

PERBANDINGAN KINERJA MODEL MACHINE LEARNING UNTUK PREDIKSI RISIKO PUTUS SEKOLAH TINGKAT PROVINSI DI INDONESIA DAN INTERPRETASI FITUR DENGAN SHAP

Aswan Supriyadi Sunge^{1*}, Andi Nugroho², Suzanna³, Ahmad Turmudi Zy⁴, Agung Nugroho⁵,
Ismasari Nawangsih⁶, Edora⁷

^{1,4,5,6}Teknik Informatika, Universitas Pelita Bangsa, Bekasi, Indonesia

²Sistem Informasi, Universitas Mercu Buana, Jakarta, Indonesia

³Binus Online Learning, Universitas Bina Nusantara, Jakarta, Indonesia

⁷Pendidikan Guru Sekolah Dasar, Universitas Pelita Bangsa, Bekasi, Indonesia

*Penulis korespondensi: aswan.sunge@pelitabangsa.ac.id

ABSTRAK

Angka putus sekolah di Indonesia masih menjadi salah satu permasalahan pendidikan yang perlu mendapat perhatian karena terdapat perbedaan kondisi pendidikan antar provinsi yang dipengaruhi oleh berbagai faktor, seperti kualitas sumber daya manusia, jumlah siswa, serta kondisi sarana dan prasarana sekolah. Penelitian ini bertujuan untuk membandingkan kinerja beberapa algoritma machine learning dalam memprediksi risiko putus sekolah tingkat provinsi di Indonesia dan mengidentifikasi faktor-faktor yang paling berpengaruh menggunakan metode *shapley additive explanations* (shap). Data yang digunakan merupakan data sekunder pendidikan sekolah dasar tahun 2023–2024 yang bersumber dari kementerian pendidikan, kebudayaan, riset, dan teknologi republik Indonesia dan diperoleh melalui platform kaggle. Dataset mencakup 39 provinsi dengan 12 variabel prediktor yang meliputi jumlah siswa mengulang, jumlah rombongan belajar, jumlah kepala sekolah dan guru berdasarkan kualifikasi pendidikan, jumlah tenaga kependidikan, jumlah siswa, jumlah sekolah, serta kondisi ruang kelas. Tahapan penelitian meliputi *preprocessing* data, pembentukan kelas risiko putus sekolah, pembagian data menggunakan rasio 80:20, serta pelatihan empat *Model Machine Learning*, yaitu *random forest*, *decision tree*, *naive bayes*, dan *logistic regression*. Hasil pengujian menunjukkan bahwa *random forest* menghasilkan performa terbaik dengan *Accuracy*, *Precision*, *Recall*, dan *F1-Score* masing-masing sebesar 1.0000, diikuti oleh *naive bayes* dengan *Accuracy* sebesar 0.8750 dan *F1-Score* sebesar 0.8571, sedangkan *decision tree* dan *logistic regression* memperoleh *Accuracy* sebesar 0.7500 dan *F1-Score* sebesar 0.7500. Hasil analisis *roc curve* menunjukkan bahwa *random forest* dan *naive bayes* menghasilkan nilai auc sebesar 1.00, sedangkan *decision tree* dan *logistic regression* memperoleh nilai auc sebesar 0.75. Uji McNemar menunjukkan bahwa tidak terdapat perbedaan performa yang signifikan secara statistik antar model yang dibandingkan. Analisis shap menunjukkan bahwa variabel mengulang, rombongan belajar, kepala sekolah dan guru ($< s1$), serta kepala sekolah dan guru ($\geq s1$) merupakan faktor yang paling berpengaruh terhadap prediksi risiko putus sekolah. Hasil penelitian ini menunjukkan bahwa *machine learning* dapat dimanfaatkan sebagai alat bantu untuk mengidentifikasi risiko putus sekolah serta memberikan informasi mengenai faktor-faktor yang perlu menjadi perhatian dalam penyusunan kebijakan pendidikan berbasis data.

Kata kunci: Prediksi, Putus Sekolah, Komparasi, Machine Learning, SHAP

1 PENDAHULUAN

Pendidikan memiliki peran penting dalam meningkatkan kualitas sumber daya manusia dan mendukung pembangunan sosial maupun ekonomi suatu negara (Kadrina, 2025). Meskipun akses pendidikan di Indonesia terus mengalami peningkatan, masalah putus sekolah masih menjadi tantangan yang perlu mendapat perhatian. Tingginya angka putus sekolah dapat berdampak pada rendahnya kualitas tenaga kerja, meningkatnya kesenjangan sosial, serta berkurangnya kesempatan masyarakat untuk memperoleh pendidikan yang lebih baik (Nadila, 2025). Oleh karena itu, identifikasi risiko putus sekolah sejak dini menjadi penting agar kebijakan dan program pendidikan dapat dirancang secara lebih tepat sasaran beberapa faktor yang saling berkaitan, selain faktor yang berasal dari peserta didik, kondisi sekolah dan kualitas sumber daya pendidikan juga memiliki peran penting dalam mendukung keberlangsungan proses pembelajaran (Temu, 2016). Ekonomi, budaya, jumlah siswa yang mengulang, kualitas kepala sekolah dan guru, jumlah tenaga kependidikan, serta kondisi sarana dan prasarana sekolah merupakan beberapa faktor yang dapat mempengaruhi tingkat keberhasilan Pendidikan (Alifa, 2023). Karakteristik tersebut dapat berbeda antar provinsi di Indonesia sehingga diperlukan pendekatan yang mampu mengolah berbagai informasi secara efektif untuk mendukung pengambilan keputusan yang lebih baik.

