dan jumlah_saudara_kandung. Hasil analisis menunjukkan bahwa korelasi antara fitur demografi dan kinerja akademik (class) menunjukkan variasi nilai yang signifikan, di mana terdapat dua kategori korelasi yang teridentifikasi. Korelasi negatif dengan urutan tertinggi

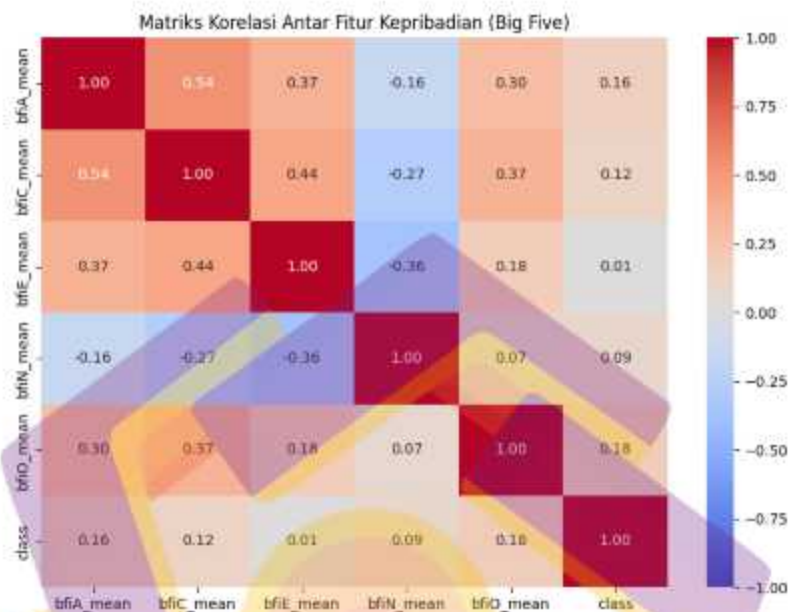
dimulai dari jenis_kelamin, diikuti oleh jumlah_saudara_kandung, anak_ke dan jenis_tinggal. Korelasi positif tertinggi dimulai dari pendidikan_ibu, pekerjaan_ibu, penghasilan_ayah, penghasilan_ibu, pendidikan_ayah, pekerjaan_ayah, dan status_pernikahan_ortu. Berdasarkan temuan tersebut, sejumlah fitur demografi menunjukkan adanya korelasi negatif, yang mengindikasikan bahwa peningkatan nilai pada fitur-fitur tersebut cenderung berkaitan dengan penurunan nilai pada fitur kinerja akademik (class). Sedangkan, pada fitur-fitur lain yang menunjukkan korelasi positif, ini menunjukkan bahwa peningkatan nilai pada fitur-fitur tersebut berhubungan dengan peningkatan nilai kinerja akademik. Secara garis besar latar belakang pendidikan dan ekonomi keluarga memiliki pengaruh positif terhadap nilai kelas, sedangkan faktor-faktor internal masing-masing pribadi seperti jenis kelamin dan kondisi atau lingkungan keluarga memberikan dampak negatif terhadap kinerja akademik.



Gambar 4. 6. Matriks Korelasi Fitur Demografi dengan Fitur Kinerja Akademik

Analisis korelasi juga dilakukan antar dimensi kepribadian big five dan fitur kinerja akademik (class), seperti yang ditunjukkan dalam Gambar 4.7. Kepribadian model Big Five sendiri terdiri dari lima aspek utama yaitu bfiA_mean (Agreeableness), bfiC_mean (Conscientiousness), bfiE_mean (Extraversion), bfiN_mean (Neuroticism), dan bfiO_mean (Openness). Dimensi bfiA_mean (Agreeableness) merupakan cerminan dari sikap kooperatif, bfiC_mean (Conscientiousness) merupakan cerminan dari sikap disiplin, bfiE_mean (Extraversion) merupakan cerminan dari sikap kehidupan sosial, bfiO_mean (Openness) merupakan cerminan dari sikap terbuka terhadap hal-hal baru, dan

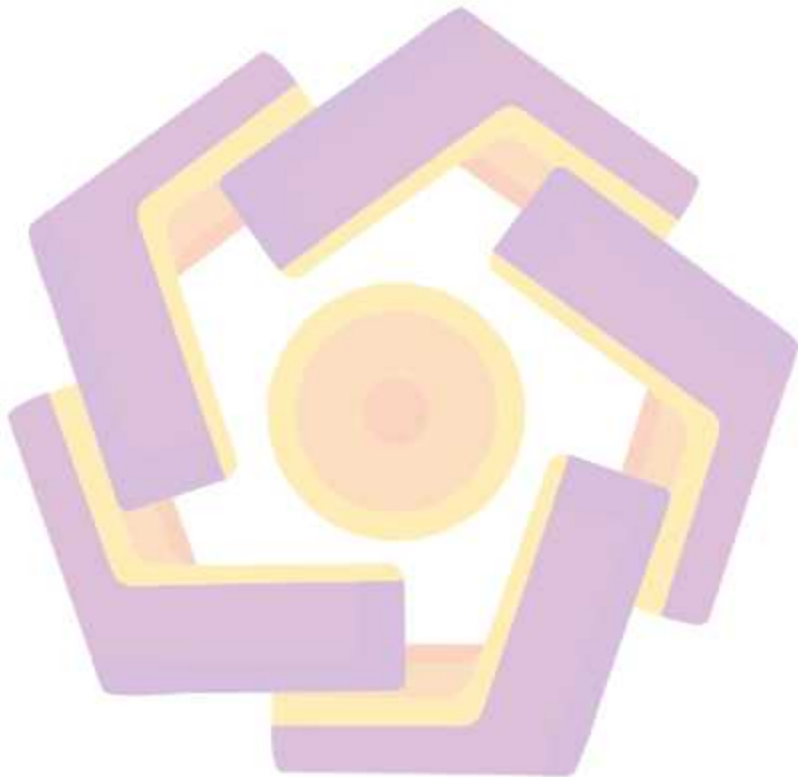
bfiN_mean (Neuroticism) merupakan cerminan dari sikap emosional seseorang artinya sejauh mana emosional dapat mempengaruhi tindakan sehari-hari mereka. Analisis korelasi yang dilakukan antara dimensi kepribadian dan kinerja akademik (class) menunjukkan bahwa dari lima dimensi kepribadian, terdapat tiga dimensi yang memiliki korelasi positif signifikan, sedangkan tiga dimensi lainnya menunjukkan korelasi positif yang non-signifikan. Dimensi dengan urutan korelasi positif tertinggi ialah bfiO_mean (Openness) diikuti oleh bfiA_mean (Agreeableness) dan bfiC_mean (Conscientiousness). Sedangkan untuk dimensi bfiN_mean (Neuroticism) dan bfiE_mean (Extraversion) ditemukan dengan nilai korelasi yang sangat kecil, bahkan mendekati 0. Temuan ini menunjukkan bahwa individu dengan tingkat keterbukaan (Openness), individu dengan tingkat kooperatif (Agreeableness), dan individu dengan tingkat kedisiplinan (Conscientiousness) yang tinggi cenderung menunjukkan performa akademik yang lebih baik. Sebaliknya, tingkat emosional (Neuroticism), dan sifat ekstrovert (Extraversion) individu dapat dikatakan tidak memiliki korelasi atau tidak memiliki pengaruh yang berarti terhadap kinerja akademik siswa.



Gambar 4. 7. Matriks Korelasi Fitur Kepribadian dengan Fitur Kinerja Akademik

Berdasarkan hasil eksperimen yang telah dilakukan dalam penelitian ini, temuan menunjukkan bahwa sejumlah fitur seperti jenis_kelamin, pendidikan_ibu, pendidikan_ayah, penghasilan_ayah, penghasilan_ibu, nilai_PAS, nilai_harian1, nilai_harian2, nilai_harian3, serta dimensi kepribadian bfiC_mean (Conscientiousness), bfiO_mean (Openness) dan bfiA_mean (Agreeableness), memiliki pengaruh yang signifikan terhadap performa model prediksi kinerja akademik siswa. Temuan ini juga didukung oleh analisis kepentingan hubungan yang ditunjukkan dalam Gambar 4.8, di mana hasil menunjukkan bahwa faktor akademik menunjukkan korelasi yang paling signifikan terhadap kinerja akademik, di mana nilai_PAS mencatat korelasi positif tertinggi, diikuti oleh nilai_harian1, nilai_harian2, dan nilai_harian3. Temuan ini mengindikasikan bahwa dibandingkan

dengan faktor lainnya, faktor akademik menjadi indikator utama, di mana hal ini juga sejalan dengan temuan (El-Keiey et al., 2022; Khairy et al., 2024; Nachouki et al., 2023; Roslan & Chen, 2023) yang menunjukkan bahwa diantara fitur lainnya, fitur akademik memiliki hubungan paling kuat dengan kinerja akademik.





Gambar 4. 8 Nilai Korelasi Semua Fitur dengan Fitur Kinerja Akademik

Berdasarkan hasil eksperimen yang dilakukan dengan menggabungkan beberapa fitur kepribadian dalam model prediksi, terungkap bahwa setiap fitur memiliki pengaruh yang berbeda terhadap kinerja model, baik secara positif maupun negatif. Hasil penelitian ini menunjukkan bahwa penerapan fitur dimensi *bfiA_mean* (*Agreeableness*), *bfiC_mean* (*Conscientiousness*) dan *bfiO_mean* (*Openness*) dapat meningkatkan kinerja model, sedangkan penambahan kedua dimensi lainnya justru mengakibatkan penurunan kinerja model prediksi.

Temuan ini juga sejalan dengan temuan (Meyer et al., 2023; Rico-Juan et al., 2024; Shaninah & Mohd Noor, 2024), di mana siswa yang memiliki tingkat *conscientiousness* tinggi atau yang mencerminkan sifat teratur dan disiplin, cenderung memperoleh kinerja akademik yang lebih baik. Selain itu, siswa dengan tingkat *openness* tinggi atau yang mencerminkan sikap terbuka terhadap pengalaman baru juga cenderung mencapai kinerja akademik yang lebih baik (Rico-Juan et al., 2024; Shaninah & Mohd Noor, 2024). Sedangkan, dimensi *agreeableness* berperan dalam membangun hubungan sosial yang harmonis di dalam kelas, sehingga menghasilkan pandangan positif dari guru dan mendukung terbentuknya suasana belajar yang kondusif (Meyer et al., 2023; Shaninah & Mohd Noor, 2024). Penelitian yang dilakukan oleh (Meyer et al., 2023) menjelaskan lebih lanjut bahwa dimensi *openness* sangat berkaitan erat dengan kinerja akademik, karena seseorang akan memiliki ketertarikan dan keinginan lebih untuk menjelajahi hal-hal baru, terutama dalam memperdalam suatu materi pembelajaran sebelum ujian akhir dilakukan. Sebaliknya, dimensi *conscientiousness* yang berkaitan dengan kepatuhan dan disiplin, akan menunjukkan hubungan yang lebih kuat

dengan nilai akademik di kelas. Serta, dimensi *agreeableness* juga menunjukkan hubungan sosial yang lebih harmonis dan kondusif akan menunjukkan hubungan yang lebih kuat dengan nilai akademik di kelas. Sehingga, pada dasarnya penjelasan tersebut memperkuat temuan dalam penelitian ini, di mana hasil korelasi (Gambar 4.8) menunjukkan bahwa peningkatan pada dimensi *agreeableness*, *openness* dan *conscientiousness* cenderung diikuti oleh peningkatan nilai akhir siswa. Hal ini terjadi karena siswa yang disiplin, mampu membentuk lingkungan sosial yang harmonis dan memiliki kemauan untuk belajar akan lebih mudah menyerap materi, sehingga lebih siap dalam menghadapi ujian. Sementara itu, untuk dimensi *bf1N_mean* (Neuroticism), dan *bf1E_mean* (Extraversion) ditemukan sebagai fitur yang tidak memiliki pengaruh terhadap kinerja akademik siswa. Temuan ini diperkuat oleh penelitian (Shaninah & Mohd Noor, 2024) yang mengindikasikan bahwa dimensi *neuroticism* memiliki dampak yang tergolong kecil terhadap kinerja akademik mahasiswa di Libya, ini disebabkan oleh stabilitas emosional yang sangat dipengaruhi oleh berbagai faktor eksternal, salah satunya lingkungan yang bervariasi antara individu satu dengan yang lainnya. Variasi ini menyebabkan pengaruh *neuroticism* sulit untuk digeneralisasi dan hasilnya kurang konsisten sebagai indikator kinerja akademik. Selain itu, penelitian lain (Rico-Juan et al., 2024) juga menemukan bahwa dimensi *extraversion* tidak berpengaruh secara signifikan terhadap kinerja akademik seseorang, ini menunjukkan bahwa nilai korelasinya tidak terlalu kuat. Namun, penelitian-penelitian tersebut (Meyer et al., 2023; Rico-Juan et al., 2024; Shaninah & Mohd Noor, 2024) menyoroti bahwa hasil korelasi yang diperoleh dapat dipengaruhi oleh tingkat keberagaman data yang

digunakan dalam pengembangan model. Oleh karena itu, penting untuk dicatat bahwa hasil dari satu penelitian mungkin berbeda dari penelitian lainnya, tergantung pada jenis data yang digunakan. Selain itu, jenis penelitian ini biasanya memanfaatkan data yang bersifat pribadi (*privacy-sensitive*), sehingga tidak dapat dengan mudah digeneralisasikan di antara berbagai studi. Dengan demikian, berdasarkan analisis hubungan faktor dan eksperimen yang dilakukan, penelitian ini menunjukkan bahwa dari lima dimensi kepribadian, hanya dua dimensi yang memiliki pengaruh signifikan terhadap kinerja model, meskipun tingkat pengaruhnya masih tergolong kecil.

Sementara itu, dalam aspek demografis, temuan menunjukkan bahwa fitur jenis kelamin dan faktor sosial-ekonomi, termasuk tingkat pendidikan orang tua serta penghasilan orang tua, memberikan dampak yang lebih signifikan kepada performa model prediksi dibandingkan dengan faktor demografis lainnya. Temuan ini juga didukung oleh penelitian (Roslan & Chen, 2023; Shaninah & Mohd Noor, 2024), yang menunjukkan bahwa faktor ekonomi, pendidikan orang tua, dan jenis kelamin memiliki hubungan yang paling kuat dengan kinerja akademik. Sementara itu, untuk fitur lainnya seperti jumlah saudara kandung, jenis tempat tinggal, pekerjaan ayah, pekerjaan ibu, status pernikahan orang tua, dan anak ke memiliki hubungan yang lemah dengan kinerja akademik siswa. Temuan ini menunjukkan bahwa faktor-faktor tersebut tidak memberikan dampak yang signifikan terhadap kinerja akademik siswa. Hasil ini juga sejalan dengan beberapa studi sebelumnya, di mana penelitian (El-Keiey et al., 2022) berhasil menunjukkan bahwa jumlah saudara kandung mengakibatkan penurunan tingkat akurasi dalam

model prediksi, sehingga dianggap tidak berpengaruh secara signifikan. Penelitian (Roslan & Chen, 2023) berhasil menunjukkan bahwa dalam model prediksi yang dikembangkan, fitur status pernikahan orang tua hanya teridentifikasi satu kali dari delapan prediksi, sehingga fitur ini dianggap tidak relevan. Selain itu, penelitian (Roslan & Chen, 2023) juga mengungkapkan bahwa tingkat pendidikan orang tua lebih berperan penting dibandingkan dengan status pekerjaan dalam menentukan keberhasilan akademik. Hal ini menunjukkan bahwa faktor sosial-ekonomi tidak selalu menjadi penentu utama pencapaian kinerja akademik siswa, sementara pendidikan orang tua lebih berkontribusi dalam mendukung keberhasilan akademik siswa.

