Journal of Information System Research (JOSH)
ISSN 2686-228X (media online)
Volume 8, No. 1, Oktober 2026, pp 157−166
DOI 10.47065/josh.v8i1.11037
https://ejurnal.seminar-id.com/index.php/josh/

Komparasi Kinerja Decision Tree dan Random Forest untuk Menentukan Kelayakan Kandidat Peserta Lomba Robotik

A Performance Comparison of Decision Tree and Random Forest for Determining the Eligibility of Robotics Competition Candidates

Aura Nurshifa, Entin Sutinah*, Nani Agustina

Fakultas Teknik dan Informatika, Prodi Sistem Informasi, Universitas Bina Sarana Informtika, Jakarta, Indonesia
Jl. Kramat Raya No.98, RT.2/RW.9, Kwitang, Kec. Senen, Kota Jakarta Pusat, Daerah Khusus Ibukota Jakarta, Indonesia
Email: [1]19220777@bsi.ac.id, [2,*]entin.esh@bsi.ac.id, [3]nani.nna@bsi.ac.id
Email Penulis Korespondensi: entin.esh@bsi.ac.id
Received: **21/08/2026**, Revised: **24/10/2026**, Accepted: **26/10/2026**, Published**: 29/10/2026**

Abstrak−Proses seleksi kandidat peserta lomba robotik di PT. Sonnic Robotic masih mengandalkan penilaian subjektif pengajar dan belum didukung kriteria terstandarisasi, sehingga berpotensi menghasilkan keputusan yang kurang optimal. Penelitian ini bertujuan membandingkan performa algoritma Decision Tree dan Random Forest dalam mengklasifikasikan kelayakan kandidat peserta lomba robotik. Data penelitian terdiri atas 99 siswa aktif sekolah dasar dengan sembilan variabel, meliputi persentase kehadiran, nilai bulanan selama lima periode, serta kreativitas, pemahaman materi, dan kemandirian yang dinilai menggunakan rubrik skala 1–4. Data dinormalisasi menggunakan Min-Max Normalization, kemudian skor total dihitung berdasarkan bobot expert judgment untuk menentukan label Layak dan Tidak Layak dengan ambang batas (threshold) 75. Pemodelan dilakukan menggunakan Altair AI Studio dengan metode 10-fold Cross Validation. Hasil perbandingan menunjukkan bahwa Random Forest memperoleh akurasi 91,89%, weighted precision 92,79%, weighted recall 91,92%, dan weighted F1-score 91,93%. Sementara itu, Decision Tree memperoleh akurasi 90,89%, weighted precision 91,99%, weighted recall 90,25%, dan weighted F1-score 90,91%. Hasil tersebut menunjukkan bahwa Random Forest memiliki kinerja yang lebih baik berdasarkan seluruh metrik yang dibandingkan, meskipun selisih akurasi kedua algoritma relatif kecil. Namun, penelitian ini memiliki keterbatasan pada ukuran sampel yang hanya mencakup 99 siswa, sehingga keunggulan Random Forest belum dapat dinyatakan signifikan secara statistik dan generalisasi hasil masih perlu diuji menggunakan data yang lebih besar dan beragam. Kontribusi penelitian ini adalah menyediakan komparasi kinerja kedua algoritma serta mengidentifikasi pendekatan klasifikasi yang berpotensi mendukung proses seleksi kandidat lomba robotik secara lebih objektif, terukur, dan konsisten.

Kata Kunci: Decision Tree; Klasifikasi; Lomba Robotik; Random Forest; Sistem Pendukung Keputusan.

Abstract−The selection process for robotics competition candidates at PT. Sonnic Robotic still relies on instructors' subjective assessments and lacks standardized criteria, potentially leading to suboptimal decisions. This study aims to compare the performance of the Decision Tree and Random Forest algorithms in classifying the eligibility of robotics competition candidates. The dataset consists of 99 active elementary school students and nine variables, including attendance percentage, monthly scores over five periods, and creativity, subject comprehension, and independence, assessed using a 1–4 rating scale. The data are normalized using Min-Max Normalization, and total scores are then calculated based on expert judgment weights to determine the Eligible and Not Eligible labels using a threshold of 75. Modeling is conducted using Altair AI Studio with 10-fold cross-validation. The comparative results show that Random Forest achieves an accuracy of 91.89%, weighted precision of 92.79%, weighted recall of 91.92%, and a weighted F1-score of 91.93%. Meanwhile, Decision Tree achieves an accuracy of 90.89%, weighted precision of 91.99%, weighted recall of 90.25%, and a weighted F1-score of 90.91%. These results indicate that Random Forest performs better across all the evaluated metrics, although the difference in accuracy between the two algorithms is relatively small. However, this study is limited by its sample size of only 99 students. Therefore, the superiority of Random Forest cannot yet be considered statistically significant, and the generalizability of the findings needs to be further evaluated using larger and more diverse datasets. The contribution of this study is to provide a comparative evaluation of both algorithms and identify a classification approach with the potential to support a more objective, measurable, and consistent selection process for robotics competition candidates.