Perkembangan teknologi dan ketersediaan data telah mendorong pemanfaatan *Machine Learning* (ML) dalam berbagai bidang, salah satu pendidikan yang memungkinkan komputer mempelajari pola dari data historis untuk menghasilkan prediksi terhadap data baru tanpa memerlukan aturan yang ditentukan secara eksplisit (Trujillo, 2024). Kemampuan tersebut menjadikan sebagai salah satu pendekatan yang banyak digunakan dalam mendukung pengambilan keputusan berbasis data (Prayoga, 2025). Dalam bidang pendidikan, dapat dimanfaatkan untuk membantu mengidentifikasi wilayah yang memiliki risiko putus sekolah yang lebih tinggi sehingga upaya pencegahan dapat dilakukan secara lebih dini (Habibi, 2025).

Berbagai penelitian sebelumnya telah memanfaatkan ML untuk memprediksi performa akademik maupun risiko putus sekolah dan sebagian besar penelitian tersebut dilakukan pada tingkat individu siswa atau lingkungan pendidikan tinggi (Elbasi, 2025). Padahal, karakteristik pendidikan pada tingkat wilayah memiliki pola yang berbeda sehingga diperlukan kajian yang lebih spesifik untuk memperoleh gambaran yang sesuai dengan kondisi di Indonesia. Selain itu, sebagian besar penelitian lebih berfokus pada kemampuan prediksi model, sedangkan kajian yang menghubungkan hasil prediksi dengan faktor-faktor yang mempengaruhi risiko putus sekolah pada tingkat provinsi masih relatif terbatas.

Berdasarkan kondisi tersebut, masih terdapat keterbatasan penelitian yang membahas prediksi risiko putus sekolah pada tingkat provinsi di Indonesia dengan membandingkan beberapa model ML sekaligus mengidentifikasi faktor-faktor yang berkontribusi terhadap hasil prediksi. Oleh karena itu, penelitian ini bertujuan untuk membandingkan kinerja beberapa model dalam memprediksi risiko putus sekolah tingkat provinsi di Indonesia serta mengidentifikasi faktor-faktor yang paling berpengaruh terhadap risiko tersebut. Hasil penelitian ini diharapkan dapat memberikan gambaran yang lebih komprehensif mengenai pola risiko putus sekolah dan menjadi masukan bagi pemerintah maupun pemangku kepentingan dalam menyusun kebijakan pendidikan yang lebih tepat sasaran. Selain memberikan perbandingan performa beberapa model ML, penelitian ini juga berkontribusi dalam memberikan pemahaman mengenai faktor-faktor yang berkaitan dengan risiko putus sekolah pada tingkat provinsi di Indonesia sehingga dapat mendukung pengambilan keputusan yang lebih berbasis data.

2 PENELITIAN TERDAHULU

Dalam bidang pendidikan, pendekatan ML telah digunakan untuk berbagai tujuan, seperti memprediksi performa akademik, mengidentifikasi siswa yang berisiko putus sekolah, mengevaluasi tingkat keberhasilan pembelajaran, serta mendukung pengambilan keputusan berbasis data, dan menjadi salah satu pendekatan yang paling banyak digunakan dalam educational data mining karena mampu mengolah data yang kompleks dan menghasilkan prediksi yang lebih baik dibandingkan pendekatan konvensional (Wu, 2020).

Random Forest (RF) merupakan gabungan sejumlah pohon keputusan atau decision tree untuk meningkatkan kemampuan prediksi, dan setiap pohon dibangun dari sampel data yang berbeda dan hasil prediksi akhir diperoleh melalui mekanisme voting (Louppe, 2014). Kelebihan utama berkemampuan dalam menangani data berdimensi tinggi, ketahanan terhadap noise, serta performa yang relatif stabil (Breiman, 2001). Dalam bidang pendidikan, model ini telah digunakan untuk memprediksi siswa yang berisiko putus sekolah dan menunjukkan performa yang baik dibandingkan beberapa metode klasifikasi lainnya (Dermol, 2026) (Santos, 2026).

Decision Tree (DT) merupakan salah satu algoritma klasifikasi yang membentuk struktur pohon berdasarkan proses pemisahan data secara bertahap menggunakan atribut yang paling informatif. Algoritma ini memiliki keunggulan berupa kemudahan interpretasi dan kemampuan menghasilkan aturan klasifikasi yang mudah dipahami oleh pengguna (Matzavela, 2021) (Izza, 2020), di bidang pendidikan, digunakan untuk memprediksi performa akademik maupun risiko putus sekolah (Mariano, 2022).