4.3. Pemodelan dan Evaluasi

Pada tahap ini, dijelaskan hasil dan analisis dari model prediksi kinerja akademik yang telah dikembangkan. Pemodelan dilakukan dengan menerapkan empat skenario yang berbeda, yang di mana menggunakan dataset yang sama. Evaluasi model juga dilakukan melalui pengujian menggunakan matriks evaluasi untuk mengukur tingkat akurasi dari model prediksi yang telah dibangun. Proses pengembangan model ini melibatkan pembagian dataset menjadi 70% untuk data pelatihan dan 30% untuk data pengujian. Uji coba dalam penelitian ini dilakukan dengan memanfaatkan Python serta Google Collab.

4.3.1 Skenario 1

Dalam skenario pertama, pemodelan untuk memprediksi kinerja akademik dilakukan dengan menggunakan metode klasifikasi Random Forest, serta dataset yang digunakan menunjukkan ketidakseimbangan kelas. Dataset dibiarkan dalam

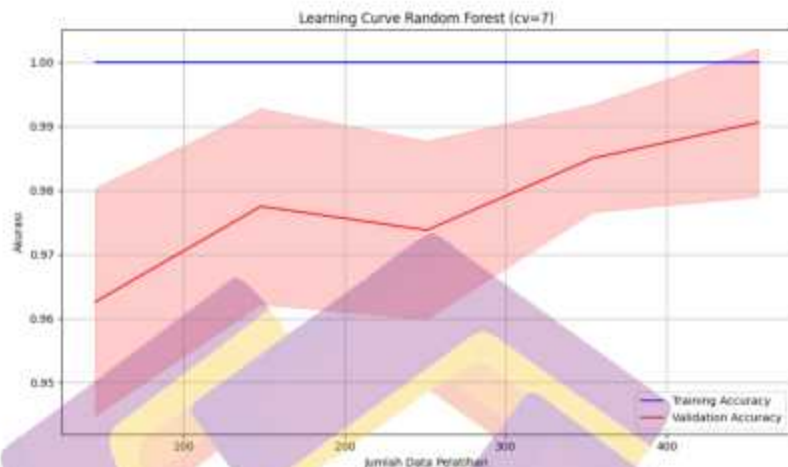
kondisi tidak seimbang untuk mengevaluasi dampak ketidakseimbangan tersebut terhadap kinerja model prediksi. Kemudian, sebelum melanjutkan ke tahap pelatihan dan prediksi, evaluasi awal dilakukan pada data latih dengan membandingkan kinerja k-fold validation 2 hingga 10 untuk mendapatkan nilai k-fold yang paling optimal dalam membangun model prediksi. Pengujian ini juga dimaksudkan untuk mencegah terjadinya overfitting, serta mengukur kestabilan data. Hasil pengujian cross validation sendiri dapat dilihat pada Tabel 4.9.

Tabel 4. 9. Hasil Pengujian dengan K-Fold Cross Validation Skenario 1

K-fold Validation	Average Cross-Validation
2	0.9832
3	0.9869
4	0.9869
5	0.9888
6	0.9888
7	0.9906
8	0.9869
9	0.9868
10	0.9888

Pada tahap ini, pengujian dilakukan terhadap nilai k-fold dari 2 hingga 10 untuk menentukan nilai k-fold yang paling optimal dalam membangun model prediksi menggunakan Random Forest. Berdasarkan hasil percobaan, data latih yang dievaluasi menunjukkan performa terbaik ketika diterapkan nilai k-fold = 7, di mana tingkat akurasi yang dicapai sebesar 0,9906 (99,06%). Ini menunjukkan bahwa data yang digunakan berada pada titik yang cukup stabil. Selain itu, meskipun beberapa nilai cross-validation pada berbagai k-fold menunjukkan tingkat akurasi yang serupa, terdapat perbedaan dalam hal kestabilan validasinya.

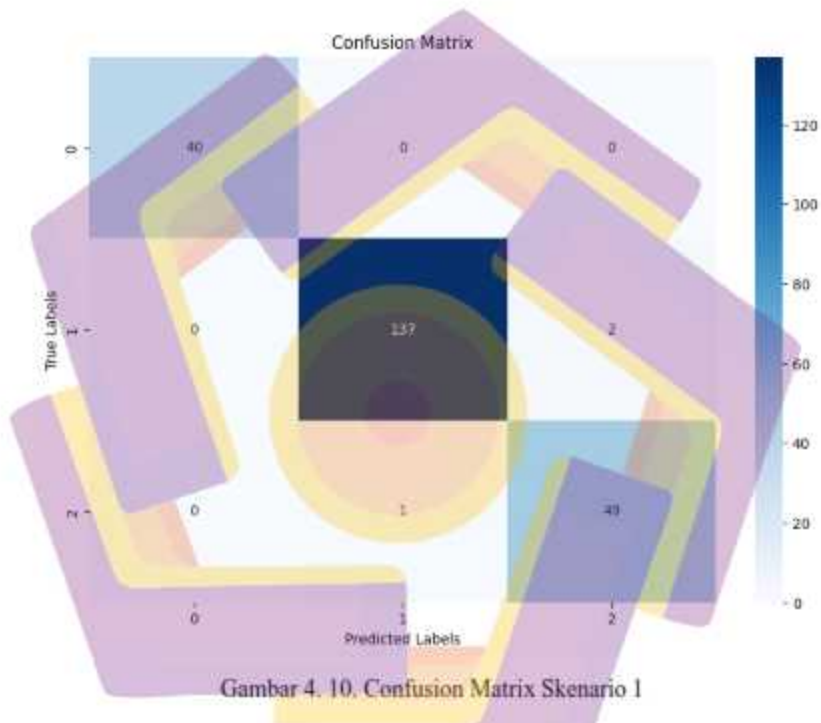
Kestabilan ini tercermin pada grafik learning curve yang ditunjukkan oleh Gambar 4.9. Gambar grafik tersebut menggambarkan proses training score pada model dengan menerapkan $k\text{-fold}=7$. Grafik tersebut menunjukkan perbandingan antara akurasi data pelatihan (kurva biru) dan akurasi validasi (kurva merah) seiring dengan peningkatan jumlah data pelatihan. Kurva biru yang mewakili akurasi pelatihan menunjukkan bahwa akurasi konsisten berada di angka 100% untuk seluruh jumlah data pelatihan. Ini menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mengingat data pelatihan, sehingga ada kemungkinan model mengalami overfit pada data pelatihan. Sementara itu, kurva merah yang menggambarkan akurasi validasi menunjukkan kestabilan yang cukup baik, meskipun terjadi peningkatan dan penurunan secara terus-menerus seiring dengan bertambahnya jumlah data pelatihan. Meskipun begitu, grafik menunjukkan bahwa nilai akurasi model cukup stabil di atas 0,97 (97%) pada semua jumlah data pelatihan. Secara garis besar, analisis ini menunjukkan bahwa terdapat perbedaan kinerja yang signifikan antara pelatihan dan validasi. Meskipun begitu, secara keseluruhan model menunjukkan kinerja yang cukup baik, serta perlu di garis bawah bahwa jumlah data pelatihan memiliki pengaruh yang signifikan terhadap performa model.



Gambar 4. 9. Grafik Learning Curve k-fold = 7 Skenario 1

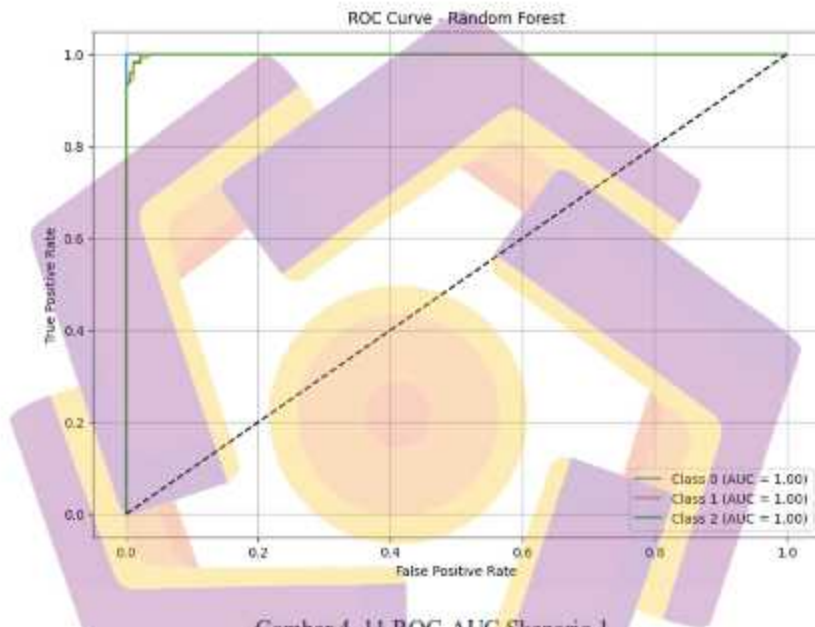
Berikutnya, pembahasan mengenai hasil evaluasi yang diperoleh melalui metrics evaluasi menggunakan confusion matrix, sebagaimana ditampilkan dalam Gambar 4.10. Gambar tersebut memberikan ilustrasi mengenai kinerja model dalam mengklasifikasikan tiga kategori kinerja akademik. Pada kelas 0 (cukup), model menunjukkan kinerja yang sangat memuaskan dengan semua data berhasil diklasifikasikan sesuai dengan kelasnya. Sementara untuk kelas 1 (baik), terdapat 137 data diklasifikasikan sesuai dengan kelasnya, sedangkan 2 data salah diklasifikasikan sebagai kelas 2 (sangat baik). Sementara itu, untuk kelas 2 (sangat baik), dari 50 data yang ada, 49 benar diklasifikasikan sesuai kelasnya dan hanya 1 data yang salah diklasifikasikan sebagai kelas 0 (baik). Selain itu, hasil evaluasi menggunakan beberapa metrics evaluasi juga menunjukkan bahwa model berhasil memperoleh nilai akurasi sebesar 98,69%, nilai precision sebesar 98,704%, nilai recall sebesar 98,69%, dan nilai f1-score sebesar 98,694%. Hasil ini didapatkan

dengan menggunakan model Random Forest yang menerapkan nilai parameter secara default, di mana parameter yang digunakan dalam skenario ini mencakup: $n_estimators = 100$, $max_depth = none$, $max_features = sqrt$, $min_samples_leaf = 1$, $min_samples_split = 2$, dan $random_state = 42$.



Metode ROC-AUC juga diterapkan untuk merangkum tingkat akurasi keseluruhan dari kinerja model. Skor ROC-AUC (Gambar 4.11) memiliki nilai yang berkisar antara 0 hingga 1, di mana nilai 0 menunjukkan bahwa model tidak akurat sama sekali, sedangkan nilai 1 menunjukkan akurasi yang sangat tinggi. Hasil evaluasi dengan metode ini menunjukkan bahwa model mencapai skor 99,927%. Ini menunjukkan bahwa model dapat melakukan klasifikasi dengan

tingkat keakuratan yang cukup baik. Temuan ini menunjukkan bahwa model mampu mengklasifikasikan data yang belum pernah dilihat sebelumnya dengan tingkat keakuratan yang cukup baik, meskipun terdapat ketidakseimbangan dalam kelas pada dataset.



Gambar 4. 11 ROC-AUC Skenario 1

Berdasarkan hasil eksperimen yang dilakukan pada skenario pertama, dapat diketahui juga bahwa setiap fitur yang digunakan dalam pemodelan memiliki kontribusi yang berbeda dalam mempengaruhi keputusan model, seperti yang terlihat pada Tabel 4.10. Hasil pemodelan menunjukkan bahwa dari total 12 fitur yang diterapkan, empat fitur yang berhubungan dengan aspek akademik menunjukkan pengaruh yang paling signifikan dibandingkan dengan fitur lainnya. Meskipun begitu, fitur-fitur kepribadian dan demografis tetap memberikan

kontribusi, meskipun dalam proporsi yang sangat kecil. Berdasarkan hubungan yang ada, dapat disimpulkan bahwa faktor akademik menunjukkan pengaruh yang paling signifikan dalam pengambilan keputusan.

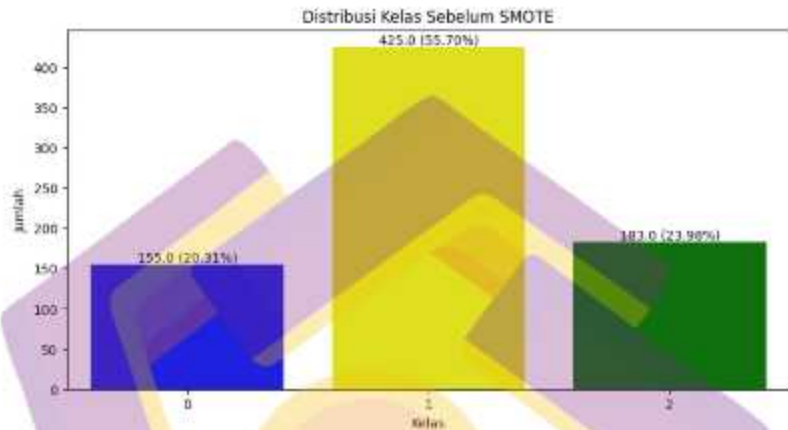
Tabel 4. 10. Bobot Kepentingan Fitur (Feature Importance) Skenario 1

NO	Nama Fitur	Kepentingan Fitur (Importance)
1	nilai PAS	0.323578
2	nilai harian1	0.238357
3	nilai harian3	0.200057
4	nilai harian2	0.186058
5	bfiO mean	0.012343
6	bfiA mean	0.011518
7	bfiC mean	0.009294
8	penghasilan ayah	0.005047
9	jenis kelamin	0.004128
10	pendidikan ibu	0.004125
11	pendidikan ayah	0.003515
12	penghasilan ibu	0.001980

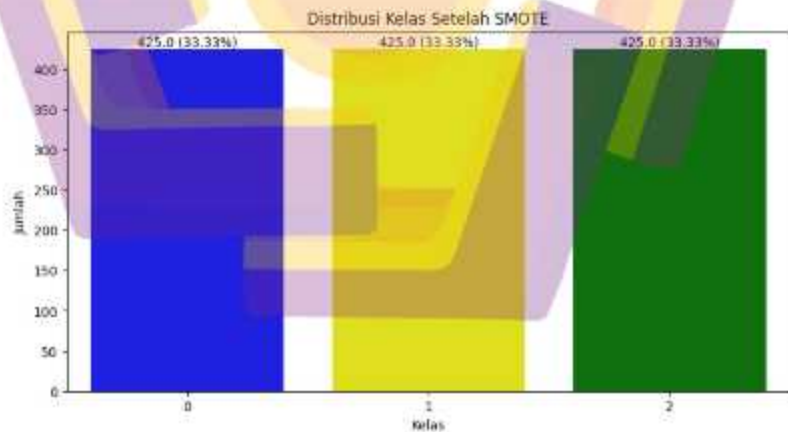
4.3.2 Skenario 2

Dalam skenario kedua, pemodelan untuk memprediksi kinerja akademik dilakukan dengan menggunakan metode klasifikasi Random Forest, serta menerapkan teknik SMOTE untuk menyelesaikan permasalahan ketidakseimbangan kelas. Penerapan SMOTE ini dilakukan dengan cara tidak menghapus data yang ada, melainkan membuat data sintesis baru dari contoh data yang sudah ada pada data latih minoritas. Namun, perlu diketahui bahwa sebelum evaluasi dilakukan, teknik SMOTE akan diterapkan terlebih dahulu pada dataset. Seperti yang terlihat pada Gambar 4.12 dan Gambar 4.13, di mana penerapan SMOTE dilakukan pada seluruh dataset, yaitu sebelum data dibagi menjadi data latih dan data uji. Hasil dari proses ini menunjukkan bahwa jumlah data pada kelas

minoritas telah disintesis hingga seimbang dengan jumlah data pada kelas mayoritas, sehingga distribusi kelas menjadi lebih proporsional.