Keywords: Decision Tree; Classification; Robotics Competition; Random Forest; Decision Support System.

1. PENDAHULUAN

Perkembangan kecerdasan buatan dan pembelajaran mesin dalam satu dekade terakhir telah mendorong transformasi signifikan dalam proses pengambilan keputusan di berbagai sektor, termasuk pendidikan dan pengembangan sumber daya manusia [1]. Kemampuan algoritma pembelajaran mesin dalam mengidentifikasi pola dari data yang kompleks menjadikannya relevan untuk diterapkan dalam berbagai konteks klasifikasi dan seleksi berbasis kriteria yang terukur [2]. Penerapan pendekatan berbasis data dalam proses pengambilan keputusan diharapkan mampu meningkatkan objektivitas, konsistensi, dan ketepatan keputusan, khususnya pada permasalahan yang melibatkan banyak kriteria. Salah satu domain yang mulai mengadopsi pendekatan berbasis data adalah pengelolaan program ekstrakurikuler, khususnya program yang berorientasi pada kompetisi. Seleksi peserta lomba merupakan tahap kritis karena menentukan kualitas representasi suatu lembaga dalam ajang kompetitif [3]. Ketepatan proses seleksi tidak hanya berpengaruh terhadap peluang pencapaian prestasi dalam perlombaan, tetapi juga berkaitan dengan efisiensi sumber daya pelatihan, pembentukan komposisi tim, serta

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 157
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Information System Research (JOSH)
ISSN 2686-228X (media online)
Volume 8, No. 1, Oktober 2026, pp 157−166
DOI 10.47065/josh.v8i1.11037
https://ejurnal.seminar-id.com/index.php/josh/

motivasi peserta didik secara keseluruhan. Oleh karena itu, diperlukan mekanisme seleksi yang mampu mempertimbangkan berbagai aspek kemampuan peserta secara sistematis dan objektif.

PT. Sonnic Robotic merupakan lembaga penyelenggara ekstrakurikuler robotik bagi siswa sekolah dasar dengan cakupan kompetisi tingkat antarkota, provinsi, hingga internasional. Namun, proses seleksi kandidat lomba masih bergantung pada penilaian subjektif pengajar berdasarkan pengamatan selama pembelajaran. Belum adanya kriteria seleksi yang terdokumentasi menyebabkan keputusan pemilihan peserta cenderung didasarkan pada kesan umum terhadap keaktifan dan kemampuan siswa. Ketiadaan kerangka seleksi yang terstruktur berpotensi menimbulkan keputusan yang kurang optimal. Pada lomba Desember 2025, dua kandidat dengan kemampuan saling melengkapi ditempatkan pada tim berbeda, dan kekeliruan baru diketahui setelah pendaftaran selesai. Kondisi ini menunjukkan bahwa seleksi masih rentan terhadap kelalaian dan penilaian subjektif, yang berpotensi menimbulkan inkonsistensi dan bias dalam pengambilan keputusan [4].

Dalam konteks tersebut, pendekatan machine learning dapat dimanfaatkan untuk membangun model klasifikasi yang mengidentifikasi pola kelayakan kandidat berdasarkan variabel yang telah ditentukan. Di antara berbagai algoritma klasifikasi, Decision Tree dan Random Forest merupakan dua metode yang dapat digunakan untuk menangani klasifikasi biner serta memiliki karakteristik berbeda dalam membentuk model prediksi [5]. Decision Tree bekerja dengan membangun struktur pohon keputusan berdasarkan nilai atribut data secara rekursif sehingga menghasilkan aturan keputusan yang relatif mudah diinterpretasikan [6]. Decision Tree dikenal mudah dipahami dan diterapkan, namun cenderung rentan terhadap overfitting pada dataset yang besar [7]. Sementara itu, Random Forest merupakan metode ensemble yang membangun sejumlah pohon keputusan secara paralel dan menggabungkan hasil prediksi dari pohon-pohon tersebut untuk menghasilkan prediksi yang lebih stabil dan akurat [8].