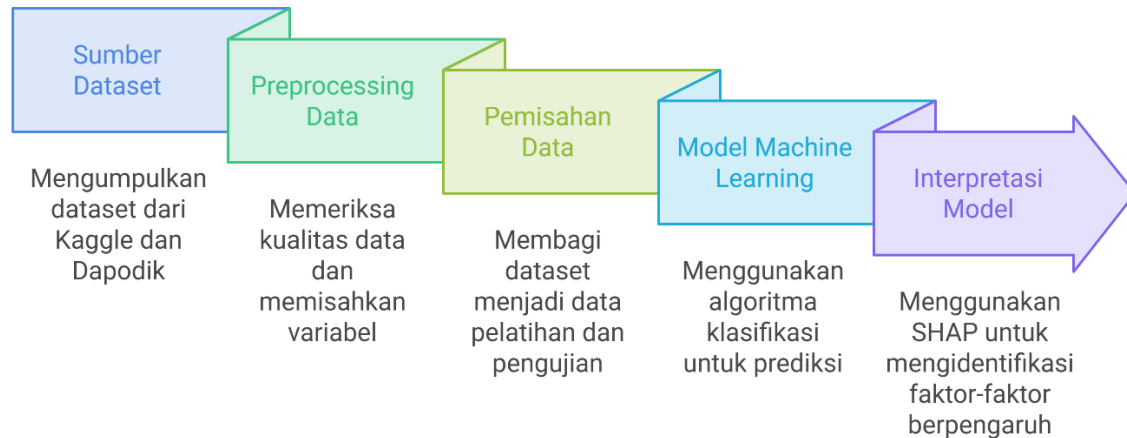
Naive Bayes (NB) merupakan algoritma klasifikasi berbasis probabilitas yang menggunakan Teorema Bayes dengan asumsi bahwa setiap variabel bersifat independen satu sama lain. Algoritma ini dikenal memiliki proses pelatihan yang sederhana, kebutuhan komputasi yang rendah, serta mampu memberikan hasil yang baik pada berbagai permasalahan klasifikasi (Peretz, 2024), dalam bidang pendidikan, dan model ini telah digunakan untuk memprediksi keberhasilan akademik maupun tingkat putus sekolah (Kirana, 2024).

Logistic Regression (LR) merupakan metode klasifikasi yang banyak digunakan untuk memprediksi probabilitas suatu kejadian berdasarkan hubungan antara variabel bebas dan variabel target, dan memiliki keunggulan berupa implementasi yang sederhana dan hasil yang relatif mudah diinterpretasikan (Chung, 2020), di bidang pendidikan, digunakan dalam memprediksi tingkat keberhasilan siswa dan risiko putus sekolah (Singh, 2022).

Seiring meningkatnya penggunaan ML dalam berbagai bidang, kebutuhan terhadap interpretasi hasil prediksi juga menjadi semakin penting. Salah satu model SHapley Additive exPlanations (SHAP) merupakan salah satu metode yang banyak digunakan untuk menjelaskan kontribusi setiap variabel terhadap hasil prediksi model, dan konsep dari teori permainan dan mampu memberikan informasi mengenai pengaruh masing-masing fitur secara konsisten, dan interpretabilitas menjadi salah satu aspek penting dalam penelitian prediksi putus sekolah karena membantu memahami faktor-faktor yang mempengaruhi hasil prediksi (Lundberg, 2017) (Ponce-Bobadilla, 2024).

3 MODEL YANG DIUSULKAN

Penelitian ini mengusulkan suatu kerangka kerja untuk memprediksi risiko putus sekolah tingkat provinsi di Indonesia dari pengumpulan data, *Preprocessing*, pembagian data, pembangunan model klasifikasi, hingga interpretasi hasil prediksi menggunakan SHAP dan alur penelitian diusulkan dapat dilihat pada Gambar 1.



Gambar 1. Alur Penelitian

3.1 Sumber Dataset

Dataset yang digunakan dalam penelitian ini merupakan dataset pendidikan sekolah dasar tingkat provinsi di Indonesia tahun 2023–2024 yang diperoleh melalui platform Kaggle. Data yang tersedia mencakup berbagai indikator pendidikan pada 38 provinsi di Indonesia (Pastika, 2024).

3.2 Preprocessing

Tahap ini mempersiapkan data sebelum digunakan dalam proses pembangunan model ML dengan melakukan pemeriksaan terhadap kualitas data untuk memastikan tidak terdapat nilai yang hilang maupun data yang tidak konsisten. Selanjutnya dilakukan pemisahan antara variabel prediktor dan variabel target. Variabel target kemudian ditransformasikan ke dalam bentuk kategorikal yang merepresentasikan tingkat risiko putus sekolah, hal ini bertujuan untuk menghasilkan data yang lebih terstruktur sehingga dapat meningkatkan kualitas model yang dibangun.

3.3 Splitting

Setelah tahap *Preprocessing* dilakukan, dataset dibagi menjadi data pelatihan dan data pengujian. Data pelatihan digunakan untuk membangun model ML, sedangkan data pengujian digunakan untuk mengevaluasi kemampuan model dalam melakukan prediksi terhadap data yang belum pernah digunakan sebelumnya. Dalam penelitian ini, pembagian data sebanyak 80% data digunakan sebagai data pelatihan dan 20% sisanya digunakan sebagai data pengujian.

3.4 Model Machine Learning

Penelitian ini menggunakan empat algoritma ML yang umum digunakan pada permasalahan klasifikasi, hal ini bertujuan untuk memperoleh model yang memiliki kemampuan terbaik dalam memprediksi risiko putus sekolah.

3.5 Model Interpretasi

Selain menghasilkan prediksi, penelitian ini juga bertujuan untuk mengidentifikasi faktor-faktor yang paling berpengaruh terhadap risiko putus sekolah. Oleh karena itu SHAP digunakan sebagai metode interpretasi model, lalu menghitung kontribusi setiap variabel terhadap hasil prediksi.