Gambar 4. 12. Distribusi Kelas Sebelum SMOTE Skenario 2



Gambar 4. 13. Distribusi Kelas Setelah SMOTE Skenario 2

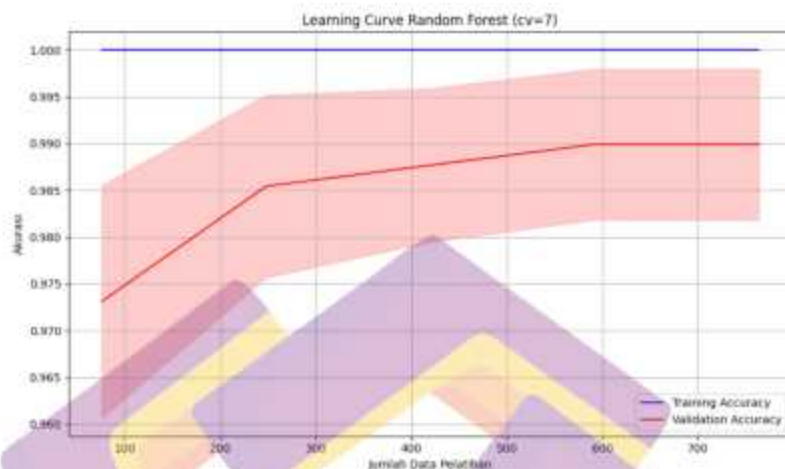
Evaluasi awal menggunakan teknik cross validation dilakukan terhadap data latih dengan membandingkan performa nilai k-folds validation 2 hingga 10 untuk mendapatkan nilai k-fold yang paling optimal dalam membangun model prediksi menggunakan Random Forest. Pengujian ini juga dimaksudkan untuk mencegah terjadinya *overfitting*, serta mengukur kestabilan data. Hasil pengujian cross validation sendiri dapat dilihat pada Tabel 4.11 .

Tabel 4. 11. Hasil Pengujian dengan K-Fold Cross Validation Skenario 2

K-fold Validation	Average Cross-Validation
2	0,9878
3	0,9888
4	0,9854
5	0,9899
6	0,9910
7	0,9922
8	0,9910
9	0,9899
10	0,9910

Berdasarkan hasil percobaan, data latih yang dievaluasi menunjukkan performa terbaik ketika diterapkan nilai k-fold = 7, di mana tingkat akurasi yang dicapai sebesar 0,9922 (99,22%). Ini menunjukkan bahwa data yang digunakan berada pada titik yang cukup stabil. Selain itu, meskipun beberapa nilai cross-validation pada berbagai k-fold menunjukkan tingkat akurasi yang serupa, terdapat perbedaan dalam hal kestabilan validasinya. Kestabilan ini tercermin pada grafik learning curve yang ditunjukkan oleh Gambar 4.14. Gambar grafik tersebut menggambarkan proses training score pada model dengan menerapkan k-fold=7. Grafik tersebut menunjukkan perbandingan antara akurasi data pelatihan

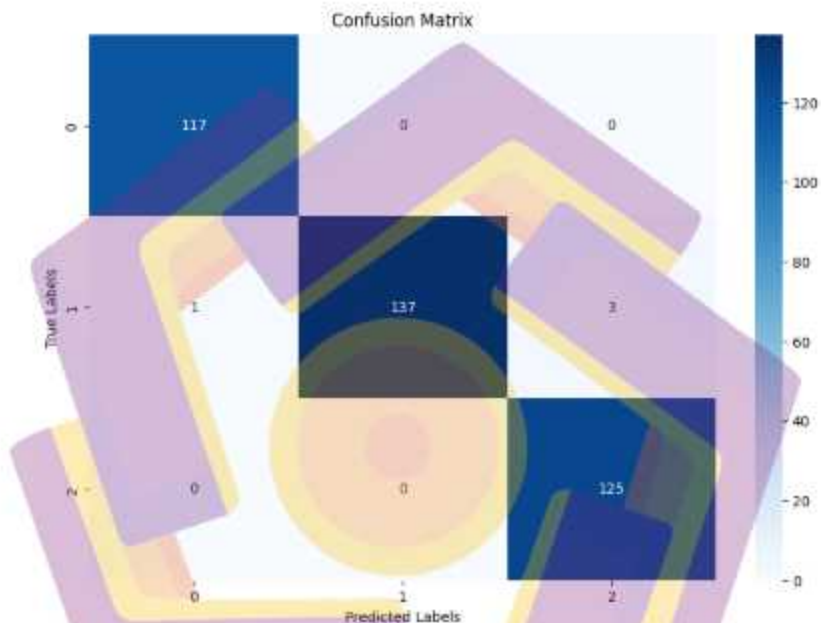
(kurva biru) dan akurasi validasi (kurva merah) seiring dengan peningkatan jumlah data pelatihan. Kurva biru yang mewakili akurasi pelatihan menunjukkan bahwa akurasi konsisten berada di angka 100% untuk seluruh jumlah data pelatihan. Ini menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mengingat data pelatihan, sehingga ada kemungkinan model mengalami overfit pada data pelatihan. Sementara itu, kurva merah yang menggambarkan akurasi validasi menunjukkan kestabilan yang cukup baik, serta terjadi peningkatan secara terus menerus seiring dengan bertambahnya jumlah data pelatihan. Grafik tersebut menunjukkan bahwa nilai akurasi model cukup stabil diatas 0,98 (98%) pada semua jumlah data pelatihan. Secara garis besar, analisis ini menunjukkan bahwa terdapat perbedaan kinerja yang signifikan antara pelatihan dan validasi. Meskipun begitu, secara keseluruhan model menunjukkan kinerja yang cukup baik, serta perlu di garis bawahi bahwa jumlah data pelatihan memiliki pengaruh yang signifikan terhadap performa model.



Gambar 4. 14. Grafik Learning Curve k-fold = 10 Skenario 2

Berikutnya, pembahasan mengenai hasil evaluasi yang diperoleh melalui metrics evaluasi menggunakan confusion matrix, sebagaimana ditampilkan dalam Gambar 4.15. Gambar tersebut memberikan ilustrasi mengenai kinerja model dalam mengklasifikasikan tiga kategori kinerja akademik. Pada kelas 1 (baik), terdapat 137 data yang berhasil diklasifikasikan dengan tepat, sementara 3 data salah dikategorikan sebagai kelas 2 (baik sekali) dan 1 data salah dikategorikan sebagai kelas 1 (cukup). Sementara itu, untuk kelas 0 (cukup) dan kelas 2 (sangat baik), model menunjukkan kinerja yang sangat memuaskan dengan semua data berhasil diklasifikasikan dengan benar. Selain itu, hasil evaluasi menggunakan beberapa metrics evaluasi juga menunjukkan bahwa model berhasil memperoleh nilai akurasi sebesar 98,96%, nilai precision sebesar 98,98%, nilai recall sebesar 98,96%, dan nilai f1-score sebesar 98,95%. Hasil ini didapatkan dengan menggunakan model Random Forest yang menerapkan nilai parameter secara default, di mana parameter

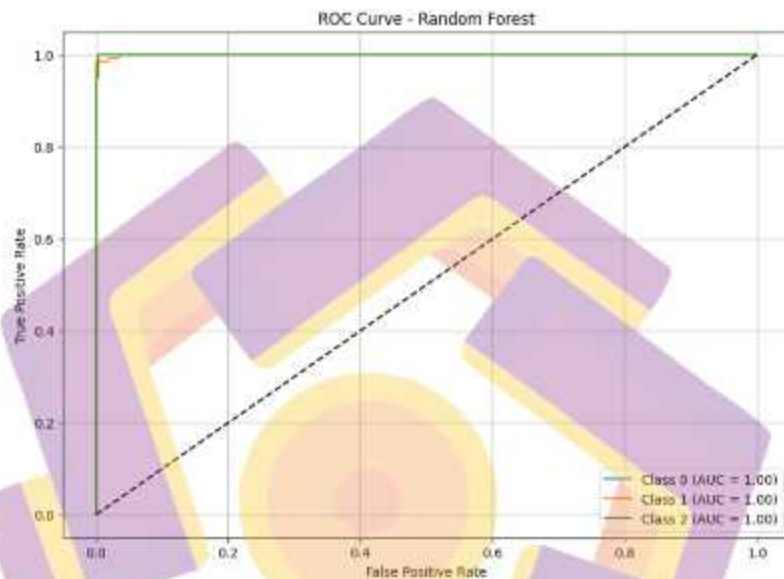
yang digunakan dalam skenario ini mencakup: $n_estimators = 100$, $max_depth = none$, $max_features = sqrt$, $min_samples_leaf = 1$, $min_samples_split = 2$, dan $random_state = 42$.



Gambar 4. 15. Confusion Matrix Skenario 2

Metode ROC-AUC juga diterapkan untuk merangkum tingkat akurasi keseluruhan dari kinerja model. Skor ROC-AUC (Gambar 4.16) memiliki nilai yang berkisar antara 0 hingga 1, di mana nilai 0 menunjukkan bahwa model tidak akurat sama sekali, sedangkan nilai 1 menunjukkan akurasi yang sangat tinggi. Hasil evaluasi dengan metode ini menunjukkan bahwa model mencapai skor 99,978%. Ini menunjukkan bahwa model dapat melakukan klasifikasi dengan tingkat keakuratan yang cukup baik. Temuan ini menunjukkan bahwa model

mampu mengklasifikasikan data yang belum pernah dilihat sebelumnya dengan tingkat keakuratan yang cukup baik.



Gambar 4. 16 ROC-AUC Skenario 2

Berdasarkan hasil eksperimen yang dilakukan pada skenario kedua, dapat diketahui juga bahwa setiap fitur yang digunakan dalam pemodelan memiliki kontribusi yang berbeda dalam mempengaruhi keputusan model, seperti yang terlihat pada Tabel 4.12. Hasil pemodelan menunjukkan bahwa dari total 12 fitur yang diterapkan, empat fitur yang berhubungan dengan aspek akademik menunjukkan pengaruh yang paling signifikan dibandingkan dengan fitur lainnya. Meskipun begitu, fitur-fitur kepribadian dan demografis tetap memberikan kontribusi, meskipun dalam proporsi yang sangat kecil. Berdasarkan hubungan

yang ada, dapat disimpulkan bahwa faktor akademik menunjukkan pengaruh yang paling signifikan dalam pengambilan keputusan.

Tabel 4. 12. Bobot Kepentingan Fitur (Feature Importance) Skenario 2

NO	Nama Fitur	Kepentingan Fitur (Importance)
1	nilai PAS	0.302185
2	nilai harian1	0.267676
3	nilai harian3	0.191878
4	nilai harian2	0.181889
5	bfiO mean	0.014406
6	bfiA mean	0.010380
7	jenis kelamin	0.009839
8	bfiC mean	0.009707
9	penghasilan ayah	0.004106
10	pendidikan ayah	0.003982
11	pendidikan ibu	0.002266
12	penghasilan ibu	0.001686

4.3.3 Skenario 3

Dalam skenario ketiga, pemodelan untuk memprediksi kinerja akademik dilakukan dengan menggunakan metode klasifikasi Random Forest, serta dataset yang digunakan menunjukkan ketidakseimbangan kelas. Dataset dibiarkan dalam kondisi tidak seimbang untuk mengevaluasi dampak ketidakseimbangan tersebut terhadap kinerja model prediksi. Selain itu, pendekatan GridSearchCV juga dilakukan untuk mendapatkan kombinasi nilai *hyperparameter* yang paling optimal dari Random Forest. Dalam prosesnya, GridSearchCV diterapkan bersamaan dengan cross validation untuk menilai stabilitas data pelatihan yang digunakan dalam pengembangan model. Parameter yang diterapkan dalam tahapan ini mencakup *n_estimators*, *max_depth*, *max_features*, *min_samples_leaf*, dan *min_samples_split*. Namun, sebelum melanjutkan ke tahap pelatihan dan prediksi, evaluasi awal dilakukan pada data latih dengan membandingkan kinerja k-fold validation 2 hingga 10 untuk mendapatkan nilai k-fold yang paling optimal dalam

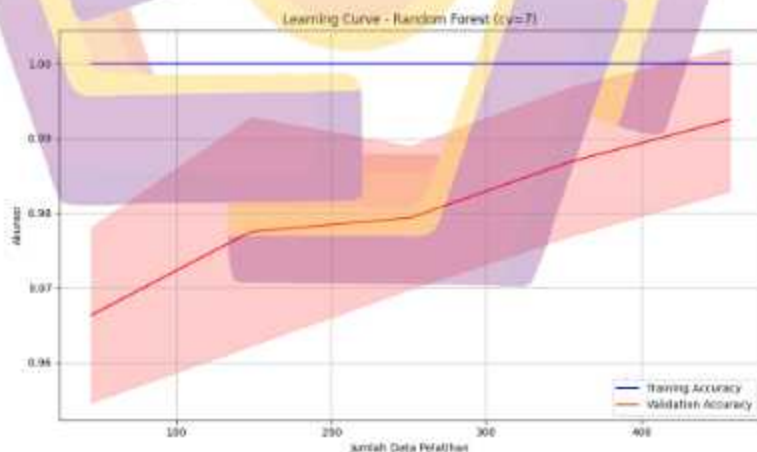
membangun model prediksi. Pengujian ini juga dimaksudkan untuk mencegah terjadinya overfitting, serta mengukur kestabilan data. Hasil pengujian cross validation sendiri dapat dilihat pada Tabel 4.13 .

Tabel 4. 13. Hasil Pengujian dengan K-Fold Cross Validation Skenario 3

K-fold Validation	Average Cross-Validation
2	0,9906
3	0,9906
4	0,9906
5	0,9924
6	0,9906
7	0,9925
8	0,9924
9	0,9906
10	0,9906

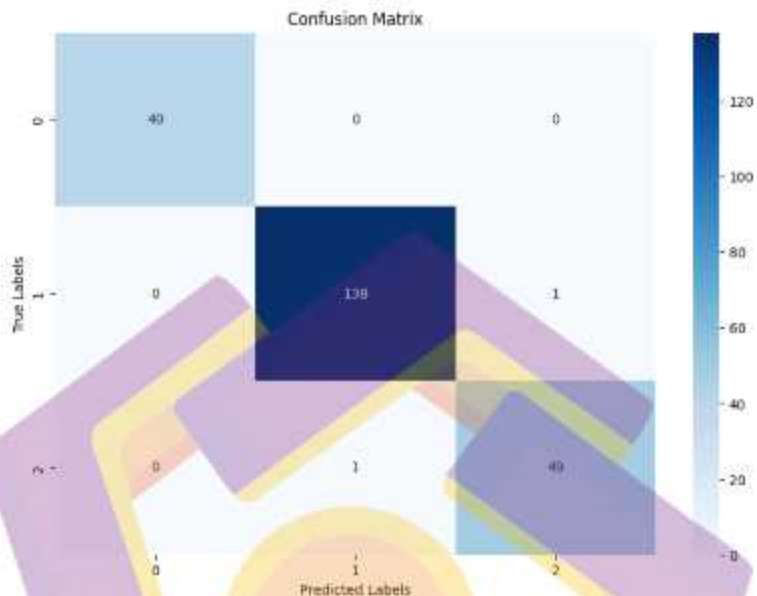
Berdasarkan hasil percobaan, data latih yang dievaluasi menunjukkan performa terbaik ketika diterapkan nilai k-fold = 7, di mana tingkat akurasi yang dicapai sebesar 0,9925 (99,25%). Ini menunjukkan bahwa data yang digunakan berada pada titik yang cukup stabil. Selain itu, meskipun beberapa nilai cross-validation pada berbagai k-fold menunjukkan tingkat akurasi yang serupa, terdapat perbedaan dalam hal kestabilan validasinya. Kestabilan ini tercermin pada grafik learning curve yang ditunjukkan oleh Gambar 4.17. Gambar grafik tersebut menggambarkan proses training score pada model dengan menerapkan k-fold=7. Grafik tersebut menunjukkan perbandingan antara akurasi data pelatihan (kurva biru) dan akurasi validasi (kurva merah) seiring dengan peningkatan jumlah data pelatihan. Secara garis besar grafik menunjukkan terdapat gap kecil antara

akurasi pelatihan dan validasi, ini menandakan bahwa model belajar cukup baik tetapi masih terdapat ketergantungan pada data pelatihan. Kurva biru yang mewakili akurasi pelatihan menunjukkan bahwa akurasi hampir konstan, yaitu nilai berada di kisaran 97-100% meskipun terdapat penurunan dan peningkatan seiring bertambahnya data. Sementara itu, kurva merah yang menggambarkan akurasi validasi menunjukkan kestabilan yang cukup baik, serta terjadi peningkatan secara terus menerus seiring dengan bertambahnya jumlah data pelatihan. Grafik tersebut menunjukkan bahwa nilai akurasi model cukup stabil pada semua jumlah data pelatihan. Secara garis besar, analisis ini menunjukkan bahwa terdapat perbedaan kinerja yang signifikan antara pelatihan dan validasi. Meskipun begitu, secara keseluruhan model menunjukkan kinerja yang cukup baik, serta perlu di garis bawahi bahwa jumlah data pelatihan memiliki pengaruh yang signifikan terhadap performa model.



