

ANALISIS DATA PENERIMAAN PESERTA DIDIK BARU MENGUNAKAN *CROSS VALIDATION* DAN ALGORITMA *DECISION TREE* DI SMA NEGERI 1 BANDUNG

Wini Andriyani¹, Rudi Kurniawan², Yudhistira Arie Wijaya³

^{1,3} Program Studi Teknik Informatika, STMIK IKMI Cirebon

² Program Studi Rekayasa Perangkat Lunak, STMIK IKMI Cirebon

Jln Perjuangan No 10 B Karyamulya Kesambi Kota Cirebon Jawa Barat Indonesia

winiandriyani088@gmail.com

ABSTRAK

Setiap tahun, SMA Negeri 1 Bandung menerima calon peserta didik baru melalui tiga jalur pendaftaran: penelusuran minat, Prestasi Akademik (PA), dan ujian tertulis. Namun, proses seleksi dalam Penerimaan Peserta Didik Baru (PPDB) menghadapi kendala dalam menetapkan kriteria kelulusan. Untuk mengatasi masalah ini, diperlukan prediksi data guna meningkatkan kualitas penerimaan siswa. Tujuan penelitian ini adalah menganalisis data PPDB untuk menyederhanakan proses seleksi calon siswa dengan menerapkan teknik *data mining*, khususnya algoritma *Decision Tree*. Metode ini menghasilkan pohon keputusan yang membantu mengidentifikasi siswa yang lolos dan tidak lolos dengan lebih efisien. Hasil penelitian menggunakan aplikasi *RapidMiner Studio* menunjukkan bahwa dari 282 data, sebanyak 86 data tidak lolos dan 191 data lolos, dengan tingkat akurasi mencapai 98.24%, precision 97.45%, dan recall 94.51%. Pemahaman faktor-faktor yang mempengaruhi keberhasilan akademik siswa juga dapat membantu pengembangan strategi pendidikan yang lebih efektif. Melalui teknik *Cross-Validation*, validitas model *Decision Tree* diuji untuk memastikan kinerjanya yang tinggi dalam berbagai lingkungan. Dengan demikian, penelitian ini berpotensi memberikan kontribusi yang signifikan terhadap kebijakan penerimaan siswa baru yang efisien dan berkualitas di SMA Negeri 1 Bandung.

Kata kunci : *Cross validation, Data mining, Decision tree, PPDB, Faktor keberhasilan akademik siswa*

1. PENDAHULUAN

Penerimaan Peserta Didik Baru (PPDB) di SMA Negeri 1 Bandung melibatkan jalur Penelusuran Minat dan Prestasi Akademik (PMPA) serta jalur ujian tertulis. Seleksi jalur PMPA didasarkan pada prestasi akademik dan non-akademik, dengan rekomendasi dari Kepala Sekolah peserta didik SMP/MTs [1]. Setiap tahun, PPDB menghadapi tantangan di mana calon siswa kesulitan memilih jurusan, seringkali disebabkan oleh pengaruh luar seperti saran orang tua, tanpa mempertimbangkan potensi diri, berdampak negatif pada hasil akademis dan peluang kerja [2].

PPDB SMA Negeri 1 Bandung terkait erat dengan faktor-faktor seperti tenaga pendidikan, peserta, sekolah, kuota, dan nilai peserta (Juman, 2021). Kendala muncul dalam pemilihan siswa baru karena ketidak-konsistenan dan sifat subjektif penilaian. Oleh karena itu, diperlukan metode identifikasi calon siswa baru di sekolah [3].

Faktor penting yang mendukung prestasi belajar siswa adalah manajemen pembelajaran, kerjasama dengan dunia kerja, kualitas guru, serta peran orang tua dan siswa [4]. Pemahaman terhadap faktor-faktor ini krusial untuk meningkatkan mutu pendidikan. Kondisi di mana siswa memilih jurusan sesuai minatnya pada SMA memiliki dampak besar terhadap aktivitas dan hasil akademik mereka [5]. Penelitian ini menggunakan algoritma klasifikasi untuk menghubungkan atribut-atribut tersebut dengan prestasi akademik, menunjukkan bahwa faktor-faktor ini berpengaruh signifikan terhadap keberhasilan

akademik siswa, memberikan panduan berharga bagi SMA Negeri 1 Bandung dalam meningkatkan kualitas pendidikan dan hasil akademik siswa.