Sejumlah penelitian terdahulu menunjukkan bahwa Decision Tree dan Random Forest memiliki kemampuan yang baik dalam menyelesaikan berbagai permasalahan klasifikasi. Sebagai contoh, penelitian yang dilakukan oleh Dana et all pada tahun (2024) membandingkan Decision Tree dan Random Forest menggunakan Grid Search dan SMOTE untuk memprediksi risiko stroke. Random Forest menghasilkan akurasi 98%, lebih tinggi dibandingkan Decision Tree sebesar 96%, menunjukkan potensi optimasi parameter dalam meningkatkan kinerja klasifikasi.dan keterbatasan data penelitian [9]. Penelitian lain yang dilakukan oleh Adrian et all pada tahun (2025) membandingkan Logistic Regression dan Random Forest menggunakan NearMiss v2 untuk klasifikasi kelayakan kredit. Random Forest unggul pada data non-linear dengan akurasi 95%, F1-score 0,96, dan AUC-ROC 0,98. Penelitian ini relevan karena sama-sama menggunakan klasifikasi biner, yaitu kategori Layak dan Tidak Layak [10]. Pada penelitian lainnya, Setiono pada tahun (2022) Penelitian membandingkan Decision Tree, Random Forest, SVM, dan K-NN menggunakan Grid Search CV untuk memprediksi kepuasan penumpang maskapai. Random Forest memperoleh akurasi tertinggi sebesar 96%, menunjukkan kinerja klasifikasi yang kompetitif dibandingkan algoritma lainnya [11].

Pada domain pendidikan, penelitian Rahmallia pada tahun (2024) mengembangkan sistem pendukung keputusan berbasis Random Forest untuk mengidentifikasi minat dan bakat siswa SMA berdasarkan nilai akademik dan aktivitas ekstrakurikuler. Model mencapai akurasi 84,1% dan F1-score 0,913. Penelitian ini relevan karena sama-sama menggunakan variabel akademik dan non-akademik sebagai dasar pengambilan keputusan terkait siswa [12]. Selain itu, penelitian sebelumnya oleh Susilawati et all pada tahun (2025) membandingkan sistem pendukung keputusan berbasis machine learning dan rule-based. Random Forest memperoleh akurasi tertinggi sebesar 91%, diikuti SVM sebesar 87% dan Decision Tree sebesar 84%. Hasil ini menunjukkan potensi machine learning dalam meningkatkan kinerja sistem pendukung keputusan dibandingkan pendekatan berbasis aturan konvensional [13].

Penelitian lainnya oleh Haryanto et all pada tahun (2023) merancang sistem pendukung keputusan berbasis web menggunakan metode SMART untuk pemilihan peserta LKS SMK berdasarkan nilai akademik, kejuruan, praktik, dan peringkat kelas. Sistem ini membantu memberikan rekomendasi kandidat secara otomatis dan objektif melalui pemeringkatan multikriteria [3]. Namun, pendekatan yang digunakan belum menerapkan machine learning maupun menghasilkan model klasifikasi berbasis data historis. Selanjutnya, komparasi algoritma Decision Tree dan Random Forest pada domain yang berbeda secara konsisten menemukan bahwa Random Forest menghasilkan akurasi tertinggi [11], [13]. Temuan tersebut memperkuat relevansi pemilihan kedua algoritma dalam penelitian ini, meskipun domain penerapannya berbeda dari konteks lomba robotik.

Berdasarkan kajian penelitian terdahulu, terdapat sejumlah keterbatasan yang menjadi dasar research gap penelitian ini. Studi-studi sebelumnya telah menerapkan algoritma Decision Tree dan Random Forest pada berbagai domain, seperti kesehatan, kelayakan kredit, kepuasan pelanggan, serta identifikasi minat dan bakat siswa. Namun, hasil dari domain tersebut belum dapat langsung digeneralisasi pada seleksi kandidat lomba robotik yang melibatkan kombinasi kemampuan akademik dan keterampilan non-akademik. Selain itu, penelitian mengenai pemilihan peserta kompetisi cenderung menggunakan pendekatan sistem pendukung keputusan berbasis pembobotan, seperti SMART, sehingga perbandingan empiris antara algoritma klasifikasi Decision Tree dan Random Forest dalam konteks tersebut masih terbatas. Penggunaan variabel multidimensional dengan rubrik penilaian terstruktur pada lembaga ekstrakurikuler robotik juga belum banyak dibahas dalam studi-studi yang dikaji. Keterbatasan lainnya terletak pada penentuan kriteria dan label klasifikasi. Penelitian terdahulu yang dikaji belum secara konsisten menjelaskan keterlibatan praktisi domain dalam menetapkan bobot variabel dan ambang