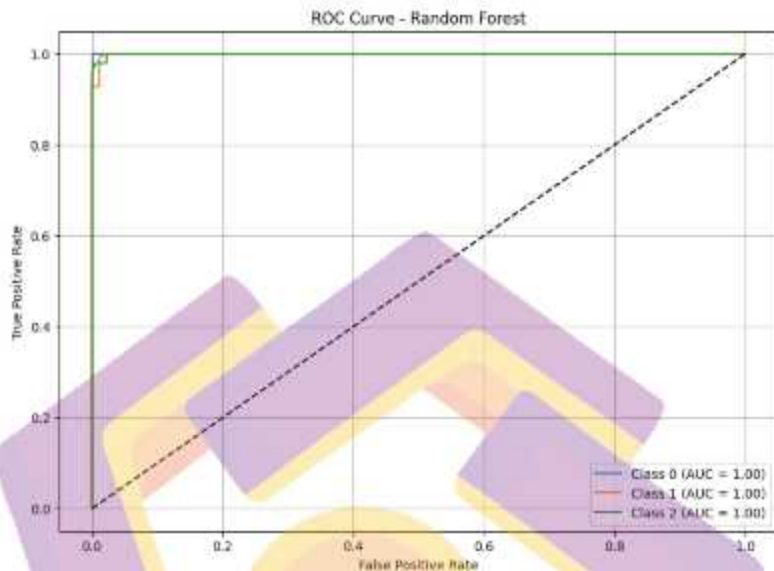
Gambar 4. 17. Grafik Learning Curve k-fold = 7 Skenario 3

Berikutnya, pembahasan mengenai hasil evaluasi yang diperoleh melalui metrics evaluasi menggunakan confusion matrix, sebagaimana ditampilkan dalam Gambar 4.18. Gambar tersebut memberikan ilustrasi mengenai kinerja model dalam mengklasifikasikan tiga kategori kinerja akademik. Pada kelas 0 (cukup), terdapat 40 data yang berhasil diklasifikasikan dengan tepat, Pada kelas 1 (baik), dari 139 data terdapat satu data diklasifikasikan salah menjadi kelas 2 (sangat baik). Sementara itu, untuk kelas 2 (sangat baik), dari 50 data terdapat satu data diklasifikasikan salah menjadi kelas 1 (baik). Selain itu, hasil evaluasi menggunakan beberapa metrics evaluasi juga menunjukkan bahwa model berhasil memperoleh nilai akurasi, precision, recall, dan F1-Score sebesar 99,13%. Hasil ini diperoleh dengan melakukan pencarian hyperparameter yang paling optimal menggunakan GridSearchCV, di mana ditemukan bahwa model prediksi mencapai performa terbaik dengan kombinasi parameter yaitu, `random_state = 42`, `max_depth = 15`, `max_features = sqrt`, `min_samples_leaf = 1`, `min_samples_split = 2`, dan `n_estimators = 500`.



Gambar 4. 18. Confusion Matrix Skenario 3

Metode ROC-AUC juga diterapkan untuk merangkum tingkat akurasi keseluruhan dari kinerja model. Skor ROC-AUC (Gambar 4.19) memiliki nilai yang berkisar antara 0 hingga 1, di mana nilai 0 menunjukkan bahwa model tidak akurat sama sekali, sedangkan nilai 1 menunjukkan akurasi yang sangat tinggi. Hasil evaluasi dengan metode ini menunjukkan bahwa model mencapai skor 99,938%. Ini menunjukkan bahwa model dapat melakukan klasifikasi dengan tingkat keakuratan yang cukup baik.. Temuan ini menunjukkan bahwa model mampu mengklasifikasikan data yang belum pernah dilihat sebelumnya dengan tingkat keakuratan yang cukup baik, meskipun terdapat ketidakseimbangan dalam kelas pada dataset.



Gambar 4. 19. ROC-AUC Skenario 3

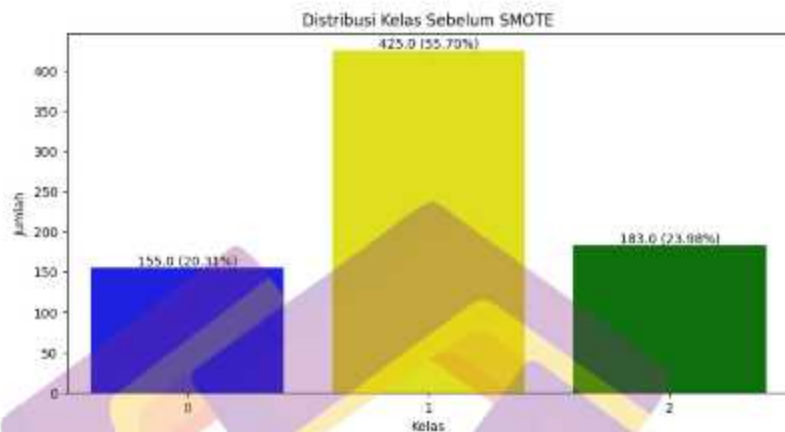
Berdasarkan hasil eksperimen yang dilakukan pada skenario ketiga, dapat diketahui juga bahwa setiap fitur yang digunakan dalam pemodelan memiliki kontribusi yang berbeda dalam mempengaruhi keputusan model, seperti yang terlihat pada Tabel 4.14 Hasil pemodelan menunjukkan bahwa dari total 12 fitur yang diterapkan, empat fitur yang berhubungan dengan aspek akademik menunjukkan pengaruh yang paling signifikan dibandingkan dengan fitur lainnya. Meskipun begitu, fitur-fitur kepribadian dan demografis tetap memberikan kontribusi, meskipun dalam proporsi yang sangat kecil. Berdasarkan hubungan yang ada, dapat disimpulkan bahwa faktor akademik menunjukkan pengaruh yang paling signifikan dalam pengambilan keputusan.

Tabel 4. 14. Bobot Kepentingan Fitur (Feature Importance) Skenario 3

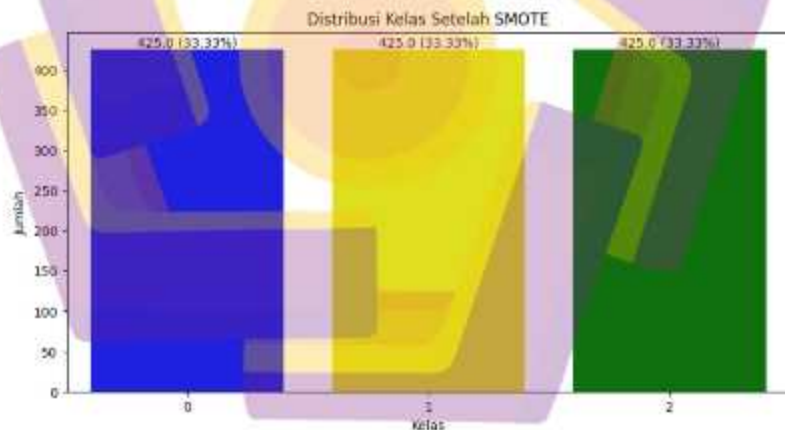
Nama Fitur	Kepentingan Fitur (Importance)
nilai PAS	0.319390
nilai harian1	0.264120
nilai harian3	0.222796
nilai harian2	0.142516
bfiO mean	0.010548
bfiA mean	0.010428
bfiC mean	0.009157
jenis kelamin	0.005571
penghasilan ayah	0.004703
pendidikan ibu	0.004248
pendidikan ayah	0.003970
penghasilan ibu	0.002553

4.3.4 Skenario 4

Dalam skenario keempat, pemodelan untuk memprediksi kinerja akademik dilakukan dengan menggunakan metode klasifikasi Random Forest, serta menerapkan teknik SMOTE untuk menyelesaikan permasalahan kelas tidak seimbang. Penerapan SMOTE ini dilakukan dengan cara tidak menghapus data yang ada, melainkan membuat data sintesis baru dari contoh data yang sudah ada pada data latih minoritas. Namun, perlu diketahui bahwa sebelum evaluasi dilakukan, teknik SMOTE akan diterapkan terlebih dahulu pada dataset. Seperti yang terlihat pada Gambar 4.20 dan Gambar 4.21, di mana penerapan SMOTE dilakukan pada seluruh dataset, yaitu sebelum data dibagi menjadi data latih dan data uji. Hasil dari proses ini menunjukkan bahwa jumlah data pada kelas minoritas telah disintesis hingga seimbang dengan jumlah data pada kelas mayoritas, sehingga distribusi kelas menjadi lebih proporsional.



Gambar 4. 20. Distribusi Kelas Sebelum SMOTE Skenario 4



Gambar 4. 21. Distribusi Kelas Sesudah SMOTE Skenario 4

Selain itu, pendekatan GridSearchCV juga dilakukan untuk mendapatkan kombinasi nilai *hyperparameter* yang paling optimal dari Random Forest. Dalam prosesnya, GridSearchCV diterapkan bersamaan dengan cross validation untuk

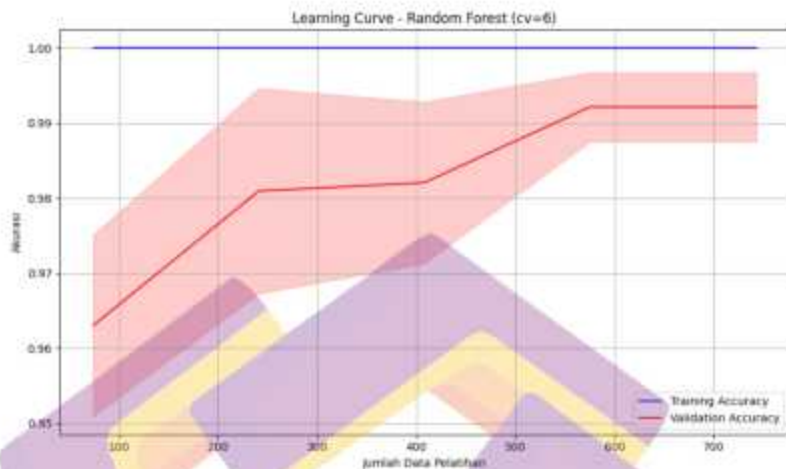
menilai stabilitas data pelatihan yang digunakan dalam pengembangan model. Parameter yang diterapkan dalam tahapan ini mencakup `n_estimators`, `max_depth`, `max_features`, `min_samples_leaf`, dan `min_samples_split`. Namun, sebelum melanjutkan ke tahap pelatihan dan prediksi, evaluasi awal dilakukan pada data latih dengan membandingkan kinerja `k-fold validation` 2 hingga 10 untuk mendapatkan nilai `k-fold` yang paling optimal dalam membangun model prediksi. Pengujian ini juga dimaksudkan untuk mencegah terjadinya `overfitting`, serta mengukur kestabilan data. Hasil pengujian `cross validation` sendiri dapat dilihat pada Tabel 4.15.

Tabel 4. 15. Hasil Pengujian dengan K-Fold Cross Validation Skenario 4

K-fold Validation	Average Cross-Validation
2	0.9922
3	0.9922
4	0.9899
5	0.9899
6	0.9922
7	0.9922
8	0.9910
9	0.9921
10	0.9921

Berdasarkan hasil percobaan, data latih yang dievaluasi menunjukkan performa terbaik ketika diterapkan nilai `k-fold` = 6, di mana tingkat akurasi yang dicapai sebesar 0,9922 (99,22%). Ini menunjukkan bahwa data yang digunakan berada pada titik yang cukup stabil. Kestabilan ini tercermin pada grafik `learning curve` yang ditunjukkan oleh Gambar 4.22. Gambar grafik tersebut menggambarkan

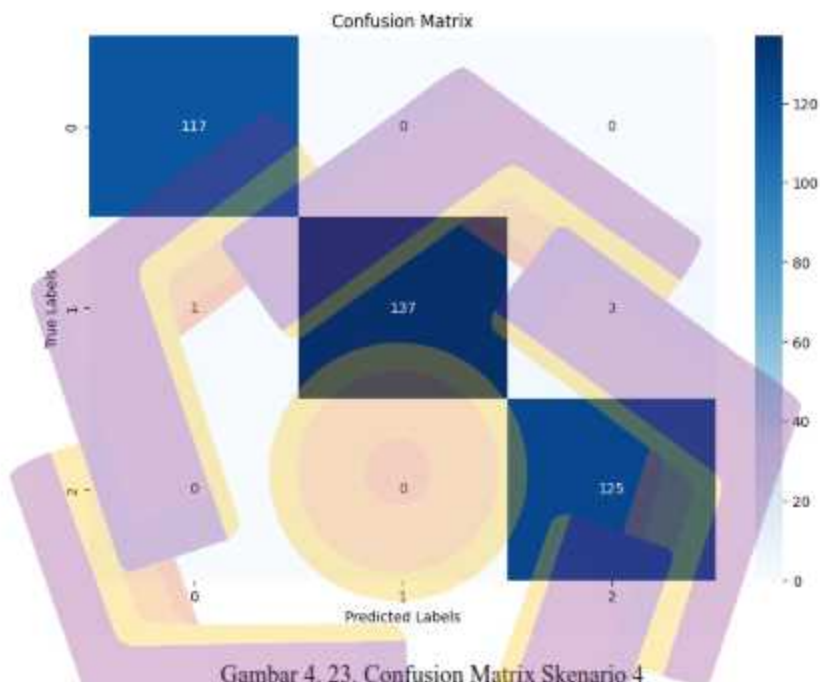
proses training score pada model dengan menerapkan $k\text{-fold} = 6$. Grafik tersebut menunjukkan perbandingan antara akurasi data pelatihan (kurva biru) dan akurasi validasi (kurva merah) seiring dengan peningkatan jumlah data pelatihan. Kurva biru yang mewakili akurasi pelatihan menunjukkan bahwa akurasi konsisten berada di angka 100% untuk seluruh jumlah data pelatihan. Ini menunjukkan bahwa model memiliki kemampuan yang sangat baik dalam mengingat data pelatihan, sehingga ada kemungkinan model mengalami overfit pada data pelatihan. Sementara itu, kurva merah yang menggambarkan akurasi validasi menunjukkan kestabilan yang cukup baik, di mana terjadi peningkatan seiring dengan bertambahnya jumlah data pelatihan, namun terdapat penurunan performa pada pertengahan proses validasi. Meskipun begitu, secara keseluruhan grafik tersebut menunjukkan bahwa nilai akurasi model cukup stabil di atas 0,98 (98%) pada semua jumlah data pelatihan. Secara garis besar, analisis ini menunjukkan bahwa terdapat perbedaan kinerja yang signifikan antara pelatihan dan validasi. Meskipun begitu, secara keseluruhan model menunjukkan kinerja yang cukup baik, serta perlu di garis bawah bahwa jumlah data pelatihan memiliki pengaruh yang signifikan terhadap performa model.