Memfaatkan nilai UN SMPN 9 Jakarta tahun ajaran 2011/2012 dan *passing grade* SMAN 99 Jakarta pada tahun-tahun sebelumnya, penelitian ini menggunakan teknik *data mining* dengan metode *decision tree* untuk mengungkap hasil seleksi masuk SMAN 99 Jakarta [6]. *Decision tree*, sebuah metode klasifikasi dan prediksi yang efektif, mampu mengubah data besar menjadi pohon keputusan yang merepresentasikan aturan dengan bahasa alami, memungkinkan identifikasi hubungan tersembunyi antara variabel input dan target [7]. Struktur pohon keputusan memecah data besar menjadi himpunan *record* kecil melalui serangkaian aturan keputusan.

Penelitian sebelumnya menunjukkan bahwa banyak faktor, seperti tingkat kehadiran, dukungan keluarga, dan metode belajar, dapat mempengaruhi kinerja siswa. Namun, penelitian ini memperluas wawasan dengan menggunakan metode klasifikasi data untuk secara sistematis mengidentifikasi faktor-faktor yang paling mempengaruhi keberhasilan akademik siswa, mengisi kesenjangan pengetahuan yang ada.

Penelitian ini bertujuan untuk memprediksi data PPDB SMA Negeri 1 Bandung dengan menggunakan algoritma *decision tree*, yang fokus pada pohon keputusan untuk memberikan hasil yang akurat dan spesifik terhadap data polinomial. Tujuan utamanya adalah memberikan landasan ilmiah untuk perbaikan

kebijakan pendidikan, dengan analisis data pendaftaran siswa menggunakan metode klasifikasi untuk mengidentifikasi faktor-faktor yang memengaruhi keberhasilan akademik.

Dengan pemahaman yang lebih baik terhadap faktor-faktor tersebut, institusi pendidikan dan pemangku kepentingan dapat mengembangkan strategi yang lebih efektif untuk meningkatkan keberhasilan akademik siswa. Penelitian ini juga menggunakan teknik *Cross-validation* untuk evaluasi kinerja model, umumnya digunakan dalam *machine learning*, guna mengukur generalisasi model ke data baru.

Hasil penelitian menunjukkan bahwa dari 282 data, 191 siswa berhasil dan 86 siswa tidak lolos dalam PPDB. Pemahaman mendalam terhadap faktor-faktor yang memengaruhi keberhasilan siswa pada penerimaan peserta didik baru memberikan kontribusi positif bagi calon siswa yang lebih efektif. Implikasi praktis penelitian ini melibatkan peningkatan kualitas penerimaan siswa melalui prediksi akurat, memberikan saran kepada calon siswa dan orang tua untuk mempertimbangkan pilihan mereka sebelum mendaftar, serta membantu menghindari kesalahan pada saat mendaftar PPDB SMA Negeri 1 Bandung.

2. TINJAUAN PUSTAKA

Dari hasil *literature review* terhadap jurnal-jurnal terkait Analisis Data Penerimaan Peserta Didik Baru (PPDB), dapat disimpulkan berbagai pendekatan dan metode yang digunakan untuk meningkatkan efektivitas dan akurasi dalam proses PPDB serta pengambilan keputusan di bidang pendidikan. [3] mengaplikasikan algoritma C4.5 pada sampel data dari SMK Hasanah tahun 2018/2019 untuk memprediksi penerimaan siswa baru. Rachmadi menyarankan pendekatan *Waterfall* dan Algoritma *Decision Tree* sebagai dukungan pengambilan keputusan pemilihan jurusan bagi siswa baru. [1] fokus pada prediksi kelulusan PMPA dengan memanfaatkan metode data mining, khususnya algoritma *Decision Tree* dan C4.5.

D. Yusuf, Y. Mubarak, A. R. Pangesti, N. Wulansari, dan R. Zulqornain melakukan perbandingan metode *Naive Bayes* dan *Decision Tree* C4.5 untuk membantu pemilihan jurusan siswa [8]. Y. S. Eirlangga dan A. E. Syaputra mengimplementasikan algoritma C4.5 untuk menentukan jurusan siswa dengan akurat dan efektif melalui identifikasi masalah, analisis masalah, studi literatur, pengumpulan data, dan pengujian hasil [9]. Yulhendri and K. K. Juman mengadopsi tiga algoritma klasifikasi (*Support Vector Machine*, *Naive Bayes*, dan *Decision Tree*) untuk memprediksi penerimaan siswa baru di SMA [10].