4 HASIL DAN PEMBAHASAN

4.1 Perbandingan Kinerja Model Machine Learning

Dari hasil pengujian dari empat model ML maka didapat dari *Accuracy*, *Precision*, *Recall*, dan *F1-Score*, untuk hasil lebih lengkap pengujian masing-masing model ditunjukkan pada Tabel 1.

Tabel 1. Hasil Perbandingan Kinerja Model

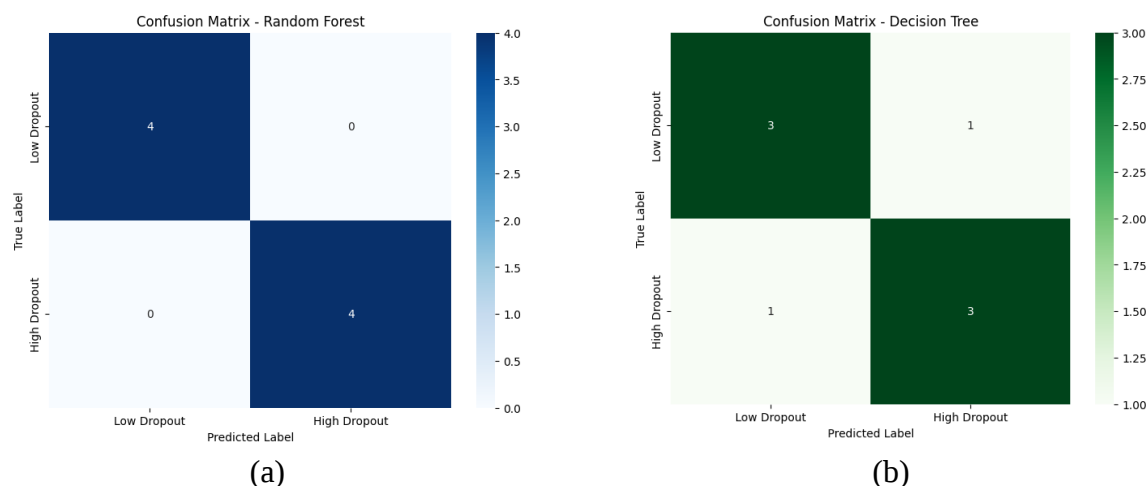
Model	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>
Random Forest	1.0000	1.0000	1.0000	1.0000
Decision Tree	0.7500	0.7500	0.7500	0.7500
Naive Bayes	0.8750	1.0000	0.7500	0.8571
Logistic Regression	0.7500	0.7500	0.7500	0.7500

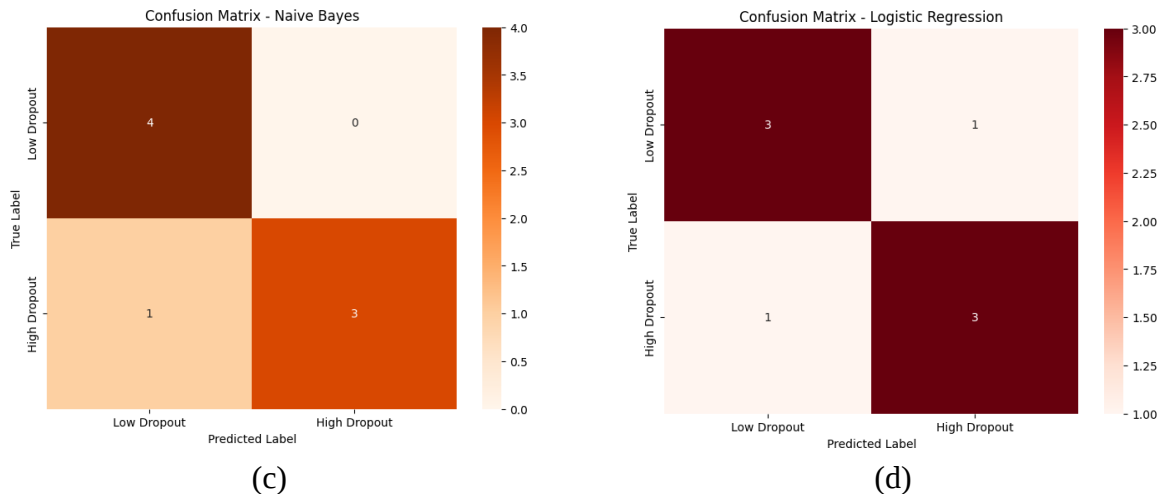
Sumber: Hasil Pengujian Dengan Google Colab

Berdasarkan Tabel 1, RF menghasilkan performa terbaik semua metric sebesar 1.0000, dan hasil tersebut menunjukkan bahwa seluruh data pengujian dapat diklasifikasikan dengan benar. NB menempati urutan kedua dengan *Accuracy* sebesar 0.8750 dan *F1-Score* sebesar 0.8571. Sementara itu, DT dan LR menghasilkan performa yang sama dengan *Accuracy* sebesar 0.7500. Secara umum, hasil tersebut menunjukkan bahwa pendekatan ensemble pada RF mampu memberikan kemampuan klasifikasi yang lebih baik dibandingkan model lainnya. Namun, jumlah data pengujian yang relatif terbatas menyebabkan hasil ini perlu diinterpretasikan secara hati-hati dan memerlukan validasi lebih lanjut menggunakan jumlah data yang lebih besar.