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 158
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Information System Research (JOSH)
ISSN 2686-228X (media online)
Volume 8, No. 1, Oktober 2026, pp 157−166
DOI 10.47065/josh.v8i1.11037
https://ejurnal.seminar-id.com/index.php/josh/

batas klasifikasi. Padahal, keputusan kelayakan kandidat lomba robotik perlu mempertimbangkan bukan hanya pola yang dipelajari dari data, tetapi juga pengetahuan praktisi mengenai kemampuan peserta.

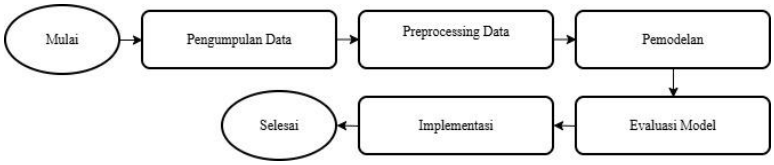
Oleh karena itu, novelty dari penelitian ini membandingkan kinerja Decision Tree dan Random Forest dengan menggunakan bobot variabel dan threshold yang ditetapkan berdasarkan expert judgment dalam konteks seleksi kandidat lomba robotik. Pendekatan ini memungkinkan evaluasi terhadap kinerja kedua algoritma dalam konteks seleksi kandidat yang spesifik, sekaligus menghubungkan hasil klasifikasi dengan kriteria penilaian yang relevan secara praktis. Meskipun demikian, penggunaan expert judgment tidak dengan sendirinya menjamin objektivitas atau validitas model, sehingga hasil penelitian tetap perlu ditafsirkan sesuai dengan karakteristik dan keterbatasan dataset yang digunakan.

Penelitian ini bertujuan untuk merancang sistem klasifikasi kelayakan kandidat lomba robotik di PT. Sonnic Robotic menggunakan algoritma Decision Tree dan Random Forest berdasarkan variabel akademik dan non-akademik yang terstruktur. Penerapan algoritma-algoritma tersebut dapat memberikan solusi yang lebih objektif dan berbasis data dalam menentukan kelayakan calon peserta [14]. Selanjutnya, penelitian ini membandingkan kinerja kedua algoritma berdasarkan metrik akurasi, presisi, recall, dan F1-score untuk mengidentifikasi perbedaan performa dalam mengklasifikasikan kandidat. Hasil perbandingan tersebut digunakan untuk menilai potensi masing-masing algoritma sebagai dasar sistem pendukung keputusan. Dengan demikian, penelitian ini diharapkan dapat menjadi bahan pertimbangan bagi PT. Sonnic Robotic dalam meningkatkan objektivitas dan konsistensi proses seleksi kandidat peserta lomba robotik melalui pemanfaatan klasifikasi berbasis data. Pendekatan ini dapat mendukung proses pengambilan keputusan yang lebih terstruktur dengan mengurangi ketergantungan pada penilaian subjektif pengajar, tanpa mengesampingkan pertimbangan profesional dalam menentukan kelayakan kandidat.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dilaksanakan melalui lima tahapan yang saling berkesinambungan, seperti pada Gambar 1.



Gambar 1. Tahapan penelitian

1. Pengumpulan Data
   Mengumpulkan data 99 siswa aktif PT. Sonnic Robotic berdasarkan sembilan variabel, meliputi persentase kehadiran, nilai bulanan periode pertama hingga kelima, kreativitas, pemahaman materi, dan kemandirian.
2. Preprocessing Data
   Melakukan pemeriksaan dan penyiapan data, normalisasi menggunakan Min-Max Normalization, serta penentuan label Layak dan Tidak Layak berdasarkan bobot penilaian dan ambang batas 75.
3. Pemodelan
   Membangun model klasifikasi menggunakan algoritma Decision Tree dan Random Forest dengan bantuan Altair AI Studio serta metode validasi silang 10-fold.
4. Evaluasi Model
   Membandingkan hasil kedua algoritma berdasarkan metrik akurasi, presisi, recall, dan F1-score untuk mengetahui kinerja masing-masing model.
5. Implementasi
   Menggunakan algoritma terpilih sebagai pendukung proses klasifikasi kelayakan kandidat peserta lomba robotik, dengan tetap mempertimbangkan evaluasi dan keterbatasan data penelitian.