J. Triwidianti, F. Y. A. Alfian, dan M. Prasojo memaparkan bahwa studi pustaka, observasi/survei, dan berbagai metode *data mining*, seperti *Decision Tree*, *Random Forest*, *Naive Bayes*, dapat digunakan untuk memprediksi kinerja siswa [4].

Ramadhani & Hendriyani menekankan penggunaan *Decision Tree* untuk memprediksi prestasi siswa berdasarkan faktor-faktor sosial ekonomi, disiplin, dan motivasi siswa [11]. M. Nasution menggunakan metode *Classification And Regression Tree* (CART) untuk memprediksi jumlah pendaftaran siswa baru [12]. D. Kurniawan and A. Saputra membahas tantangan dan isu-isu terkait sistem zonasi untuk penerimaan siswa baru dan mengusulkan metode *K-Nearest Neighbor* (K-NN) sebagai solusi [13].

H. D. Prasetyo, T. Pramiyati, dan I. N. Isnainiyah menerapkan algoritma *Naive Bayes* untuk menganalisis sentimen pengguna *Twitter* terhadap kebijakan Merdeka Belajar [14]. F. Achsan, M. A. MF, dan E. A. R. Rasyid menggunakan *big data* Dapodik dan algoritma C4.5 melalui e-Zoni untuk memprediksi kelulusan calon siswa pada PPDB Sistem Zonasi [15]. A. Nilogiri dan R. Umilasari memanfaatkan metode *K-Nearest Neighbor* (KNN) untuk mengklasifikasikan jurusan siswa di SMA Negeri 2 Situbondo [16].

S. Rizal, P. Studi, T. Informatika, dan U. Yudharta mengimplementasikan algoritma *Naive Bayes* untuk memprediksi penerimaan siswa baru di SMK Al-Amien Wonorejo [17]. S. P. Nabila melakukan prediksi kelulusan tepat waktu mahasiswa berdasarkan data registrasi menggunakan algoritma [18]. T. M. S. Mulyana menggunakan algoritma *K-Nearest Neighbor* untuk mengklasifikasikan jurusan siswa baru berdasarkan data seperti nama siswa, jenis kelamin, minat, asal sekolah, posisi dalam keluarga, jumlah saudara, dan nilai ujian [19].

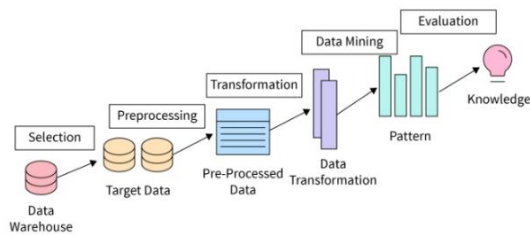
Berdasarkan hasil *literature review*, penggunaan metode *data mining* seperti algoritma C4.5, *Decision Tree*, dan *Random Forest* terbukti signifikan dalam meningkatkan efektivitas proses penerimaan peserta didik baru di berbagai institusi pendidikan. Praktik umum pengumpulan data, melibatkan partisipasi langsung dari sekolah atau institusi, digunakan untuk prediksi kelulusan, penentuan jurusan siswa, dan seleksi jalur prestasi akademik. Metode *Waterfall* dalam pengembangan sistem informasi memberikan pendekatan sistematis dalam mendukung pengambilan keputusan terkait pemilihan jurusan. Penerapan beragam algoritma *data mining*, seperti kombinasi *Knowledge Discovery in Databases* (KDD), *Decision Tree*, dan Algoritma C4.5, mencerminkan upaya meningkatkan kompleksitas dan akurasi prediksi, dengan hasil *literature review* menunjukkan variasi permasalahan yang diselesaikan, mulai dari pemilihan jurusan hingga kinerja siswa. Pendekatan ini memberikan kontribusi pada peningkatan efisiensi dan kualitas keputusan dalam konteks penerimaan peserta didik baru.