4.2 Analisis Confusion Matrix

Hasil Confusion Matrix dari masing-masing model ditunjukkan Gambar 2(a) menunjukkan RF, Gambar 2(b) menunjukkan DT, Gambar 2(c) menunjukkan NB, sedangkan Gambar 2(d) menunjukkan LR. Analisis bertujuan untuk melihat kemampuan masing-masing model dalam mengklasifikasikan data risiko putus sekolah ke dalam kelas risiko rendah dan risiko tinggi.





Gambar 2. Confusion Matrix Keseluruhan Model

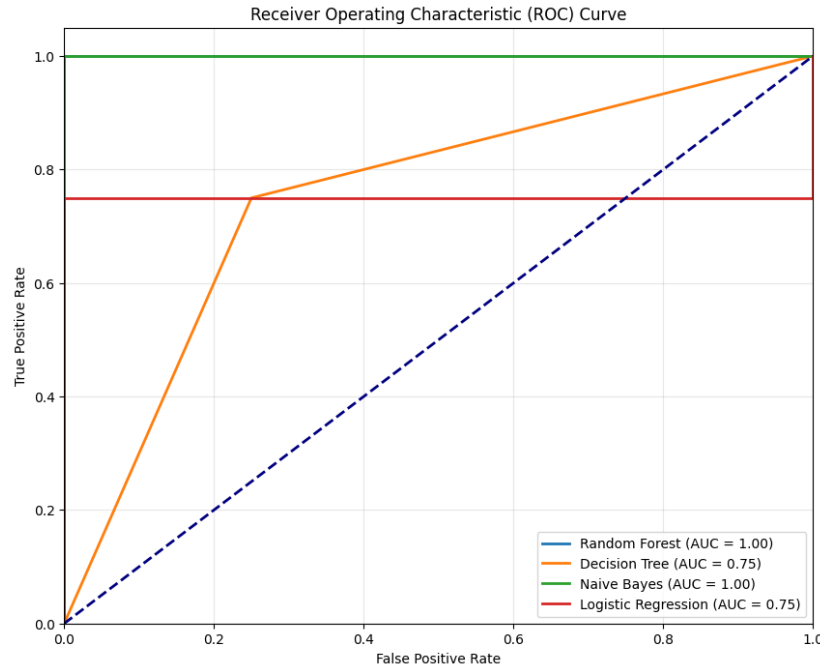
Sumber: Hasil Pengujian Dengan Google Colab

Berdasarkan Gambar 2(a), RF mampu mengklasifikasikan seluruh delapan data pengujian dengan benar, yang terdiri dari empat data pada kelas risiko rendah dan empat data pada kelas risiko tinggi. Tidak terdapat kesalahan klasifikasi sehingga model ini menghasilkan nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score* sebesar 1.0000. Hasil tersebut menunjukkan bahwa memiliki kemampuan yang sangat baik dalam membedakan kedua kelas pada dataset yang digunakan. Pada Gambar 2(b), DT menghasilkan enam prediksi yang benar dan dua kesalahan klasifikasi. Satu data dari kelas risiko rendah diprediksi sebagai risiko tinggi dan satu data dari kelas risiko tinggi diprediksi sebagai risiko rendah. Pola yang sama juga ditunjukkan oleh LR pada Gambar 2(d), yang menghasilkan enam prediksi benar dan dua prediksi yang tidak sesuai. Kondisi tersebut menyebabkan kedua model memperoleh nilai *Accuracy* sebesar 0.7500. Sementara itu, Gambar 2(c) menunjukkan bahwa NB mampu mengidentifikasi seluruh data pada kelas risiko rendah dengan benar, namun terdapat satu data pada kelas risiko tinggi yang diprediksi sebagai kelas risiko rendah. Dengan demikian, NB menghasilkan tujuh prediksi yang benar dari delapan data pengujian dan memperoleh *Accuracy* sebesar 0.8750. Meskipun nilai *Precision* yang dihasilkan mencapai 1.0000, adanya satu kesalahan klasifikasi menyebabkan nilai *Recall* dan *F1-Score* menjadi lebih rendah dibandingkan RF.

Secara keseluruhan, hasil *Confusion Matrix* menunjukkan bahwa RF memiliki kemampuan klasifikasi terbaik dibandingkan model lainnya. NB berada pada urutan kedua dengan hanya satu kesalahan klasifikasi, sedangkan DT dan LR menghasilkan pola prediksi yang sama dengan dua kesalahan klasifikasi. Meskipun demikian, jumlah data pengujian yang relatif kecil menyebabkan hasil yang diperoleh perlu diinterpretasikan secara hati-hati dan memerlukan pengujian lebih lanjut menggunakan jumlah data yang lebih besar untuk memastikan konsistensi performa model.

4.3 Analisis ROC Curve

Hasil Gambar 3 menunjukkan kurva *Receiver Operating Characteristic* (ROC) dari empat model ML yang digunakan dalam penelitian ini, hal ini digunakan untuk menggambarkan kemampuan model dalam membedakan antara kelas risiko putus sekolah rendah dan risiko putus sekolah tinggi berdasarkan hubungan antara *True Positive Rate* (TPR) dan *False Positive Rate* (FPR).