Populasi penelitian terdiri atas seluruh siswa aktif jenjang sekolah dasar yang mengikuti program ekstrakurikuler robotik di PT. Sonnic Robotic, sebanyak 99 siswa. Teknik pengambilan sampel menggunakan total sampling, yaitu seluruh anggota populasi dijadikan sampel penelitian. Teknik ini dipilih karena jumlah populasi relatif kecil sehingga dapat memaksimalkan representativitas data [15]. Pemilihan total sampling memastikan seluruh variasi karakteristik siswa aktif terekam tanpa adanya bias pemilihan sampel.

Data penelitian terdiri atas sembilan variabel input, meliputi persentase kehadiran, nilai bulanan periode pertama hingga kelima, serta tiga variabel non-akademik, yaitu kreativitas, pemahaman materi, dan kemandirian. Ketiga variabel non-akademik tersebut dinilai menggunakan rubrik terstruktur dengan skala 1–4.

Tahap preprocessing dilakukan melalui dua proses utama. Pertama, normalisasi variabel rubrik berskala 1–4 menggunakan metode min-max normalization. Hasil normalisasi kemudian dikalikan 100 agar berada pada rentang yang setara dengan variabel lain yang memiliki skala 0–100. Formula normalisasi ditunjukkan pada Persamaan (1):

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 159
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Information System Research (JOSH)
ISSN 2686-228X (media online)
Volume 8, No. 1, Oktober 2026, pp 157−166
DOI 10.47065/josh.v8i1.11037
https://ejurnal.seminar-id.com/index.php/josh/

$$x' = \frac{x-1}{4-1} = (x-1)/3 \quad (1)$$

Kedua, dilakukan penghitungan skor total setiap siswa berdasarkan bobot masing-masing variabel yang ditetapkan melalui expert judgment dengan melibatkan pimpinan PT. Sonnic Robotic, sebagaimana ditampilkan pada Tabel 1. Skor total tersebut selanjutnya digunakan untuk menentukan label klasifikasi. Siswa dengan skor total ≥75 dikategorikan Layak, sedangkan siswa dengan skor <75 dikategorikan Tidak Layak. Threshold sebesar 75 ditetapkan melalui expert judgment sebagai representasi standar kompetensi minimum untuk mengikuti kompetisi robotik tingkat antarkota ke atas.

Tabel 1. Bobot Variabel Berdasarkan Expert Judgment

Variabel	Bobot
% Kehadiran	15%
Nilai Bulan 1	5%
Nilai Bulan 2	5%
Nilai Bulan 3	5%
Nilai Bulan 4	5%
Nilai Bulan 5	5%
Kreativitas	25%
Pemahaman Materi	20%
Kemandirian	15%
Total	100%

Formula skor total merupakan hasil penjumlahan perkalian antara nilai setiap variabel (rubrik telah dinormalisasi) dengan bobotnya masing-masing sebagaimana disajikan pada persamaan (2).

Skor = 0,15·Kehadiran + 0,05·(Bln1+Bln2+Bln3+Bln4+Bln5) + 0,25·Kreatif' + 0,20·Paham' + 0,15·Mandiri'

$$Skor = 0{,}15 - Kehadiran + 0{,}05(Bln1 + Bln2 + Bln3 + Bln4 + Bln5) + 0{,}25 - Kreatif' + 0{,}20 - Paham' + 0{,}15 - Mandiri' \quad (2)$$

Pemodelan dilakukan menggunakan perangkat lunak Altair AI Studio (RapidMiner) versi 2026.0.5 dengan metode evaluasi 10-fold Cross Validation, di mana data dibagi menjadi 10 partisi secara acak dan model dilatih serta diuji sebanyak 10 kali dengan partisi berbeda pada setiap iterasi. Pendekatan ini dipilih karena lebih representatif dibandingkan simple split mengingat jumlah data yang relatif terbatas. Konfigurasi kedua model disajikan pada Tabel 2.