3. METODE PENELITIAN

3.1. Metode Penelitian

Gambar 1 menunjukkan penggunaan Metode Penelitian *Knowledge Discovery in Databases* (KDD) dengan Algoritma *Decision Tree* pada Analisis Data

PPDB SMA Negeri 1 Bandung, memberikan gambaran visual langkah-langkah penelitian prediksi penerimaan peserta didik baru di sekolah tersebut.



Gambar 1. Tahapan Metode Penelitian

Tabel 1. Tahapan Metode Penelitian

Tahapan	Aktivitas	Deskripsi Aktivitas
Selection	Pemilihan Data	Memilih <i>dataset</i> yang sesuai dengan tujuan penelitian dan kebutuhan analisis di SMA Negeri 1 Bandung.
Pre-processing	Pembersihan Data	Menangani <i>missing values</i> , <i>outliers</i> , dan <i>noise</i> untuk memastikan data yang digunakan dalam analisis berkualitas tinggi.
Transformation	Transformasi Data	Mengubah format atau struktur data agar sesuai dengan kebutuhan analisis, seperti normalisasi, pengkodean ulang, atau restrukturisasi.
Data Mining	Proses Data Mining	Menerapkan algoritma data mining untuk mengeksplorasi pola atau hubungan yang terdapat dalam data.
Evaluation	Evaluasi Model	Melakukan pengujian dan evaluasi kinerja model <i>data mining</i> terhadap dataset yang berbeda untuk memastikan validitas dan akurasi.

Tabel 1 merupakan detail aktivitas dari setiap tahapan metode penelitian yang dilaksanakan.

3.2. Sumber Data

Sumber data penelitian ini menggunakan data sekunder dimana merupakan data yang diperoleh dari sumber yang sudah ada yaitu melalui *website* opendata.jabarprov.go.id. *Dataset* diunduh dari *website* *Open Data Jabar* pada tanggal 5 Oktober 2023. Pengambilan data dilakukan secara manual dengan melakukan pengunduhan data yang terdapat di *Open Data Jabar*.

3.3. Populasi dan Sampel

Populasi penelitian mencakup pendaftar yang lulus dan tidak lulus pada Penerimaan Peserta Didik Baru di SMA Negeri 1 Bandung. Dari populasi tersebut, diambil sampel sebanyak 282 data yang teridentifikasi dan terukur dengan atribut "NISN, Nama, Sekolah Asal, Sekolah Tujuan, Zonasi, Poin, Jarak (meter), Usia, dan Status" untuk keperluan penelitian.

3.4. Teknik Pengumpulan Data

Teknik pengumpulan data pada penelitian ini adalah data sekunder yang diambil dari *website* resmi pemerintah yaitu *Open Data Jabar* melalui *download* pada *link* <https://opendata.jabarprov.go.id>.

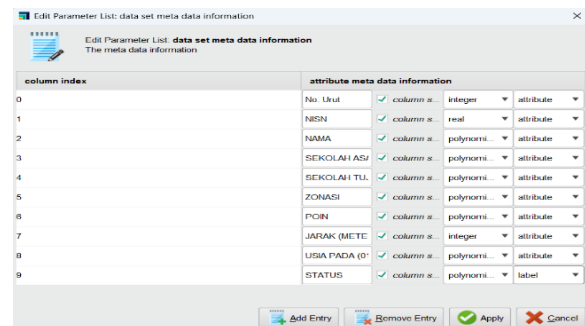
3.5. Teknik Analisis Data

Teknik analisis data dalam penelitian ini mengikuti metode penelitian KDD, yang melibatkan tahapan perancangan dari awal sampai akhir. Proses ini mencakup langkah-langkah sistematis untuk mendapatkan wawasan dan informasi yang berharga dari data yang ada.

4. HASIL DAN PEMBAHASAN

4.1. Data Selection

Pada *data selection* dilakukan pemilihan data yang akan diolah untuk kebutuhan penelitian ini mencakup atribut yang tersedia pada *dataset* Penerimaan Peserta Didik Baru (PPDB) di SMA Negeri 1 Bandung, terdiri dari 10 atribut.