Gambar 2. Receiver Operating Characteristic Keseluruhan Model
Sumber: Hasil Pengujian Dengan Google Colab

Berdasarkan Gambar 3, RF dan NB menghasilkan nilai AUC sebesar 1.00. Nilai tersebut menunjukkan bahwa kedua model memiliki kemampuan yang sangat baik dalam membedakan kedua kelas pada data pengujian. Kurva ROC kedua model berada jauh di atas garis diagonal dan mendekati sudut kiri atas, yang mengindikasikan kemampuan diskriminasi yang sangat tinggi. Sementara itu, DT dan LR menghasilkan nilai AUC sebesar 0.75. Meskipun nilai tersebut lebih rendah dibandingkan RF dan NB, kedua model masih menunjukkan kemampuan klasifikasi yang cukup baik dalam membedakan kelas risiko putus sekolah rendah dan tinggi. Selain itu, pola kurva ROC yang dihasilkan oleh DT dan LR relatif serupa, sejalan dengan hasil confusion matrix dan nilai evaluasi yang menunjukkan performa yang hampir sama.

Secara keseluruhan, hasil ROC curve memperlihatkan bahwa RF memiliki kemampuan klasifikasi yang paling baik dibandingkan model lainnya. NB juga menunjukkan kemampuan diskriminasi yang sangat baik meskipun masih terdapat satu kesalahan klasifikasi pada confusion matrix. Di sisi lain, DT dan LR menghasilkan kemampuan klasifikasi yang cukup baik, namun belum sebaik RF dan NB. Perlu diperhatikan bahwa jumlah data pengujian yang relatif kecil dapat mempengaruhi bentuk kurva ROC dan nilai AUC yang diperoleh, sehingga hasil ini perlu dikaji lebih lanjut menggunakan dataset yang lebih besar untuk memperoleh gambaran performa model yang lebih representatif.

4.4 Analisis Uji McNemar

Untuk mengetahui apakah terdapat perbedaan performa yang signifikan antar model, dilakukan pengujian menggunakan McNemar's Test. Hasil pengujian ditunjukkan pada Tabel 2.

Tabel 2. Hasil Uji McNemar

Perbandingan Model	p-value	Kesimpulan
Random Forest vs Decision Tree	0.5000	Tidak terdapat perbedaan signifikan

Random Forest vs Naive Bayes	1.0000	Tidak terdapat perbedaan signifikan
Random Forest vs Logistic Regression	0.5000	Tidak terdapat perbedaan signifikan
Decision Tree vs Naive Bayes	1.0000	Tidak terdapat perbedaan signifikan
Decision Tree vs Logistic Regression	1.0000	Tidak terdapat perbedaan signifikan
Naive Bayes vs Logistic Regression	1.0000	Tidak terdapat perbedaan signifikan

Sumber: Hasil Pengujian Dengan Google Colab

Hasil pengujian menunjukkan bahwa seluruh pasangan model memiliki nilai p-value lebih besar dari 0.05. Dengan demikian, tidak terdapat perbedaan performa yang signifikan secara statistik antara model yang dibandingkan. Meskipun RF menghasilkan nilai evaluasi yang lebih tinggi, keunggulan tersebut belum dapat dinyatakan berbeda secara signifikan dibandingkan model lainnya. Hal ini kemungkinan disebabkan oleh jumlah data pengujian yang relatif kecil sehingga variasi performa antar model menjadi tidak terlalu terlihat.

4.5 Analisis Uji SHAP

Untuk mengetahui faktor-faktor yang paling berpengaruh terhadap risiko putus sekolah, dilakukan interpretasi model menggunakan SHAP, dan nilai rata-rata masing-masing variabel ditunjukkan pada Tabel 3.

Tabel 2. Hasil Uji SHAP

Ranking	Fitur	Mean SHAP
1	Mengulang	0.079391
2	Rombongan Belajar	0.058245
3	Kepala Sekolah dan Guru (<S1)	0.056848
4	Kepala Sekolah dan Guru ($\geq$ S1)	0.051378
5	Tenaga Kependidikan (SM)	0.043282
6	Siswa	0.040022
7	Sekolah	0.029166
8	Ruang kelas (baik)	0.026181
9	Tenaga Kependidikan (>SM)	0.020213
10	Ruang kelas (rusak berat)	0.018395
11	Ruang kelas (rusak sedang)	0.015403
12	Ruang kelas (rusak ringan)	0.006124

Sumber: Hasil Pengujian Dengan Google Colab

Berdasarkan hasil interpretasi SHAP, variabel Mengulang memiliki nilai rata-rata SHAP tertinggi sebesar 0.079391 sehingga menjadi faktor yang paling berpengaruh terhadap prediksi risiko putus sekolah. Temuan ini menunjukkan bahwa tingginya jumlah siswa yang mengulang kelas memiliki keterkaitan yang kuat dengan peningkatan risiko putus sekolah. Variabel Rombongan Belajar menempati urutan kedua dengan nilai SHAP sebesar 0.058245, diikuti oleh variabel Kepala Sekolah dan Guru (<S1) sebesar 0.056848 dan Kepala Sekolah dan Guru ($\geq$ S1) sebesar 0.051378. Hasil tersebut menunjukkan bahwa kualitas sumber daya manusia di sekolah memiliki peran penting dalam mempengaruhi tingkat risiko putus sekolah. Selain itu, jumlah tenaga kependidikan dan jumlah siswa juga memberikan kontribusi yang cukup besar terhadap

hasil prediksi. Sementara itu, kondisi sarana dan prasarana seperti ruang kelas rusak berat, rusak sedang, maupun rusak ringan memiliki kontribusi yang relatif lebih kecil dibandingkan variabel akademik dan sumber daya manusia.