Tabel 2. Konfigurasi Pemodelan

Parameter	Decision Tree	Random Forest
Algoritma	C4.5	Random Forest
Criterion	Gain Ratio	-
Jumlah Pohon	-	100
Metode Evaluasi	10-fold Cross Validation	10-fold Cross Validation
Software	Altair AI Studio 2026.0.5	Altair AI Studio 2026.0.5

Evaluasi performa kedua model dilakukan menggunakan operator Performance (Classification) pada Altair AI Studio yang menghasilkan confusion matrix beserta metrik akurasi, presisi, dan recall. Algoritma dengan performa terbaik selanjutnya digunakan sebagai dasar sistem pendukung keputusan dalam proses seleksi kandidat lomba robotik di PT. Sonnic Robotic, tanpa menggantikan sepenuhnya peran dan penilaian pengajar dalam keputusan akhir.

2.2 Metode Usulan

Data mining merupakan proses ekstraksi pola dan pengetahuan tersembunyi dari kumpulan data berukuran besar melalui teknik statistik, matematika, dan pembelajaran mesin [16]. Penelitian ini menggunakan pendekatan supervised learning [17], di mana model dilatih menggunakan data berlabel untuk memprediksi label pada data baru [18], dengan klasifikasi biner dua kelas: Layak dan Tidak Layak.

Decision Tree yang digunakan adalah varian C4.5 dengan criterion Gain Ratio, yang dipilih karena mampu mengatasi bias Information Gain terhadap atribut dengan jumlah nilai yang banyak [19] dan kemampuannya dalam menghasilkan keputusan yang transparan dan mudah dipahami [20]. Random Forest merupakan algoritma ensemble learning yang membangun sejumlah pohon keputusan [21] secara paralel pada subset data hasil bootstrap sampling, dengan prediksi akhir ditentukan melalui mekanisme voting mayoritas [22] sehingga lebih tahan terhadap overfitting dibandingkan Decision Tree tunggal . Dalam penelitian ini, Random Forest dibangun menggunakan 100 pohon keputusan, oleh karena itu untuk meningkatkan akurasi prediksi dan mencegah overfitting, metode ensemble seperti Random Forest dapat digunakan [23].

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 160
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Information System Research (JOSH)
ISSN 2686-228X (media online)
Volume 8, No. 1, Oktober 2026, pp 157−166
DOI 10.47065/josh.v8i1.11037
https://ejurnal.seminar-id.com/index.php/josh/

Normalisasi merupakan proses transformasi nilai variabel ke dalam rentang tertentu guna menyamakan skala antar variabel sehingga tidak terjadi dominasi variabel berskala besar terhadap hasil pemodelan [24] Evaluasi performa model menggunakan empat metrik yang diturunkan dari confusion matrix, yaitu akurasi, presisi, recall, dan F1-score, sebagaimana disajikan pada persamaan (3) hingga (6).

$$\text{Akurasi} = (TP + TN) / (TP + TN + FP + FN) \quad (3)$$

$$\text{Presisi} = TP / (TP + FP) \quad (4)$$

$$\text{Recall} = TP / (TP + FN) \quad (5)$$

$$\text{F1-Score} = 2 \times (\text{Presisi} \times \text{Recall}) / (\text{Presisi} + \text{Recall}) \quad (6)$$

Expert judgment merupakan pendekatan pengumpulan informasi yang mengandalkan penilaian dari individu yang memiliki kompetensi dan pengalaman mendalam di bidang yang relevan [25]. Dalam penelitian ini expert judgment digunakan untuk menetapkan bobot variabel dan threshold klasifikasi, dengan pihak yang dilibatkan adalah pimpinan PT. Sonnic Robotic yang memiliki pengalaman langsung dalam penyelenggaraan kompetisi robotik. Output penelitian diwujudkan sebagai bahan sistem pendukung keputusan (SPK) berupa hasil klasifikasi terstruktur yang menyajikan label kelayakan dan skor masing-masing siswa sebagai bahan pertimbangan pengajar, tanpa menggantikan peran pengajar dalam keputusan akhir.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pengumpulan dan Preprocessing Data

Pengumpulan data menghasilkan dataset sebanyak 99 siswa dengan sembilan variabel input yang mencakup dua kelompok dimensi penilaian: variabel akademik (persentase kehadiran dan nilai bulanan bulan pertama hingga kelima) dan variabel non-akademik (kreativitas, pemahaman materi, dan kemandirian) yang dinilai melalui rubrik terstruktur berskala 1-4. Cuplikan dataset disajikan pada Tabel 3.