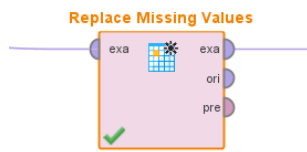
Gambar 2. Data Selection

Berdasarkan data yang tersedia, dipilih atribut status sebagai label untuk membentuk model *RapidMiner* menggunakan metode *decision tree*, seperti yang terlihat pada Gambar 2.

4.2. Pre-processing

Pada *pre-processing*, akan ditangani *missing values* untuk memastikan data yang digunakan dalam analisis tidak *null* atau kosong. Dalam *dataset*, atribut POIN memiliki 138 data dengan *missing value*. Berdasarkan atribut poin dengan *missing value*, digunakan operator *Replace Missing Values* pada tahap ini untuk mengatasi nilai *null* pada data. Berikut adalah gambar operator yang digunakan dalam *pre-processing* untuk menangani *missing value* pada data

Penerimaan Peserta Didik Baru (PPDB) di SMA Negeri 1 Bandung.

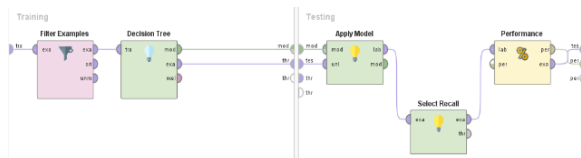


Gambar 3. *Replace Missing Values*

Pada tahap ini, operator *Replace Missing Values* seperti yang tertera pada Gambar 3 digunakan untuk mengisi atau mengganti *missing value* dengan metode imputasi seperti nilai rata-rata, median, dan modus yang disediakan oleh *tools RapidMiner*. Dengan atribut poin yang memiliki tipe nominal, digunakan metode mengganti nilai hilang dengan nilai rata-rata untuk menangani nilai yang hilang.

4.3. Data Mining

Proses data mining menggunakan operator *Cross Validation*, yang terdiri dari dua subproses: Pelatihan dan Pengujian. Evaluasi performa model diperoleh setelah melalui tahap pelatihan dan pengujian.



Gambar 4. *Cross validation*

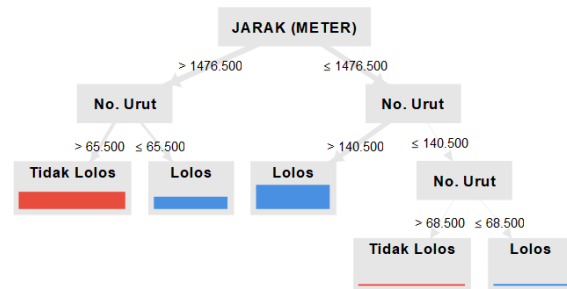
Gambar 4 merupakan proses *Cross Validation* melibatkan iterasi bergantian antara tahap pelatihan dan pengujian untuk mengevaluasi kinerja model dengan tahapan sebagai berikut:

4.3.1. Training

- Filter Example:** Tahap pra-pemrosesan data untuk pemfilteran, normalisasi, atau transformasi lainnya pada data pelatihan sebelum digunakan untuk melatih model.
- Decision Tree:** Algoritma pembelajaran mesin yang membangun model prediksi dalam bentuk struktur pohon, mempelajari pola dari data pelatihan, dan membuat keputusan atau prediksi pada data baru.
- Apply Model:** Penggunaan model yang sudah dilatih untuk memprediksi label atau nilai target pada data pengujian.
- Select Recall:** Metrik evaluasi untuk kinerja model klasifikasi, menghitung dan menampilkan nilai recall dari hasil prediksi model.
- Performance:** Evaluasi kinerja model secara keseluruhan untuk memahami seberapa baik model memprediksi label atau nilai target pada data pengujian.

4.4. Evaluation

Operator *cross-validation* dalam proses data mining menghasilkan hasil evaluasi kinerja model, terutama dengan menggunakan algoritma *decision tree* pada tahap pelatihan. Algoritma ini membagi kumpulan data besar menjadi himpunan rekaman yang lebih kecil, dengan mempertimbangkan variabel tujuan yang ditentukan melalui filterisasi data contoh.