Gambar 4. 22. Grafik Learning Curve k-fold = 6 Skenario 4

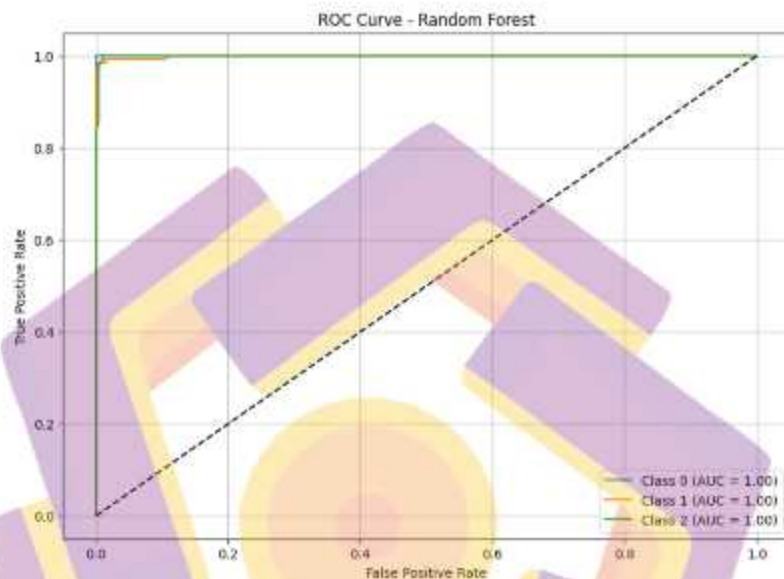
Berikutnya, pembahasan mengenai hasil evaluasi yang diperoleh melalui metrics evaluasi menggunakan confusion matrix, sebagaimana ditampilkan dalam Gambar 4.23. Gambar tersebut memberikan ilustrasi mengenai kinerja model dalam mengklasifikasikan tiga kategori kinerja akademik. Pada kelas 0 (cukup) dan kelas 2 (sangat baik), model menunjukkan kinerja yang sangat memuaskan dengan semua data berhasil diklasifikasikan dengan benar. Sementara itu, untuk kelas 1 (baik) terdapat 137 data diklasifikasikan dengan tepat, dan terdapat 4 data yang salah diklasifikasikan, 3 data salah dikategorikan sebagai kelas 2 (baik sekali) dan 1 data salah dikategorikan sebagai kelas 0 (cukup). Selain itu, hasil evaluasi menggunakan beberapa metrics evaluasi juga menunjukkan bahwa model berhasil memperoleh nilai akurasi sebesar 98,96%, nilai precision sebesar 99,23%, nilai recall sebesar 98,96%, dan nilai f1-score sebesar 98,95%. Hasil ini didapatkan dengan melakukan pencarian hyperparameter yang paling optimal menggunakan GridSearchCV, di

mana ditemukan bahwa model prediksi mencapai performa terbaik dengan kombinasi parameter yaitu, `random_state = 42`, `max_depth = 15`, `max_features = sqrt`, `min_samples_leaf = 1`, `min_samples_split = 2`, dan `n_estimators = 250`.



Metode ROC-AUC juga diterapkan untuk merangkul tingkat akurasi keseluruhan dari kinerja model. Skor ROC-AUC (Gambar 4.24) memiliki nilai yang berkisar antara 0 hingga 1, di mana nilai 0 menunjukkan bahwa model tidak akurat sama sekali, sedangkan nilai 1 menunjukkan akurasi yang sangat tinggi. Hasil evaluasi dengan metode ini menunjukkan bahwa model mencapai skor 99,943%. Ini menunjukkan bahwa model dapat melakukan klasifikasi dengan tingkat keakuratan yang cukup baik.. Temuan ini menunjukkan bahwa model

mampu mengklasifikasikan data yang belum pernah dilihat sebelumnya dengan tingkat keakuratan yang cukup baik.



Gambar 4. 24. ROC-AUC Skenario 4

Berdasarkan hasil eksperimen yang dilakukan pada skenario keempat, dapat diketahui juga bahwa setiap fitur yang digunakan dalam pemodelan memiliki kontribusi yang berbeda dalam mempengaruhi keputusan model, seperti yang terlihat pada Tabel 4.16. Hasil pemodelan menunjukkan bahwa dari total 12 fitur yang diterapkan, empat fitur yang berhubungan dengan aspek akademik menunjukkan pengaruh yang paling signifikan dibandingkan dengan fitur lainnya. Meskipun begitu, fitur-fitur kepribadian dan demografis tetap memberikan kontribusi, meskipun dalam proporsi yang sangat kecil. Berdasarkan hubungan

yang ada, dapat disimpulkan bahwa faktor akademik menunjukkan pengaruh yang paling signifikan dalam pengambilan keputusan.

Tabel 4. 16. Bobot Kepentingan Fitur (Feature Importance) Skenario 4

Nama Fitur	Kepentingan Fitur (Importance)
nilai PAS	0.458638
nilai harian1	0.331166
nilai harian3	0.171683
nilai harian2	0.026748
bfiC mean	0.003503
bfiA mean	0.002498
bfiO mean	0.001942
pendidikan ayah	0.001865
penghasilan ayah	0.000808
pendidikan ibu	0.000729
penghasilan ibu	0.000333
jenis kelamin	0.000088

4.4. Analisis Perbandingan

Model prediksi kinerja akademik menggunakan algoritma Random Forest dikembangkan melalui perbandingan hasil dari empat skenario pengujian yang berbeda, dengan memanfaatkan dataset yang sama tetapi dalam kondisi yang berbeda. Ini berarti terdapat model dengan dataset yang tidak seimbang dan model dengan dataset yang telah diseimbangkan. Secara umum, data dibagi menjadi 70% untuk proses pelatihan dan 30% untuk pengujian. Eksperimen yang dilakukan berhasil menjawab pertanyaan penelitian yang telah dirumuskan sebelumnya, mencakup analisis hubungan antara berbagai fitur dengan fitur target (nilai akhir atau kinerja akademik siswa), evaluasi dampak penggunaan teknik SMOTE terhadap kinerja model Random Forest, serta kinerja model Random Forest secara keseluruhan dalam memprediksi nilai akhir siswa berdasarkan metrik evaluasi.

Berdasarkan keseluruhan hasil uji coba, seperti yang disajikan dalam Tabel 4.10, Tabel 4.12, Tabel 4.14, dan Tabel 4.16, ditemukan bahwa faktor akademik bertindak sebagai indikator utama dalam pengambilan keputusan terkait kinerja akademik. Setiap fitur akademik yang dianalisis menunjukkan tingkat kepentingan yang konsisten di berbagai skenario pengujian, di mana nilai_PAS memiliki bobot tertinggi, diikuti oleh nilai_harian1, nilai_harian3, dan nilai_harian2. Hasil ini mengindikasikan bahwa setiap fitur memiliki hubungan yang stabil dan signifikan dengan kinerja akademik. Temuan ini sejalan dengan penelitian yang dilakukan oleh (Nachouki et al., 2023), yang menunjukkan adanya hubungan positif antara nilai akademik di jenjang sebelumnya dan kinerja akademik mahasiswa dalam suatu mata kuliah. Penelitian lain yang dilakukan oleh (El-Keiey et al., 2022; Khairy et al., 2024; Roslan & Chen, 2023) juga mengungkapkan hal serupa, di mana nilai akademik sebelumnya terbukti memiliki pengaruh secara signifikan terhadap prediksi kinerja akademik siswa. Oleh karena itu, dalam konteks prediksi kinerja akademik dalam penelitian ini, dapat disimpulkan bahwa terdapat hubungan signifikan antara nilai akademik siswa di masa lalu dan kinerja akademik mereka di masa yang akan datang. Meskipun nilai akademik menjadi indikator utama dalam memprediksi kinerja akademik, faktor demografi dan kepribadian tetap memiliki peran dalam pengambilan keputusan, meskipun pengaruhnya relatif sangat kecil. Hasil penelitian ini menunjukkan bahwa di antara faktor-faktor tersebut, dapat disimpulkan bahwa kelima dimensi kepribadian dalam model big five memiliki tingkat kepentingan yang konsisten dengan tingkat pengaruh yang bervariasi. Dari lima dimensi, hanya tiga dimensi yang memiliki tingkat kepentingan serupa

terhadap model prediksi kinerja akademik. Temuan ini menunjukkan bahwa dimensi *agreeableness*, *openness*, dan *conscientiousness* memiliki pengaruh positif yang paling signifikan terhadap prestasi akademik. Penelitian (Meyer et al., 2023) menjelaskan bahwa dimensi *conscientiousness* dapat memberikan dampak positif melalui keteraturan, ketekunan, dan kerja keras yang pada gilirannya sangat penting untuk mencapai keberhasilan akademik. Dimensi *openness* juga berkontribusi secara positif, yang dicerminkan melalui keinginan dan keterbukaan individu untuk mengeksplorasi hal-hal baru di luar kebiasaan, yang pada gilirannya dapat mendukung proses pembelajaran. *Agreeableness* juga ditemukan berkontribusi dalam menciptakan hubungan sosial yang harmonis di kelas, yang berpengaruh pada penilaian guru serta menciptakan lingkungan belajar yang kondusif. Temuan ini juga didukung oleh berbagai penelitian yang menjelaskan bahwa dimensi *conscientiousness*, *agreeableness*, dan *openness* (El-Keiey et al., 2022; Meyer et al., 2023; Shaninah & Mohd Noor, 2024) ditemukan sebagai dimensi dari kepribadian Big Five yang memiliki pengaruh terhadap kinerja akademik siswa. Sementara itu, temuan juga menunjukkan bahwa dimensi *extraversion* dan *neuroticism* memiliki pengaruh yang relatif kecil dan tidak signifikan terhadap prestasi akademik, meskipun arah pengaruhnya tetap positif. Hal ini menunjukkan bahwa meskipun siswa extrovert cenderung aktif dalam interaksi sosial (El-Keiey et al., 2022; Meyer et al., 2023), dampaknya tidak terlalu signifikan terhadap kinerja akademik. Kemudian, ketidakstabilan emosional juga tidak memiliki pengaruh signifikan terhadap prestasi akademik (El-Keiey et al., 2022; Meyer et al., 2023). Meskipun faktor demografi juga ditemukan memiliki bobot kepentingan yang relatif sangat

kecil, faktor sosial-ekonomi, seperti pendidikan dan penghasilan orang tua, sebagai fitur-fitur yang memiliki bobot kepentingan tertinggi di antara fitur demografi lainnya. Sedangkan, fitur jenis kelamin ditemukan sebagai fitur yang memiliki bobot kepentingan paling rendah di antara semua skenario. Temuan ini didukung oleh beberapa penelitian (El-Keiey et al., 2022; Nachouki et al., 2023; Roslan & Chen, 2023), yang menyoroti bahwa faktor akademik ditemukan menjadi fitur utama, sedangkan fitur demografis menjadi fitur pendukung dalam menentukan kinerja akademik siswa.

Berdasarkan hasil evaluasi empat skenario utama menggunakan berbagai metrik evaluasi (Tabel 4.17), model prediksi kinerja akademik menggunakan Random Forest telah menunjukkan kinerja yang cukup baik, dengan tingkat akurasi berkisar antara 98% hingga 99%. Meskipun begitu, keempat skenario yang diuji menunjukkan variasi dalam tingkat stabilitasnya.

Tabel 4. 17. Evaluasi Model Prediksi Kinerja Akademik

Evaluasi	Random Forest	Random Forest + SMOTE	Random Forest + GridSearchCV	Random Forest + GridSearchCV + SMOTE
Cross Validation %	99,06 (k=7)	99,22 (k=7)	99,25 (k=7)	99,22 (k=6)
Training Score %	100	100	100	100
Accuracy %	98,69	98,96	99,13	98,96
Precision %	98,70	98,98	99,13	99,23
Recall %	98,69	98,96	99,13	98,96
F1-Score %	98,69	98,95	99,13	98,96
ROC-AU %	99,927	99,978	99,938	99,943

Hasil evaluasi keempat skenario menunjukkan bahwa nilai cross validation bervariasi sesuai dengan nilai k-fold yang digunakan. Dari keempat skenario,

scenario ketiga menghasilkan nilai terbaik sebesar 99,25% pada $k_fold=7$, disusul oleh scenario dua dan scenario empat yang sama-sama mencapai nilai 99,22% dengan nilai k_fold berbeda, di mana scenario dua memiliki nilai $k_fold=7$ dan scenario keempat memiliki nilai $k_fold=6$. Sementara itu, scenario satu memperoleh hasil terendah dengan 99,06% ($k_fold=7$). Namun, analisis learning curve memberikan temuan yang berbeda. Skenario kedua dan keempat menunjukkan kestabilan performa yang lebih baik dibandingkan dengan scenario kesatu dan ketiga, dengan nilai minimum sebesar 98%. Sedangkan, scenario kesatu dan ketiga mencapai nilai minimum sebesar 97%. Namun, jika membandingkan lagi scenario kedua dan keempat, ditemukan bahwa scenario keempat terbukti lebih stabil, karena performanya mendekati 100% dan selisih antara training score dan validation score relatif kecil. Evaluasi berbagai metrics juga mengkonfirmasi perbedaan performa antar skenario. Hasil menunjukkan bahwa skenario ketiga unggul dalam hal akurasi (99%) dibandingkan dengan skenario kesatu, kedua dan keempat. Sementara itu, skenario kedua dan keempat memiliki akurasi yang hamper serupa, tetapi skenario keempat lebih menonjol dengan nilai precision tertinggi sebesar 99,23%. Sebaliknya scenario kesatu konsisten menjadi model dengan performa terendah di antara keempat skenario. Hasil evaluasi menggunakan ROC-AUC juga menunjukkan bahwa, skenario kedua dan keempat memperlihatkan performa yang lebih baik dibandingkan dengan skenario kesatu dan ketiga. Analisis confusion matrix juga menunjukkan bahwa skenario kedua dan keempat memiliki tingkat kesalahan yang serupa, sedangkan skenario ketiga memiliki tingkat kesalahan yang lebih rendah dibandingkan skenario kesatu. Secara

keseluruhan, meskipun skenario ketiga menunjukkan tingkat performa yang cukup tinggi, kestabilan performa dan keseimbangan hasil evaluasi pada skenario keempat menunjukkan model dengan penerapan teknik SMOTE pada data tidak seimbang dan tuning membuat model menjadi lebih optimal dibandingkan dengan skenario lainnya.