Secara keseluruhan, hasil interpretasi SHAP menunjukkan bahwa faktor akademik dan kualitas sumber daya pendidikan memiliki pengaruh yang lebih dominan dibandingkan kondisi fisik sekolah dalam menentukan risiko putus sekolah tingkat provinsi di Indonesia.

5 KESIMPULAN DAN SARAN

Penelitian ini telah membandingkan kinerja empat model ML, yaitu RF, DT, NB, dan LR dalam memprediksi risiko putus sekolah tingkat provinsi di Indonesia. Berdasarkan hasil pengujian, RF menghasilkan performa terbaik dengan nilai *Accuracy*, *Precision*, *Recall*, dan *F1-Score* sebesar 1.0000. NB berada pada urutan kedua dengan *Accuracy* sebesar 0.8750 dan *F1-Score* sebesar 0.8571, sedangkan DT dan LR menghasilkan performa yang sama dengan *Accuracy* sebesar 0.7500 dan *F1-Score* sebesar 0.7500. Hasil analisis ROC curve juga menunjukkan bahwa RF dan NB memiliki kemampuan diskriminasi yang sangat baik dengan nilai AUC sebesar 1.00, sedangkan DT dan LR memperoleh nilai AUC sebesar 0.75.

Hasil uji McNemar menunjukkan bahwa tidak terdapat perbedaan yang signifikan secara statistik antar model yang dibandingkan. Meskipun RF menghasilkan performa terbaik berdasarkan metrik evaluasi, keunggulan tersebut belum dapat dinyatakan berbeda secara signifikan dibandingkan model lainnya. Kondisi ini menunjukkan bahwa seluruh model memiliki kemampuan yang relatif sebanding pada dataset yang digunakan.

Analisis menggunakan SHAP menunjukkan bahwa variabel Mengulang merupakan faktor yang paling berpengaruh terhadap prediksi risiko putus sekolah, diikuti oleh Rombongan Belajar, Kepala Sekolah dan Guru ($<S1$), serta Kepala Sekolah dan Guru ($\geq S1$). Temuan tersebut mengindikasikan bahwa faktor akademik dan kualitas sumber daya manusia di sekolah memiliki pengaruh yang lebih dominan dibandingkan faktor sarana dan prasarana dalam menentukan risiko putus sekolah tingkat provinsi di Indonesia. Secara keseluruhan, hasil penelitian ini menunjukkan bahwa pendekatan ML dapat digunakan sebagai alat bantu untuk mengidentifikasi risiko putus sekolah dan memberikan informasi mengenai faktor-faktor yang berkaitan dengan kondisi tersebut.

Penelitian ini masih memiliki beberapa keterbatasan yang dapat menjadi perhatian untuk penelitian selanjutnya. Pertama, jumlah data yang digunakan relatif terbatas sehingga hasil yang diperoleh masih perlu divalidasi menggunakan dataset dengan cakupan yang lebih besar dan periode pengamatan yang lebih panjang. Kedua, penelitian berikutnya dapat mempertimbangkan penambahan variabel lain yang berkaitan dengan kondisi sosial, ekonomi, dan demografi agar model yang dihasilkan mampu menggambarkan faktor-faktor yang mempengaruhi risiko putus sekolah secara lebih komprehensif. Ketiga, penelitian selanjutnya dapat mengembangkan model dengan menggunakan algoritma ML lainnya maupun pendekatan *Deep Learning* untuk memperoleh hasil prediksi yang lebih baik. Selain itu, penggunaan metode interpretasi lain dapat dipertimbangkan untuk memberikan pemahaman yang lebih mendalam mengenai faktor-faktor yang mempengaruhi risiko putus sekolah sehingga hasil penelitian dapat memberikan dukungan yang lebih baik dalam penyusunan kebijakan pendidikan berbasis data.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Terbuka atas kesempatan yang diberikan untuk mempresentasikan dan mempublikasikan hasil penelitian ini. Apresiasi juga

disampaikan atas dukungan serta kerja sama akademik yang terjalin dengan Universitas Bina Nusantara dan Universitas Mercu Buana dalam mendukung pelaksanaan dan penyelesaian penelitian ini. Kolaborasi antar perguruan tinggi tersebut diharapkan dapat terus mendorong pengembangan penelitian dan inovasi di bidang teknologi informasi.