Gambar 5. *Decision Tree*

Gambar 5 menampilkan hasil pohon keputusan dari evaluasi model menggunakan algoritma *decision tree*. Evaluasi ini berfokus pada prediksi kategori "Lolos" atau "Tidak Lolos" berdasarkan dua atribut: "Jarak" (meter) dan "No. Urut". Hasil evaluasi dijelaskan sebagai berikut:

- Ketika $JARAK > 1476.500$:
 - Jika $No. Urut > 65.500$, hasil prediksi: "Tidak Lolos" (88 sampel)
 - Jika $No. Urut \leq 65.500$, hasil prediksi: "Lolos" (61 sampel)
- Ketika $JARAK \leq 1476.500$:
 - Jika $No. Urut > 140.500$, hasil prediksi: "Lolos" (126 sampel)
 - Jika $No. Urut \leq 140.500$:
 - Jika $No. Urut > 68.500$, hasil prediksi: "Tidak Lolos" (3 sampel)
 - Jika $No. Urut \leq 68.500$, hasil prediksi: "Lolos" (4 sampel)

Hasil evaluasi tersebut memberikan gambaran bagaimana model melakukan prediksi berdasarkan atribut yang tersedia. Pada tahap pengujian, akan dievaluasi seberapa baik kinerja pemodelan dalam memprediksi data yang tidak dilihat untuk memperoleh nilai akurasi yang baik.

accuracy: 98.24% +/- 2.99% (micro average: 98.23%)

	true Lolos	true Tidak Lolos	class precision
pred. Lolos	191	5	97.45%
pred. Tidak Lolos	0	86	100.00%
class recall	100.00%	94.51%	

Gambar 6. Hasil evaluasi kinerja model klasifikasi

Gambar 6 menampilkan hasil akurasi dari proses data mining yang telah selesai. Evaluasi pengujian data menghasilkan metrik evaluasi untuk kinerja model klasifikasi. Hasil evaluasi ini menunjukkan tingkat akurasi yang sangat tinggi, sekitar 98.24%, yang menunjukkan bahwa model berhasil memprediksi

kelas "Lolos" dan "Tidak Lolos" dengan baik dalam dataset evaluasi yang diberikan. Berikut adalah penjelasan hasil evaluasi kinerja model klasifikasi:

a. *Accuracy*:

Akurasi mengukur seberapa sering model benar dalam memprediksi kelas dari semua prediksi yang dilakukan. Dalam kasus ini, akurasi adalah 98.24%, yang berarti model memprediksi dengan benar sekitar 98.24% dari semua kasus yang dievaluasi.

b. *Confusion Matrix*:

Matriks konfusi adalah tabel yang mengevaluasi kinerja model klasifikasi pada satu set data pengujian. Dalam kasus ini, terdapat empat sel:

- Lolos: Dari 196 sampel yang sebenarnya "Lolos", model memprediksi 191 dengan benar, namun salah memprediksi 5 sampel sebagai "Tidak Lolos".
- Tidak Lolos: Dari 86 sampel yang sebenarnya "Tidak Lolos", model memprediksi semuanya dengan benar (0 salah prediksi).

4.5. Pembahasan

Penggunaan Model Klasifikasi *Decision Tree* untuk menganalisis jumlah PPDB di SMA Negeri 1 Bandung menandakan keberhasilan dalam mengidentifikasi pola dan hubungan dalam data. Setting filter example digunakan untuk menyaring data sebelum dieksekusi, memudahkan pengelolaan data pada proses eksekusi. Proses *cross* validasi dilakukan untuk melatih dan menguji model, menunjukkan keakuratan, presisi, dan *recall* yang tinggi. Algoritma *Decision Tree* digunakan untuk eksplorasi data dan pemilihan model, memperkuat analisis dan prediksi. Hasil klasifikasi menunjukkan tingkat akurasi 98,24%, presisi 100%, dan *recall* 97,45%.

Penelitian sebelumnya menggunakan metode *data mining* seperti *C4.5*, *Decision Tree*, *Naive Bayes*, dan *Random Forest* untuk memprediksi penerimaan siswa baru dan kinerja siswa. Penelitian ini melibatkan pengumpulan sampel data dari berbagai institusi pendidikan dan penerapan berbagai tahapan dalam *Knowledge Discovery in Databases* (KDD). Menggunakan Model Klasifikasi *Decision Tree* dalam analisis PPDB di SMA Negeri 1 Bandung merupakan kontribusi terhadap eksplorasi metode-metode tersebut. Evaluasi kinerja model dan akurasi yang tinggi, serta penerapan langkah-langkah seperti *setting filter example*, proses *cross* validasi, dan pemilihan *recall*, memperkuat hasil analisis dan prediksi.