Temuan ini menunjukkan bahwa penelitian ini telah berhasil mengembangkan model prediksi dengan tingkat performa yang lebih baik dibandingkan dengan beberapa studi sebelumnya. Hasil penelitian oleh (Nachouki et al., 2023) yang menerapkan skenario dataset yang tidak seimbang dan penerapan metode Random Forest secara *default*, mengungkapkan bahwa model prediksi Random Forest yang dikembangkan berhasil mencapai akurasi sebesar 90,33% dalam mendeteksi mahasiswa yang berisiko dikeluarkan. Penelitian lain oleh (Khairy et al., 2024) juga menemukan bahwa pemodelan prediksi kinerja terhadap hasil ujian yang hanya mempertimbangkan faktor akademik dapat mencapai akurasi hingga 98,70%. Selain itu, penelitian (Santoso et al., 2024) juga menemukan bahwa model Random Forest yang dikembangkan dengan fokus pada faktor demografis hanya mampu mencapai akurasi sebesar 76% dalam memprediksi nilai akhir (GPA) mahasiswa. Penelitian oleh (El-Keiey et al., 2022) juga menunjukkan bahwa meskipun berbagai fitur yang berkaitan dengan kinerja akademik, termasuk faktor demografis, akademik, dan kepribadian Big Five telah diterapkan, hasil yang diperoleh menunjukkan bahwa dalam kondisi data yang tidak seimbang, hanya mampu mencapai akurasi sebesar 83%. Namun, penelitian tersebut juga menyoroti bahwa tanpa *hyperparameter tuning*, model Random Forest tidak mampu mencapai

kinerja yang paling optimal. Oleh karena itu, penerapan *hyperparameter tuning* diperlukan untuk meningkatkan keakuratan model secara optimal.

Berdasarkan hasil uji coba penerapan *hyperparameter tuning* menggunakan GridSearchCV, proses *tuning* ini mampu meningkatkan performa optimal pada skenario ketiga. Skenario ketiga ini mengalami peningkatan ke angka 99,25% dengan nilai $k_fold=7$, melalui pengaturan parameter dengan nilai $random_state = 42$, $max_depth = 15$, $max_features = sqrt$, $min_samples_leaf = 1$, $min_samples_split = 2$, $n_estimators = 500$. Sementara itu, untuk skenario keempat sendiri performa model juga memiliki peningkatan meskipun tidak seoptimal skenario ketiga, di mana peningkatan signifikan hanya terlihat dalam indikator precision. Hasil ini diperoleh melalui pengaturan parameter dengan nilai $random_state = 42$, $max_depth = 15$, $max_features = sqrt$, $min_samples_leaf = 1$, $min_samples_split = 2$, $n_estimators = 250$. Berdasarkan temuan tersebut, terungkap bahwa penerapan *tuning* melalui GridSearchCV mampu meningkatkan performa model, khususnya pada skenario ketiga. Meskipun begitu, jika memperhatikan metrik evaluasi lainnya, hasil evaluasi tersebut menunjukkan bahwa meskipun tingkat akurasi skenario ketiga meningkat, ketidakstabilan performa masih terlihat jelas. Ini menunjukkan bahwa optimasi yang tidak disertai dengan penanganan ketidakseimbangan kelas mengakibatkan model lebih cenderung fokus pada kelas mayoritas, sehingga dapat mengurangi kemampuan generalisasi model secara keseluruhan. Temuan ini juga sejalan dengan sejumlah penelitian sebelumnya (Khairy et al., 2024; Nachouki et al., 2023; Santoso et al., 2024) yang menunjukkan bahwa ketidakseimbangan data memiliki pengaruh yang

signifikan terhadap stabilitas model. Karena, distribusi data yang tidak seimbang di setiap kategori dapat mengakibatkan model mengalami bias algoritmik. Sehingga, meskipun Random Forest dapat mencapai akurasi yang cukup baik, ketidakseimbangan data dapat mempengaruhi keandalan model prediksi. Oleh karena itu, penanganan ketidakseimbangan data sangat penting untuk memastikan generalisasi hasil model.

Kemudian, pada skenario kedua dan keempat penerapan metode *Synthetic Minority Over-sampling Technique* (SMOTE) dalam model prediksi menunjukkan konsistensi yang baik dalam performa model. Seperti yang dapat dilihat dalam Gambar 4.12 dan Gambar 4.20, data menunjukkan ketidakseimbangan dengan dominasi kelas 1 sebesar 55,70%. Sedangkan dalam Gambar 4.13 dan Gambar 4.21, setelah penerapan SMOTE, distribusi data menjadi seimbang dengan masing-masing kelas memiliki proporsi 33,33%. SMOTE menerapkan pendekatan *k-nearest neighbors* dengan menggunakan jarak Euclidean untuk menghasilkan data sintesis dari contoh-contoh minoritas terdekat (Chawla et al., 2002), sehingga dapat merepresentasikan karakteristik kelas minoritas tersebut. Hasil evaluasi skenario kedua telah menunjukkan bahwa penerapan SMOTE dalam Random Forest default berhasil meningkatkan performa di semua aspek metrics evaluasi. Peningkatan ini juga terlihat dari tingkat kestabilan model yang lebih baik dibandingkan dengan skenario pertama. Namun demikian, skenario keempat memberikan hasil yang lebih baik, di mana kombinasi SMOTE dan tuning tidak hanya meningkatkan nilai precision, tetapi juga menghasilkan kestabilan performa yang lebih baik dibandingkan skenario lainnya. Meskipun begitu, jika dibandingkan dengan

skenario ketiga yang berfokus pada tuning, akurasi skenario keempat masih berada dibawahnya. Hal ini terjadi karena pada dasarnya SMOTE memiliki risiko menimbulkan overfitting, noise, dan bias, karena dapat menghasilkan data yang terlalu mirip dan memaksa keseimbangan tanpa mempertimbangkan distribusi alami yang ada (Abdul Bujang et al., 2023; Ghaniaviyanto Ramadhan & Adiwijaya, 2022). Sehingga, penerapan tuning saja belum mampu meningkatkan tingkat akurasi model ke nilai yang optimal. Meskipun begitu, temuan ini telah berhasil mengembangkan model prediksi dengan tingkat performa yang lebih baik dibandingkan dengan penelitian sebelumnya dengan skenario serupa. Salah satu penelitian (Ghaniaviyanto Ramadhan & Adiwijaya, 2022), yang berfokus pada analisis pelanggan ideal berdasarkan kepribadian dengan memanfaatkan data riwayat transaksi, metode Random Forest dan SMOTE mampu menghasilkan model prediksi dengan akurasi 88%, precision 79%, dan recall 70%. Penelitian serupa juga dilakukan oleh (Hamoud et al., 2022), yang melakukan uji coba pada dua dataset berbeda dengan karakteristik data yang tidak seimbang dan jumlah data yang sangat terbatas. Penelitian tersebut model prediksi kinerja siswa berdasarkan nilai akhir suatu mata pelajaran, berhasil menghasilkan model dengan akurasi masing-masing sebesar 81,6% dan 83%. Sehingga pada dasarnya penerapan SMOTE dalam model yang telah dioptimasi melalui *tuning hyperparameter* GridSearchCV berhasil meningkatkan performa model, khususnya nilai precision secara signifikan. Temuan ini juga sejalan dengan hasil penelitian sebelumnya (Pecuchova & Drlik, 2023), bahkan menunjukkan performa yang lebih baik. Dalam penelitian tersebut, model prediksi kinerja akademik yang hanya menekankan pada

faktor-faktor akademik menggunakan dataset yang relatif kecil, hanya berhasil mencapai akurasi di angka 94,5%.

Berdasarkan temuan-temuan tersebut, dapat diketahui bahwa ketidakseimbangan data memiliki peran yang tidak terlalu baik terhadap kinerja model. Penelitian ini juga menemukan bahwa penurunan kinerja pada beberapa metrik evaluasi setelah proses *tuning*, dapat disebabkan oleh perlakuan yang sama terhadap semua algoritma tanpa memperhatikan karakteristik unik masing-masing algoritma. Artinya, *hyperparameter tuning* merupakan proses yang sangat sensitif dan tidak dapat dilakukan secara umum atau dibandingkan secara langsung tanpa adanya penyesuaian yang lebih spesifik dan mendetail (Arden & Safitri, 2022). Meskipun begitu, secara khusus metode Random Forest yang memiliki lebih banyak *hyperparameter* dibandingkan metode lainnya, terbukti memberikan kontribusi signifikan terhadap kinerja model lainnya (Pecuchova & Drlik, 2023). Hal ini dapat dilihat pada skenario 4 yang menunjukkan hasil performa prediksi yang lebih baik dibandingkan tiga skenario lainnya. Namun, hal ini juga terjadi karena kinerja model tidak hanya dipengaruhi oleh penerapan metode *hyperparameter tuning*, tetapi juga melalui penyelesaian masalah data tidak seimbang menggunakan teknik SMOTE. Temuan ini menyoroti bahwa kombinasi antara Random Forest, SMOTE, dan *hyperparameter tuning* menggunakan GridSearchCV telah menghasilkan kinerja paling baik dan stabil. Hal ini mengindikasikan bahwa diantara skenario lainnya, model skenario keempat merupakan model prediksi kinerja akademik yang paling stabil, dengan tingkat akurasi secara keseluruhan berada di tingkat yang cukup konsisten. Sehingga,

penerapan *hyperparameter tuning* pada data tidak seimbang, tidak menjamin tercapainya performa dan kestabilan model yang optimal. Sementara itu, penerapan teknik SMOTE dalam dataset tidak seimbang terbukti dapat meningkatkan performa model prediksi, meskipun hasilnya belum optimal. Sehingga, penerapan SMOTE dalam dataset tidak seimbang dapat dioptimalkan dengan menerapkan *hyperparameter tuning*. Temuan ini juga diperkuat oleh penelitian yang dilakukan oleh (Ghaniaviyanto Ramadhan & Adiwijaya, 2022), yang menyatakan bahwa penggunaan SMOTE pada model Random Forest untuk data tidak seimbang tidak secara otomatis meningkatkan performa model secara signifikan. Oleh karena itu, penerapan teknik SMOTE sangat penting untuk meningkatkan kemampuan generalisasi model, terutama ketika dikombinasikan dengan metode *tuning* yang berguna untuk mengoptimalkan pemilihan *hyperparameter* pada algoritma Random Forest (Pecuchova & Drlik, 2023). Pada dasarnya proses *tuning* bisa menjadi elemen penting dalam mencapai keseimbangan performa model, karena tanpa penyesuaian *hyperparameter* yang tepat, model Random Forest tidak dapat mencapai kinerja optimalnya (Arden & Safitri, 2022; Rimal et al., 2024).

Secara keseluruhan, penelitian ini berhasil menunjukkan bahwa metode Random Forest dapat digunakan untuk mengembangkan model prediksi kinerja akademik dengan tingkat akurasi yang stabil. Penelitian ini juga menekankan bahwa di samping penggunaan metode statistik deskriptif (Shaninah & Mohd Noor, 2024), teknik data mining seperti algoritma Random Forest dapat menghasilkan model prediksi dengan akurasi yang baik. Namun, penelitian ini juga menyoroti pentingnya penggunaan teknik SMOTE dan *hyperparameter tuning* pada Random

Forest untuk mencapai performa yang optimal. Temuan menunjukkan bahwa penelitian ini memberikan kontribusi baru terhadap beberapa studi sebelumnya yang berkaitan dengan tingkat performa model Random Forest dalam konteks prediksi (El-Keiey et al., 2022; Khairy et al., 2024; Roslan & Chen, 2023; Santoso et al., 2024). Penelitian ini juga berhasil menemukan solusi terhadap masalah ketidakseimbangan data yang telah diidentifikasi dalam beberapa studi sebelumnya (Khairy et al., 2024; Nachouki et al., 2023; Santoso et al., 2024). Hasil penelitian ini menunjukkan bahwa penerapan teknik SMOTE memberikan efek positif terhadap stabilitas kinerja model di semua metrik evaluasi. Meskipun, hasil evaluasi juga menunjukkan bahwa pada metrik tertentu penerapan tuning pada model SMOTE dan Random Forest, tidak menyebabkan tingkat akurasi meningkat. Bahkan, tingkat kesalahannya pun tidak jauh berbeda dengan model yang tidak menerapkan SMOTE ataupun tuning. Hal ini penting untuk diperhatikan karena pada dasarnya, SMOTE menciptakan titik-titik sintesis dengan menggabungkan data minoritas asli dengan data tetangganya. Dengan kata lain, data baru yang dihasilkan adalah hasil kombinasi dari data yang sudah ada, sehingga tidak memberikan informasi baru yang signifikan (Abdul Bujang et al., 2023). Selain itu, teknik *oversampling* juga memaksa jumlah data pada kelas minoritas agar setara dengan kelas mayoritas, tanpa mempertimbangkan distribusi alaminya (Ghaniaviyanto Ramadhan & Adiwijaya, 2022). Akibatnya, hal ini dapat mengurangi kemampuan generalisasi model serta membuat model kurang peka terhadap karakteristik sejati dari kelas minoritas, yang pada akhirnya dapat menyebabkan bias dan meningkatkan risiko kesalahan dalam prediksi. Selanjutnya,

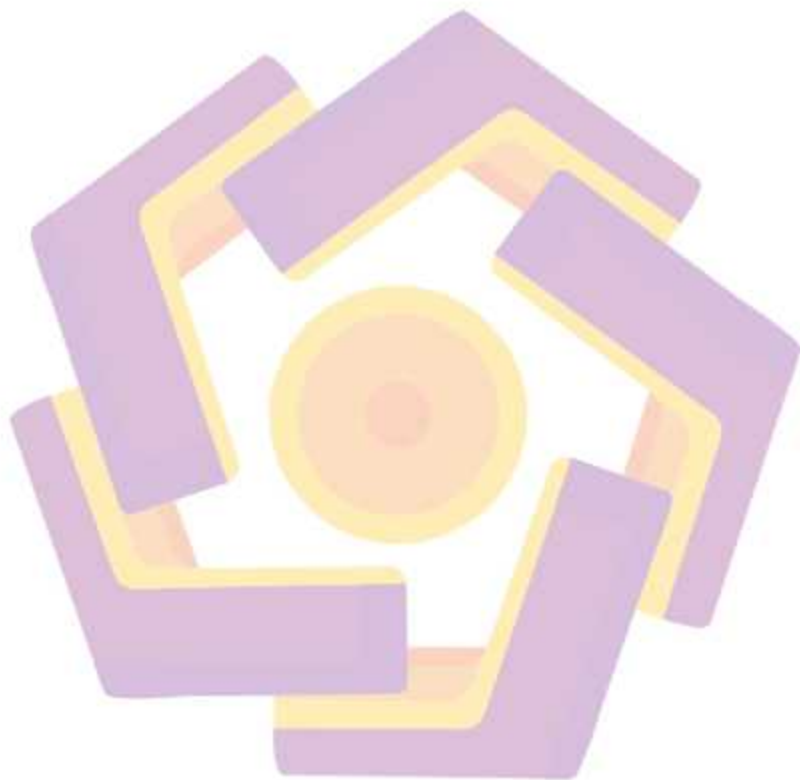
penelitian ini juga menemukan bahwa selain faktor akademik yang telah banyak dibahas dalam *literature* sebelumnya (El-Keiey et al., 2022; Khairy et al., 2024; Nachouki et al., 2023; Roslan & Chen, 2023), terdapat faktor lain yang mempengaruhi kinerja akademik. Hasil eksperimen ini menunjukkan bahwa faktor kepribadian berdasarkan model Big Five serta faktor demografi seperti jenis kelamin, tingkat pendidikan dan penghasilan orang tua memiliki dampak yang signifikan. Hal ini terlihat adanya perubahan kinerja model yang tidak konsisten ketika salah satu fitur tersebut dihilangkan dalam berbagai skenario, meskipun pengaruhnya tergolong kecil.