DAFTAR PUSTAKA

- Alifa, V. N. (2023). Analisis faktor penyebab meningkatnya angka putus sekolah di Indonesia pada tahun 2022. *Jurnal Pendidikan Sultan Agung*, 3(2), 175–182. <https://doi.org/10.30659/jp-sa.3.2.175-182>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Chung, M. K. (2020). Introduction to logistic regression (arXiv:2008.13567). arXiv. <https://doi.org/10.48550/arXiv.2008.13567>
- Dermol, V., & Skrbinjek, V. (2026). Predicting student dropout in higher education using Random Forest and survival analysis. *Discover Education*, 5(1), Artikel 1710. <https://doi.org/10.1007/s44217-026-01710-8>
- Elbasi, E., Nadeem, M., Alzoubi, Y. I., Topcu, A. E., & Varghese, G. (2025). Machine learning in education: Innovations, impacts, and ethical considerations. *IEEE Access*, 13, 128741–128770. <https://doi.org/10.1109/ACCESS.2025.3590134>
- Habibi, M. (2025). Machine learning dan AI dalam GIS modern. *TechnoGIS*. <https://www.technogis.co.id/machine-learning-dan-ai-dalam-gis-modern/>
- Izza, Y., Ignatiev, A., & Marques-Silva, J. (2020). On explaining decision trees (arXiv:2010.11034). arXiv. <https://doi.org/10.48550/arXiv.2010.11034>
- Kadrina, L. (2025). Pendidikan sebagai pondasi kemajuan bangsa. *Maliki Interdisciplinary Journal*, 3(6), 905–912. <https://urj.uin-malang.ac.id/index.php/mij/article/view/14623>
- Kirana, K. C., Wibawanto, S., Herwanto, H. W., Irianto, W. S. G., Hidayat, W. N., & Aqshal, M. (2024). Naive Bayes-based drop out recommendation system in vocational college. *Proceedings of the 5th Vocational Education International Conference (VEIC-5 2023)*, 727–733. Atlantis Press. https://doi.org/10.2991/978-2-38476-198-2_99
- Louppe, G. (2014). Understanding random forests: From theory to practice (arXiv:1407.7502). arXiv. <https://doi.org/10.48550/arXiv.1407.7502>
- Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions (arXiv:1705.07874). arXiv. <https://doi.org/10.48550/arXiv.1705.07874>
- Mariano, A. M., Ferreira, A. B. M. L., Santos, M. R., Castilho, M. L., & Bastos, A. C. F. L. C. (2022). Decision trees for predicting dropout in engineering course students in Brazil. *Procedia Computer Science*, 214, 1113–1120. <https://doi.org/10.1016/j.procs.2022.11.285>
- Matzavela, V., & Alepis, E. (2021). Decision tree learning through a predictive model for student academic performance in intelligent m-learning environments. *Computers and Education: Artificial Intelligence*, 2, Artikel 100035. <https://doi.org/10.1016/j.caeai.2021.100035>
- Nadila, A., Suratno, Ratumbusang, M. F. N. G., & Permatasari, M. A. (2025). Pengaruh tingkat kemiskinan terhadap angka putus sekolah pada jenjang SD, SMP, dan SMA di Kota Banjarbaru. *Jurnal Pendidikan Ekonomi (JUPE)*, 13(3), 234–247. <https://doi.org/10.26740/jupe.v13n3.p234-247>
- Pastika, P. B. (2024). Dataset pendidikan SD Indonesia tahun 2023-2024 [Dataset]. Kaggle. <https://www.kaggle.com/datasets/puanbeningpastika/dataset-pendidikan-sd-indonesia-2023-2024>

- Peretz, O., Koren, M., & Koren, O. (2024). Naive Bayes classifier – an ensemble procedure for *Recall* and *Precision* enrichment. *Engineering Applications of Artificial Intelligence*, 136, Artikel 108972. <https://doi.org/10.1016/j.engappai.2024.108972>
- Ponce-Bobadilla, A. V., Schmitt, V., Maier, C. S., Mensing, S., & Stodtmann, S. (2024). Practical guide to SHAP analysis: Explaining supervised machine learning model predictions in drug development. *Clinical and Translational Science*, 17(11), Artikel e70056. <https://doi.org/10.1111/cts.70056>
- Prayoga, A. A., Hasanuddin, M., Khodijah, S., & Rizki, C. A. (2025). Analisis penerapan machine learning dalam sistem prediksi dan pengambilan keputusan. *Journal of Electrical Engineering Research*, 1(3), 84–90. <https://doi.org/10.64803/joeer.v1i3.19>
- Santos, R. F., & Berthet, M. (2026). Predicting graduation and dropout rates among engineering students using AI techniques. *Discover Education*, 5, Artikel 69. <https://doi.org/10.1007/s44217-025-00069-0>
- Singh, H. P., & Alhulail, H. (2022). Predicting student-teachers dropout risk and early identification: A four-step logistic regression approach. *IEEE Access*, 10, 6470–6482. <https://doi.org/10.1109/ACCESS.2022.3141992>
- Temu, C. C., Tolok, M. S., Azmi, P. V., & Marsisno, W. (2018). Faktor-faktor yang memengaruhi putus sekolah usia SMA di Provinsi NTT tahun 2016. *Semnas Official Statistics*, 2018(1), 226–234. <https://prosiding.stis.ac.id/index.php/semnasoffstat/article/view/158>
- Trujillo, F., Pozo, M., & Suntaxi, G. (2025). Artificial intelligence in education: A systematic literature review of machine learning approaches in student career prediction. *Journal of Technology and Science Education*, 15(1), 162–185. <https://doi.org/10.3926/jotse.3124>
- Wu, J. (2020). Machine learning in education. *2020 International Conference on Modern Education and Information Management (ICMEIM)*, 56–63. <https://doi.org/10.1109/ICMEIM51375.2020.00020>