5. KESIMPULAN DAN SARAN

Berdasarkan analisis hasil penelitian, penggunaan metode *Decision Tree* dalam mengklasifikasikan "Lolos" dan "Tidak Lolos" pada data jumlah pendaftar PPDB SMA Negeri 1 Bandung memberikan hasil yang sangat positif. Model klasifikasi ini mampu memprediksi dengan akurasi tinggi, mencapai 98.24%, dengan *class precision* sebesar 97.45% dan *class recall* sebesar 94.51%. Dapat

disimpulkan bahwa atribut Jarak (Meter) memiliki pengaruh signifikan terhadap tingkat akurasi, seperti yang terlihat dalam *Confusion Matrix*.

Dari hasil analisis, beberapa saran dapat diberikan untuk peningkatan proses PPDB dan kebijakan pendidikan di tingkat instansi pendidikan, antara lain: mencari nilai akurasi yang berbeda dengan menambahkan algoritma *random forest* dan variasi metode *preprocessing* data, menggunakan atribut lain untuk klasifikasi, serta meningkatkan performa model klasifikasi untuk hasil evaluasi yang lebih baik, terutama dalam menangani ketidakseimbangan antara *precision* dan *recall* untuk kelas tertentu.

DAFTAR PUSTAKA

- [1] K. Agustine and A. Azka, "Prediksi Kelulusan Pmpa Di Sekolah Menengah Atas," *IJIS - Indones. J. Inf. Syst.*, vol. 5, no. 2, p. 144, 2020, doi: 10.36549/ijis.v5i2.118.
- [2] Y. Yee, N. Ariyanti, S. Puspita, and A. Hermawan, "Menentukan Jurusan Calon Siswa Sekolah Menengah Kejuruan (Smk) Implementation of Algorithm C . 45 To Determine the Major of Prospective Students Vocational Hight School (Vhs)," vol. 11, no. 2, pp. 175–184, 2023.
- [3] R. Winanjaya, F. Amir, and R. Doni, "Penerapan Data Mining Untuk Memprediksi Penerimaan Peserta Didik Baru Menggunakan Algoritma C4.5," *Pros. Semin. Nas. Ris. Inf. Sci.*, vol. 1, no. September, p. 1, 2019, doi: 10.30645/senaris.v1i0.1.
- [4] J. Triwidianti, F. Y. A. Alfian, and M. Prasojo, "Perbandingan Metode Data Mining Untuk Prediksi Prestasi Siswa Tingkat Pendidikan Menengah Kejuruan Pada Sekolah Menengah Kejuruan Negeri (SMKN 1) Gadingrejo Pringsewu Lampung," *Semin. Nas. Has. Penelit. dan Pengabd. Masy.*, vol. 1, no. Smkn 1, pp. 126–133, 2021.
- [5] S. M. K. Widya, D. Turen, L. N. Rachmadi, A. P. Wibawa, and U. Pujiyanto, "belantika Pendidikan Rekomendasi Jurusan dengan Menggunakan *Decision Tree* pada Sistem Penerimaan Peserta Didik Baru," vol. 4, no. 1, pp. 29–36, 2021.
- [6] a y Pratama and Y. Hanum, "Penerapan Teknik Data Mining Untuk Menentukan Hasil Seleksi Masuk Sman 99 Jakarta Untuk Siswa/Siswi Smpn 9 Jakarta," *J. Ilm. Teknol. Dan ...*, pp. 49–54, 2017, [Online]. Available: <http://www.ejournal.gunadarma.ac.id/index.php/teknol/article/view/1671>
- [7] I. Sujai, Purwanto, and H.Himawan, "Prediksi Hasil Penjurusan Siswa Sekolah Menengah Atas Dengan Menggunakan Algoritma *Decision Tree* C.4.5," vol. 12, no. April, pp. 42–53, 2016.
- [8] D. Yusuf, Y. Mubarak, A. R. Pangesti, N. Wulansari, and R. Zulqornain, "Perbandingan Metode *Naive Bayes Classifier* Dan *Decision Tree* C4.5 Dalam Mencari Pola Minat Pemilihan