Selain itu, temuan dari penelitian ini juga menunjukkan bahwa upaya mengembangkan model prediksi kinerja akademik siswa dapat memberikan dampak yang signifikan bagi pendidik dan siswa dalam meningkatkan efektivitas proses pembelajaran. Hal ini terjadi karena kemampuan model yang dikembangkan dapat menangkap dan memetakan hubungan kompleks antara variabel-variabel yang terlibat dalam prediksi kinerja siswa. Dengan menerapkan teknik *educational data mining* (EDM), institusi pendidikan dapat mengungkap pola-pola tersembunyi dalam data pendidikan yang kompleks, serta sulit untuk dianalisis secara manual (Ozyurt et al., 2023). Kemampuan ini sangat berguna untuk mendeteksi lebih awal siswa yang berisiko mengalami penurunan kinerja akademik, sehingga memungkinkan tindakan pencegahan untuk mengatasi kekurangan dalam sistem pendidikan (Issah et al., 2023; Rodrigues et al., 2022). Secara keseluruhan, model prediksi yang dibangun ini berfungsi sebagai alat bantu yang efektif bagi lembaga pendidikan dalam memahami kebutuhan siswa dengan lebih mendalam, yang pada

gilirannya dapat mendorong peningkatan hasil belajar dan prestasi akademik siswa.

Secara keseluruhan, kontribusi utama yang diberikan dalam penelitian ini terletak pada kemampuannya untuk menyelesaikan masalah ketidakseimbangan data dalam prediksi kinerja akademik, sekaligus meningkatkan tingkat akurasi dan kestabilan model prediksi melalui kombinasi metode Random Forest, teknik SMOTE, dan *hyperparameter tuning* menggunakan GridSearchCV. Penerapan SMOTE berhasil menyeimbangkan distribusi kelas minoritas dan mayoritas, sehingga model dapat lebih mewakili karakteristik semua kelompok siswa. Sementara itu, tuning memastikan model dapat beradaptasi dengan karakteristik unik data dan algoritma, sehingga menghasilkan kinerja yang lebih optimal dan stabil. Meskipun masih dalam tahapan mendasar, berdasarkan analisis hubungan antara fitur akademik, demografi, dan dimensi kepribadian Big Five mengungkapkan adanya hubungan yang cukup menarik antara variabel-variabel tersebut, terutama peran signifikan nilai akademik sebelumnya dan tiga dimensi kepribadian (*agreeableness*, *openness*, dan *conscientiousness*) terhadap prestasi akademik. Temuan ini menegaskan adanya interaksi kompleks antara faktor akademik, kepribadian, dan demografi yang mempengaruhi kinerja siswa, sehingga membuka peluang penelitian lebih lanjut untuk mengeksplorasi mekanisme hubungan antara variabel ini, termasuk bagaimana kombinasi faktor-faktor tersebut dapat dimanfaatkan untuk meningkatkan prediksi dan strategi pembelajaran yang lebih efektif. Secara garis besar, penelitian ini tidak hanya menunjukkan peningkatan performa model dibandingkan studi sebelumnya, tetapi juga

memperlihatkan pentingnya pendekatan holistik dalam pemodelan kinerja akademik yang mempertimbangkan keseimbangan data, optimasi algoritma, dan hubungan variabel yang mendasar.



BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan penjelasan yang telah disampaikan dalam bagian hasil dan pembahasan, dapat disimpulkan bahwa penelitian ini berhasil menjawab pertanyaan penelitian yang telah ditetapkan, antara lain sebagai berikut :

1. Penelitian ini berhasil mengungkapkan bahwa selain faktor akademik, faktor kepribadian model Big Five serta faktor demografi khususnya kondisi sosial-ekonomi berkontribusi terhadap kinerja akademik siswa, meskipun pengaruhnya tergolong kecil. Oleh karena itu, untuk meningkatkan tingkat keakuratan serta nilai akurasi prediksi, perlu dilakukan analisis yang lebih mendalam terhadap faktor akademik, demografi, dan kepribadian yang lebih spesifik atau eksplorasi terhadap faktor-faktor tambahan yang relevan, dengan kondisi data yang lebih beragam.
2. Penelitian ini berhasil menunjukkan bahwa teknik SMOTE memiliki dampak signifikan dalam menangani ketidakseimbangan data, menjaga kestabilan data, serta mampu meningkatkan performa model prediksi. Namun, penelitian ini juga menunjukkan bahwa penggunaan teknik SMOTE secara *independent* tidak mampu meningkatkan performa model secara optimal. Dengan demikian, penyesuaian *hyperparameter* dalam metode Random Forest sangat penting untuk mencapai performa model prediksi yang lebih optimal. Maka, temuan ini menggaris bawahi bahwa kombinasi antara metode Random Forest,

GridSearchCV, dan SMOTE merupakan pendekatan yang paling efektif untuk kondisi dataset yang tidak seimbang.

3. Penelitian ini berhasil mengungkapkan bahwa pengembangan model prediksi kinerja akademik dengan menggunakan metode Random Forest dapat mencapai tingkat akurasi yang cukup tinggi disemua metrik evaluasi, berkisar antara 98% hingga 99% dalam berbagai skenario yang diuji. Namun, penelitian ini juga menekankan bahwa hasil yang diperoleh dari model Random Forest tidak dianggap sepenuhnya independent. Temuan penelitian ini juga menunjukkan bahwa penerapan SMOTE dan penerapan *hyperparameter tuning* seperti GridSearchCV, memiliki pengaruh yang signifikan terhadap peningkatan kinerja model. Pada dasarnya penerapan teknik SMOTE dan penerapan *hyperparameter tuning* yang tepat memungkinkan model prediksi Random Forest mencapai kinerja yang lebih optimal.

5.2. Saran

Meskipun model yang dikembangkan dalam penelitian ini menunjukkan hasil yang cukup memuaskan, penelitian ini juga mengidentifikasi beberapa keterbatasan, diantaranya ialah sebagai berikut :

1. Ukuran dataset yang digunakan sangat terbatas, sehingga hasilnya kurang dapat digeneralisasi. Hal ini disebabkan oleh fakta bahwa dataset hanya berasal dari satu sekolah negeri, yang berarti jumlah siswa dan karakteristik pembelajaran yang dianalisis juga terbatas. Sehingga, untuk meningkatkan kemampuan generalisasi dari hasil kinerja model prediksi, disarankan agar penelitian selanjutnya melakukan pengambilan sampel dengan variasi yang lebih luas.

2. Meskipun *hyperparameter tuning* telah memberikan kontribusi signifikan terhadap peningkatan kinerja model, dalam prosesnya teknik ini memerlukan waktu yang cukup lama, terutama dengan penggunaan GridSearchCV yang melibatkan banyak kombinasi parameter pada metode Random Forest. Oleh karena itu, untuk meningkatkan efisiensi tanpa mengorbankan performa, penting untuk mempertimbangkan penerapan teknik tuning lainnya yang dapat mempercepat waktu pencarian parameter optimal.
3. Meskipun teknik SMOTE menunjukkan hasil yang cukup baik dalam mengatasi masalah ketidakseimbangan data, hasil penelitian ini juga menunjukkan bahwa penerapan metode tersebut masih memiliki risiko kesalahan dalam prediksi yang cukup tinggi. Oleh karena itu, penelitian terkait eksplorasi terhadap teknik sampling lainnya diperlukan untuk menentukan metode yang paling efektif dalam menangani ketidakseimbangan data, serta untuk memahami pengaruhnya terhadap kualitas model prediksi, khususnya model Random Forest.
4. Saran untuk penelitian di masa mendatang sebaiknya mempertimbangkan pendekatan yang lebih inovatif, seperti penerapan metode *deep learning* atau penggabungan random forest dengan teknik *ensemble* lainnya untuk meningkatkan performa model secara keseluruhan.

DAFTAR PUSTAKA

PUSTAKA BUKU

- Brownlee, J. (2016). *Master Machine Learning Algorithms Discover How They Work and Implement Them From Scratch*. Machine Learning Mastery. <https://machinelearningmastery.com/master-machine-learning-algorithms/>
- Budiastuti, D., & Bandur, A. (2018). *Validitas dan Reliabilitas Penelitian*. Penerbit Mitra Wacana Media.
- Garson, G. D. (2013). *Validity & Reliability*. In Statistical Associates Publishing. Statistical Publishing Associates.
- Géron, A. (2019). *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow (2nd Editio)*. O'Reilly Media, Inc. <https://powerunit-ju.com/wp-content/uploads/2021/04/Aurelien-Geron-Hands-On-Machine-Learning-with-Scikit-Learn-Keras-and-Tensorflow--Concepts-Tools-and-Techniques-to-Build-Intelligent-Systems-OReilly-Media-2019.pdf>
- Ha, J., Kambe, M., & Pe, J. (2011). *Data Mining: Concepts and Techniques*. In *Data Mining: Concepts and Techniques (3rd Editio)*. Elsevier Inc. <https://doi.org/10.1016/C2009-0-61819-5>
- Hackeling, G. (2014). *Mastering Machine Learning with scikit-learn*. Packt Publishing Ltd. <https://www.amazon.com/Mastering-Machine-Learning-Scikit-learn-Algorithms/dp/1783988363>
- He, H., & Ma, Y. M. (2013). *Imbalanced Learning: Foundations, Algorithms, and Applications*. John Wiley & Sons. <https://doi.org/10.1002/9781118646106>
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). *Paradigm Shift to the Integrative Big-Five Trait Taxonomy: History, Measurement, and Conceptual Issues*. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). New York, NY: Guilford Press.
- McCrae, R. R., & Jr Costa, P. T. (2006). *Personality in adulthood: A five-factor theory perspective (2nd Editio)*. <https://www.guilford.com/books/Personality-in-Adulthood/McCrae-Costa/9781593852603>
- Muller, A. C., & Guido, S. (2016). *Introduction to Machine Learning with Python (1st Editio)*. O'Reilly Media, Inc. [https://www.nrigrupindia.com/e-book/Introduction-to-Machine-Learning-with-Python-\(PDFDrive.com\)-min.pdf](https://www.nrigrupindia.com/e-book/Introduction-to-Machine-Learning-with-Python-(PDFDrive.com)-min.pdf)

Pena-Ayala, A. (2013). Educational Data Mining Applications and Trends. Springer Cham. <https://doi.org/10.1007/978-3-319-02738-8>

Raschka, S., & Mirjalili, V. (2017). Python Machine Learning (2nd Editio). Packt Publishing Ltd. <http://radio.eng.niigata-u.ac.jp/wp/wp-content/uploads/2020/06/python-machine-learning-2nd.pdf>

Widiger, T. A. (2015). The Oxford Handbook of the Five Factor Model (P. E. Nathan (ed.)), Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199352487.001.0001>

PUSTAKA MAJALAH, JURNAL ILMIAH ATAU PROSIDING

Abdul Bujang, S. D., Selamat, A., Krejcar, O., Mohamed, F., Cheng, L. K., Chiu, P. C., & Fujita, H. (2023). Imbalanced Classification Methods for Student Grade Prediction: A Systematic Literature Review. IEEE Access, 11(October 2022), 1970–1989. <https://doi.org/10.1109/ACCESS.2022.3225404>

Arden, F., & Safitri, C. (2022). Hyperparameter Tuning Algorithm Comparison with Machine Learning Algorithms. Proceeding - 6th International Conference on Information Technology, Information Systems and Electrical Engineering: Applying Data Sciences and Artificial Intelligence Technologies for Environmental Sustainability, ICITISEE 2022, 183–188. <https://doi.org/10.1109/ICITISEE57756.2022.10057630>

Azhar, N. A., Mohd Pozi, M. S., Din, A. M., & Jatowt, A. (2023). An investigation of SMOTE based methods for imbalanced datasets with data complexity analysis. IEEE Transactions on Knowledge and Data Engineering, 35(7), 6651–6672. <https://doi.org/10.1109/TKDE.2022.3179381>

Chakraborty, S., Islam, S. H., & Samanta, D. (2022). Introduction to Data Mining and Knowledge Discovery. In Data Classification and Incremental Clustering in Data Mining and Machine Learning. EAI/Springer Innovations in Communication and Computing (pp. 1–22). Springer, Cham. https://doi.org/10.1007/978-3-030-93088-2_1

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE : Synthetic Minority Over-sampling Technique. Journal of Artificial Intelligence Research, 16, 321–357. <https://doi.org/10.1613/jair.953>

Chowdhury, A. A., Das, A., Hoque, K. K. S., & Karmaker, D. (2022). A Comparative Study of Hyperparameter Optimization Techniques for Deep Learning. Proceedings of International Joint Conference on Advances in Computational Intelligence, January, 509–521. https://doi.org/10.1007/978-981-19-0332-8_38

- Costa, A., Moreira, D., Casanova, J., Azevedo, Â., Gonçalves, A., Oliveira, Í., Azevedo, R., & Dias, P. C. (2024). Determinants of academic achievement from the middle to secondary school education: A systematic review. In *Social Psychology of Education* (Issue 0123456789). Springer Netherlands. <https://doi.org/10.1007/s11218-024-09941-z>
- Dol, S. M., & Jawandhiya, P. M. (2022). A Review of Data Mining in Education Sector. *Journal of Engineering Education Transformations*, 36(Special Issue 2), 13–22. <https://doi.org/10.16920/jeet/2023/v36is2/23003>
- Dol, S. M., & Jawandhiya, P. M. (2024). Systematic Review and Analysis of EDM for Predicting the Academic Performance of Students. In *Journal of The Institution of Engineers (India): Series B* (Issue 0123456789). Springer India. <https://doi.org/10.1007/s40031-024-00998-0>
- El-Keiey, S., ElMenshawly, D., & Hassanein, E. (2022). Student's Performance Prediction based on Personality Traits and Intelligence Quotient using Machine Learning. *International Journal of Advanced Computer Science and Applications*, 13(9), 292–299. <https://doi.org/10.14569/IJACSA.2022.0130934>
- Ghaniaviyanto Ramadhan, N., & Adiwijaya. (2022). Data Mining Techniques in Handling Personality Analysis for Ideal Customers. *Journal of Information Systems Engineering and Business Intelligence*, 8(2), 175–181. <https://doi.org/10.20473/jisebi.8.2.175-181>
- Gori, T., Sunyoto, A., & Al Fatta, H. (2024). Preprocessing Data dan Klasifikasi untuk Prediksi Kinerja Akademik Siswa. *Jurnal Teknologi Informasi Dan Ilmu Komputer*, 11(1), 215–224. <https://doi.org/10.25126/jtiik.20241118074>
- Hamoud, A. K., Kamel, M. B. M., Gaafar, A. S., Alasady, A. S., Humadi, A. M., Awadh, W. A., & Dahr, J. M. (2022). A prediction model based machine learning algorithms with feature selection approaches over imbalanced dataset. *Indonesian Journal of Electrical Engineering and Computer Science*, 28(2), 1105–1116. <https://doi.org/10.11591/ijeecs.v28.i2.pp1105-1116>
- Issah, I., Appiah, O., Appiahene, P., & Inusah, F. (2023). A systematic review of the literature on machine learning application of determining the attributes influencing academic performance. *Decision Analytics Journal*, 7(February), 100204. <https://doi.org/10.1016/j.dajour.2023.100204>
- Jatnika, H., Waluyo, A., & Azis, A. (2024). A Comparative Study on Data Collection Methods: Investigating Optimal Datasets for Data Mining Analysis. 5(1), 16–23.