- Jurusan Di Madrasah Aliyah (Studi Kasus:MA El-Bayan Majenang),” *J. Sist. Informasi, dan Teknol. Inf.*, vol. 2, no. 1, pp. 97–106, 2023, [Online]. Available: <https://journal-siti.org/index.php/siti/PublishedByHPTAI>
- [9] Y. S. Eirlangga and A. E. Syaputra, “Klasifikasi Penjurusan pada Sekolah Menengah Atas (SMA) dengan Metode Algoritma C4.5,” *J. Inf. dan Teknol.*, vol. 4, no. 3, pp. 160–165, 2022, doi: 10.37034/jidt.v4i3.235.
- [10] Yulhendri and K. K. Juman, “Perbandingan Tiga Algoritma Classifier Untuk Penentuan Penerimaan Peserta Didik Baru Pada Sekolah Menengah Atas,” *J. Ilm. Tek. Mesin, Elektro dan Komput.*, vol. 1, no. 3, 2021, [Online]. Available: <http://ejournal.stie-trianandra.ac.id/index.php/JURITEK/article/view/267>
- [11] Ramadhani & Hendriyani, “Jurnal Vocational Teknik Elektronika dan Informatika,” *Prediksi Prestasi Siswa Berbas. Data Min. Menggunakan Algorith. Decis. Tree (Studi Kasus SMKN 2 Padang)*, vol. 9, no. 2302–3295, pp. 11–15, 2021, [Online]. Available: <http://ejournal.unp.ac.id/index.php/voteknika/>
- [12] M. Nasution, “Penerapan Principal Component Analysis (PCA) dalam Penentuan Faktor Dominan yang Mempengaruhi Prestasi Belajar Siswa (Studi Kasus: SMK Raksana 2 Medan),” *J. Teknol. Inf.*, vol. 3, no. 1, pp. 41–48, 2019.
- [13] D. Kurniawan and A. Saputra, “Penerapan K-Nearest Neighbour dalam Penerimaan Peserta Didik dengan Sistem Zonasi,” *J. Sist. Inf. Bisnis*, vol. 9, no. 2, p. 212, 2019, doi: 10.21456/vol9iss2pp212-219.
- [14] H. D. Prasetyo, T. Pramiyati, and I. N. Isnainiyah, “Sentimen Analisis Pengguna Twitter Terhadap Kebijakan Merdeka Belajar Menggunakan Algoritma Naive Bayes,” *Semin. Nas. Mhs. Ilmu Komput. dan Apl.*, no. April, pp. 559–568, 2021.
- [15] F. Achsan, M. A. MF, and E. A. R. Rasyid, “E-Zoni: Aplikasi Prediksi Kelulusan Calon Siswa Pada Ppdb Sistem Zonasi Melalui Pemanfaatan Big Data Dapodik Nasional,” *Lomba Karya Tulis ...*, pp. 15–28, 2021, [Online]. Available: <https://journal.itelkom-sby.ac.id/lkti/article/view/112%0Ahttps://journal.itelkom-sby.ac.id/lkti/article/download/112/83>
- [16] A. Nilogiri and R. Umilasari, “Klasifikasi Penjurusan Siswa Menengah Atas Menggunakan Metode K-Nearest Neighbor,” *J. Smart Teknol.*, vol. 4, no. 1, pp. 14–23, 2022.
- [17] S. Rizal, P. Studi, T. Informatika, and U. Yudharta, “Penerapan Algoritma Naive Bayes Untuk Prediksi Penerimaan Siswa Baru Di Smk Al-Amien Wonorejo,” *Explor. IT J. Keilmuan dan Apl. Tek. Inform.*, vol. 10, no. 1, pp. 14–21, 2018, doi: 10.35891/explorit.v10i1.1671.
- [18] S. P. Nabila, “Data registrasi mahasiswa dengan algoritma fuzzy C-means dan K-nearest neighbors,” 2020.
- [19] T. M. S. Mulyana, “Implementasi Algoritma K-Nearest Neighbor Untuk Rekomendasi Smartphone,” *Petir*, vol. 16, no. 1, pp. 195–207, 2023, doi: 10.33322/petir.v16i1.1707.