- John, O. P., Donahue, E. M., & Kentle, R. (1991). *The Big Five Inventory-- Versions 4a and 54*. CA: University of California, Berkeley, Institute of Personality and Social Research.
- Khairy, D., Alharbi, N., Amasha, M. A., Areed, M. F., Alkhalaf, S., & Abougala, R. A. (2024). Prediction of student exam performance using data mining classification algorithms. *Education and Information Technologies*, 0123456789. <https://doi.org/10.1007/s10639-024-12619-w>
- Meyer, J., Jansen, T., Hübner, N., & Lüdtke, O. (2023). Disentangling the Association Between the Big Five Personality Traits and Student Achievement: Meta-Analytic Evidence on the Role of Domain Specificity and Achievement Measures. In *Educational Psychology Review* (Vol. 35, Issue 1). Springer US. <https://doi.org/10.1007/s10648-023-09736-2>
- Mienye, I. D., & Jere, N. (2024). A Survey of Decision Trees: Concepts, Algorithms, and Applications. *IEEE Access*, 12, 86716–86727. <https://doi.org/10.1109/ACCESS.2024.3416838>
- Nachouki, M., Mohamed, E. A., Mehdi, R., & Abou Naaj, M. (2023). Student course grade prediction using the random forest algorithm: Analysis of predictors' importance. *Trends in Neuroscience and Education*, 33, 100214. <https://doi.org/10.1016/j.tine.2023.100214>
- Ozyurt, O., Ozyurt, H., & Mishra, D. (2023). Uncovering the Educational Data Mining Landscape and Future Perspective: A Comprehensive Analysis. *IEEE Access*, 11(October), 120192–120208. <https://doi.org/10.1109/ACCESS.2023.3327624>
- Pajila, P. J. B., Sheena, B. G., Gayathri, A., Aswini, J., Nalini, M., & Siva Subramanian, R. (2023). A Comprehensive Survey on Naive Bayes Algorithm: Advantages, Limitations and Applications. *Proceedings of the 4th International Conference on Smart Electronics and Communication, ICOSEC 2023*, 1228–1234. <https://doi.org/10.1109/ICOSEC58147.2023.10276274>
- Pecuchova, J., & Drlik, M. (2023). Predicting Students at Risk of Early Dropping Out from Course Using Ensemble Classification Methods. *Procedia Computer Science*, 225, 3223–3232. <https://doi.org/10.1016/j.procs.2023.10.316>
- Rajvanshi, S., Kaur, G., Dhatwalia, A., Arunima, Singla, A., & Bhasin, A. (2024). Research on Problems and Solutions of Overfitting in Machine Learning. In *Lecture Notes in Electrical Engineering: Vol. 1191 LNEE*. Springer Nature Singapore. https://doi.org/10.1007/978-981-97-2508-3_47
- Rico-Juan, J. R., Cachero, C., & Macià, H. (2024). Study regarding the influence of a student's personality and an LMS usage profile on learning performance

- using machine learning techniques. *Applied Intelligence*, 54(8), 6175–6197. <https://doi.org/10.1007/s10489-024-05483-1>
- Rimal, Y., Sharma, N., & Alsadoon, A. (2024). The accuracy of machine learning models relies on hyperparameter tuning: student result classification using random forest, randomized search, grid search, bayesian, genetic, and optuna algorithms. *Multimedia Tools and Applications*, 83(30), 74349–74364. <https://doi.org/10.1007/s11042-024-18426-2>
- Rodrigues, L. S., Dos Santos, M., Costa, I., & Moreira, M. A. L. (2022). Student Performance Prediction on Primary and Secondary Schools-A Systematic Literature Review. *Procedia Computer Science*, 214(C), 680–687. <https://doi.org/10.1016/j.procs.2022.11.229>
- Roslan, M. H. bin, & Chen, C. J. (2022). Educational Data Mining for Student Performance Prediction: A Systematic Literature Review (2015-2021). *International Journal of Emerging Technologies in Learning*, 17(5), 147–179. <https://doi.org/10.3991/ijet.v17i05.27685>
- Roslan, M. H. Bin, & Chen, C. J. (2023). Predicting students' performance in English and Mathematics using data mining techniques. *Education and Information Technologies*, 28(2), 1427–1453. <https://doi.org/10.1007/s10639-022-11259-2>
- Santoso, A., Retnawati, H., Kartianom, Apino, E., Rafi, I., & Rosyada, M. N. (2024). Predicting Time to Graduation of Open University Students: An Educational Data Mining Study. *Open Education Studies*, 6(1). <https://doi.org/10.1515/edu-2022-0220>
- Shaninah, F. S. E., & Mohd Noor, M. H. (2024). The impact of big five personality trait in predicting student academic performance. *Journal of Applied Research in Higher Education*, 16(2), 523–539. <https://doi.org/10.1108/JARHE-08-2022-0274>
- Sweta, S. (2024). An Overview of Sentiment Analysis and Educational Data Mining. In *Sentiment Analysis and its Application in Educational Data Mining*. SpringerBriefs in Applied Sciences and Technology (pp. 1–18). Springer, Singapore. https://doi.org/10.1007/978-981-97-2474-1_1
- Wawan, & Retnawati, H. (2022). Empirical Study of Factors Affecting the Students' Mathematics Learning Achievement. *International Journal of Instruction*, 15(2), 417–434. <https://doi.org/10.29333/iji.2022.15223a>

LAMPIRAN

❖ LEMBAR VALIDASI ANGKET KEPERIBADIAN BIG FIVE

A. Deskripsi

Angket ini bertujuan untuk memvalidasi instrument pengukuran kepribadian big five (Neuroticism (N), Extraversion (E), Openness (O), Agreeableness (A), dan Conscientiousness (C)). Validasi yang diberikan akan sangat berguna untuk mencapai kelayakan instrument pengukuran kepribadian big five yang akan dilakukan kepada siswa-siswi sekolah menengah pertama. Sehubungan dengan hal tersebut, Bapak/Ibu sebagai ahli psikolog dimohon untuk dapat memberikan tanggapan dan komentar/saran perbaikan terhadap instrument yang telah disusun.

Sasaran : Peserta Didik Sekolah Menengah Pertama (SMP)

Judul Penelitian : Prediksi Kinerja Akademik Matematika Siswa Berdasarkan Kepribadian Big Five menggunakan Random Forest dengan Teknik Synthetic Minority Over-Sampling

Peneliti : Annisa Nurul Pratiwi, S.Kom.

Evaluator :

1. Vivi Ade Cerliana, S.Psi., M.Psi., Psikolog.
2. Hikmatul Fatiyah, S.Psi., M.Psi., Psikolog.

Pekerjaan : Psikolog

B. Petunjuk Pengisian Oleh Validator

1. Lembar validasi angket kepribadian big five ini diisi oleh validator.
2. Lembar validasi ini dimaksudkan untuk mengukur kevalidan angket kepribadian big five yang akan digunakan dalam penelitian tugas akhir dengan judul “Prediksi Kinerja Akademik Matematika Siswa Berdasarkan Kepribadian Big Five menggunakan Random Forest dengan Teknik Synthetic Minority Over-Sampling”.
3. Berilah tanda *check* (√) pada kolom skor yang tersedia untuk tiap aspek yang dinilai sesuai dengan penilaian Anda dengan keadaan yang sebenarnya.
4. Jawaban yang diberikan pada kolom skala penilaian disediakan dengan skala penilaian sebagai berikut :
 - a. Point 5 dengan kriteria penilaian “sangat baik/sesuai”
 - b. Point 4 dengan kriteria penilaian “baik/sesuai”
 - c. Point 3 dengan kriteria penilaian “cukup baik/sesuai”
 - d. Point 2 dengan kriteria penilaian “kurang baik/sesuai”
 - e. Point 1 dengan kriteria penilaian “tidak baik/sesuai”
5. Komentar umum dan saran perbaikan terhadap keseluruhan isi angket kepribadian big five dapat diisi pada tempat yang telah disediakan.

Terimakasih atas ketersediaan Bapak/Ibu dalam melakukan pengujian dan mengisi lembar evaluasi ini.

C. Kisi-Kisi Angket Kepribadian

No	Domain	Deskripsi	Pernyataan	No Butir
1	Extraversion (Extraversi)	Kecenderungan individu untuk bersikap sosial, aktif, dan bersemangat. Individu cenderung bersikap ramah, energik, dan senang berinteraksi dengan orang lain.	Saya adalah seseorang yang banyak bicara	1
			Saya adalah seseorang yang pendiam	2-R
			Saya adalah seseorang yang penuh energi	3
			Saya adalah seseorang yang penuh semangat	4-R
			Saya adalah seseorang yang cenderung pendiam	5
			Saya adalah seseorang yang memiliki kepribadian yang tegas	6
			Saya adalah seseorang yang kadang-kadang pemalu dan kurang terbuka	7-R
Saya adalah seseorang yang suka bersosialisasi	8			
2	Agreeableness (Kesepakatan/Mudah Akur)	Kecenderungan individu bersikap ramah, kooperatif, dan penuh empati. Individu cenderung lebih mudah bekerja sama, penuh perhatian, dan peduli terhadap kesejahteraan orang lain.	Saya adalah seseorang yang cenderung mencari kesalahan orang lain	9-R
			Saya adalah seseorang yang suka membantu dan tidak egois terhadap orang lain	10
			Saya adalah seseorang yang suka bertengkar dengan orang lain	11-R
			Saya adalah seseorang yang pemaaf	12
			Saya adalah seseorang yang sering kali dapat dipercaya	13
			Saya adalah seseorang yang dapat bersikap dingin dan acuh tak acuh	14-R
			Saya adalah seseorang yang pengertian dan baik kepada hampir semua orang	15
Saya adalah seseorang yang kadang kasar kepada orang lain	16-R			
Saya adalah seseorang yang suka bekerja sama dengan orang lain	17			
3	Conscientiousness (Berhati-hati)	Kecenderungan individu bersikap disiplin, teratur, dan bertanggung jawab. Individu cenderung lebih teliti, terorganisir, dan memiliki motivasi yang kuat untuk mencapai tujuan.	Saya adalah seseorang yang bekerja dengan teliti	18
			Saya adalah seseorang yang kadang ceroboh	19-R
			Saya adalah seseorang yang pekerja yang dapat diandalkan	20
			Saya adalah seseorang yang cenderung tidak teratur	21-R
			Saya adalah seseorang yang cenderung malas	22-R
			Saya adalah seseorang yang gigih sampai tugas selesai	23
			Saya adalah seseorang yang melakukan segala sesuatu dengan efisien	24
Saya adalah seseorang yang menyusun rencana dan melaksanakannya	25			
Saya adalah seseorang yang mudah kehilangan fokus	26-R			
4			Saya adalah seseorang yang murung	27

	Neuroticism (Neurotisme)	Kecenderungan individu untuk merasakan emosi negatif, seperti kecemasan, kemarahan, atau depresi. Individu lebih rentan terhadap perasaan cemas, tegang, dan khawatir, serta lebih sering merasa stress atau tertekan dalam menghadapi situasi tertentu.	Saya adalah seseorang yang menangani stress dengan baik	28-R
			Saya adalah seseorang yang mudah merasa tegang	29
			Saya adalah seseorang yang sering khawatir	30
			Saya adalah seseorang yang stabil secara emosional, tidak mudah tersinggung	31-R
			Saya adalah seseorang yang terkadang suasana hatinya tidak stabil	32
			Saya adalah seseorang yang tetap tenang dalam situasi tegang	33-R
			Saya adalah seseorang yang mudah gugup	34
			Saya adalah seseorang yang menghasilkan ide-ide baru	35
			Saya adalah seseorang yang penasaran dengan berbagai hal	36
			Saya adalah seseorang yang pemikir yang mendalam	37
5	Openness (Keterbukaan terhadap hal-hal baru)	Kecenderungan individu menerima pengalaman baru, memiliki daya imajinasi yang tinggi, dan memiliki minat yang beragam. Individu cenderung lebih kreatif, rasa ingin tahu yang besar, dan terbuka terhadap ide-ide baru.	Saya adalah seseorang yang memiliki imajinasi yang aktif	38
			Saya adalah seseorang yang penuh ide	39
			Saya adalah seseorang yang memiliki apresiasi terhadap seni dan keindahan	40
			Saya adalah seseorang yang lebih suka pekerjaan yang rutin	41-R
			Saya adalah seseorang yang suka merenung untuk mengeksplorasi ide-ide	42
			Saya adalah seseorang yang kurang tertarik pada seni	43-R
			Saya adalah seseorang yang berwawasan luas dalam seni, musik, atau sastra	44

D. Penilaian

No	Aspek	Kriteria	Skor					Catatan
			1	2	3	4	5	
1	Lu	Kemampuan pernyataan dengan skenario					✓	
		Agresivitas (Berperilaku Tidak Akur)					✓	
		Kemampuan pernyataan dengan skenario (Tanggung Jawab)					✓	
		Kemampuan pernyataan dengan skenario (Cemas/kesemasan / takut/hati-hati)					✓	
		Kemampuan pernyataan dengan skenario (Optimis / Keberanian terhadap hal-hal baru)					✓	
		Kemampuan pernyataan dengan skenario (Nostalgia / Kelelahan Emosional)					✓	
		Kemampuan pernyataan dengan objek penelitian (sewa menaruh barang)					✓	
2	Kewajiban	Statis pernyataan diumumkan secara jels sehingga tidak mengancam					✓	
		Jamuran masalah dan mengajari solusi dengan situasi kehidupan sosial Day Live					✓	
		Setiap item dalam instrumen tersebut dan menunjukkan karakteristik yang sesuai dengan dimensi yang diukur					✓	
3	Kebudayaan	Kemampuan pernyataan dengan cara bahasa atau kodek Dalam bahasa					✓	
		Berisi soal kata/mula yang berlaku umum					✓	
		Efektifitas dan reliabilitas instrumen					✓	
		Penggunaan bahasa secara efektif dan efisien pada instrumen					✓	

Gambar Lampiran 1. Form Penilaian Angket Psikolog 1

E. Komentar dan Saran Perbaikan

Pernyataan terlalu singkat
 sebaiknya membuat titik responden bersikap jujur
 Sasan pernyataan yang lebih mudah dipahami, seperti mengintimidasi

Kesimpulan :
 Apakah pernyataan itu (sangat) dapat dipercaya?
 1. Sangat dapat dipercaya dengan tanpa keraguan
 2. Dapat dipercaya dengan keraguan yang kecil
 3. Tidak dapat dipercaya
 *Lingkari salah satu pada kolom

Citebon 28 November 2024
 Widiyanti
 Widiyanti, S.Psi., M.Psi., Psikolog

Gambar Lampiran 2. Form Catatan Psikolog 1

D. Penilaian

No	Aspek	Kriteria	Skor					Catatan
			1	2	3	4	5	
1	1a	Kemampuan pernyataan dengan domain Agresi/tenes (Konspikasi/Malah Akur)				✓		
		Kemampuan pernyataan dengan domain Entusias (Ektroser)				✓		
		Kemampuan pernyataan dengan domain Konsentrasi (Berfikir)				✓		
		Kemampuan pernyataan dengan domain Optimis (Ketertarikan terhadap hal-hal baru)			✓			
		Kemampuan pernyataan dengan domain Neurotizis (Ketidakhidupan)				✓		
		Kemampuan pernyataan dengan domain Kesesuaian (sa pernyataan dengan objek penelitian (tawar-menawar, seriosa)					✓	
		Hasil pernyataan ditunjukkan secara jelas sehingga tidak ambigu			✓			
2	Kontak	Intervius mengutip dan mengutip kelima dimensi utama keprilaku model Big Five				✓		
		Sebagi hasil dari wawancara relevan dan mencerminkan karakteristik yang sesuai dengan domain yang diukur				✓		
		Kemampuan pernyataan dengan sa bahasa asli Kuala Bahasa Indonesia				✓		
3	Kefahaman	Penggunaan kata istilah yang berlaku umum dalam komunikasi terapan dalam penelitian			✓			
		Penggunaan bahasa secara efektif dan efisien pada intonas			✓			

Gambar Lampiran 3. Form Penilaian Angket Psikolog 2

E. Komentar dan Saran Perubahan

① Bersikap dingin? artinya apa? gunakan istilah yang familiar

②. Susun pernyataan dengan kalimat yang lebih detail agar maksud dan tujuan tidak melenceng.

Kesimpulan :

Angket keterbacaan ini (jika ini dirombak *)

4. Layak digunakan dengan tingkat rendah


② Layak digunakan dengan tingkat rendah sesuai teori

3. Tidak layak digunakan

* (jika ini adalah satu pada lembar)

Cirebon, 28 November 2024

Validator



Haniatul Fatah, S.Pd., M.Pd., Psikolog

Gambar Lampiran 4. Form Catatan Psikolog 2