Tabel 3. Cuplikan Dataset Siswa Ekstrakurikuler Robotik PT. Sonnic Robotic

No	% Kehadiran	Bln1	Bln2	Bln3	Bln4	Bln5	Kreatif	Paham	Mandiri	Skor	Label
1	0,9	80	82	85	83	88	4	3	4	87,73	Layak
3	0,55	65	72	70	70	58	2	2	3	50	Tidak Layak
4	0,95	70	75	85	80	93	4	4	3	89,4	Layak
7	0,85	75	72	80	79	80	3	2	3	65,38	Tidak Layak
10	0,8	75	80	73	80	80	2	3	2	58,07	Tidak Layak
...	...	...	...	...	...	...	...	...	...	...	...
92	1	85	85	85	85	85	3	3	3	76,25	Layak
99	0,8	85	85	85	90	90	3	3	4	78,75	Layak

Ketiga variabel non-akademik dinilai menggunakan rubrik terstruktur berskala 1-4 yang dikembangkan oleh peneliti berdasarkan pengamatan langsung selama kegiatan ekstrakurikuler berlangsung, sehingga penilaian dapat dilakukan secara konsisten dan terstandarisasi. Variabel kreativitas mengukur kemampuan siswa dalam menciptakan modifikasi atau desain di luar contoh yang diberikan, mulai dari skor 1 (meniru persis tanpa modifikasi) hingga skor 4 (menciptakan desain original dan inovatif tanpa diminta). Variabel pemahaman materi mengukur sejauh mana siswa memahami konsep yang diajarkan, mulai dari skor 1 (belum paham dan memerlukan penjelasan berulang) hingga skor 4 (paham penuh hingga mampu menjelaskan kembali ke teman lain). Variabel kemandirian mengukur tingkat ketergantungan siswa terhadap bimbingan pengajar dalam menyelesaikan tugas, mulai dari skor 1 (tidak dapat mengerjakan tanpa pendampingan penuh) hingga skor 4 (mampu menyelesaikan tugas secara mandiri tanpa bantuan sama sekali).

Penerapan threshold skor total ≥75 menghasilkan distribusi 47 siswa terklasifikasi Layak (47,47%) dan 52 siswa terklasifikasi Tidak Layak (52,53%), dengan proporsi yang relatif seimbang sehingga tidak diperlukan penanganan class imbalance secara khusus. Jumlah dataset sebanyak 99 sampel dengan distribusi label yang relatif seimbang ini turut mengonfirmasi bahwa penggunaan 10-fold Cross Validation sebagai metode evaluasi tepat diterapkan, mengingat setiap sampel dapat dimanfaatkan secara optimal sebagai data latih maupun data uji secara bergantian pada tiap iterasi. Sebelum pemodelan, variabel kreativitas, pemahaman materi, dan kemandirian dinormalisasi menggunakan persamaan (1) sehingga nilainya berada pada rentang 0-100, setara dengan skala variabel lain. Konfigurasi Set Role turut diterapkan pada Altair AI Studio: atribut No, Nama Siswa, dan Jenjang ditetapkan sebagai id, atribut Total Skor ditetapkan sebagai id dan dikeluarkan dari pemodelan karena merupakan turunan langsung dari variabel input, sementara atribut Label ditetapkan sebagai target klasifikasi.

3.2 Hasil Pemodelan

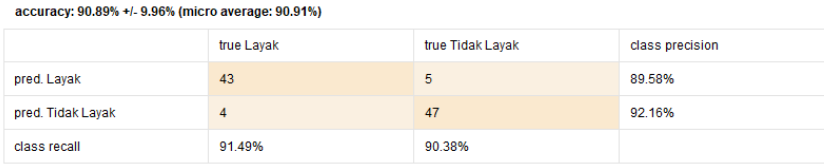
Pemodelan dilakukan untuk membangun dan mengevaluasi algoritma klasifikasi yang digunakan dalam menentukan kelayakan kandidat peserta lomba robotik.

3.2.1 Decision Tree

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 161
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Information System Research (JOSH)
ISSN 2686-228X (media online)
Volume 8, No. 1, Oktober 2026, pp 157−166
DOI 10.47065/josh.v8i1.11037
https://ejurnal.seminar-id.com/index.php/josh/

Model Decision Tree dibangun menggunakan algoritma C4.5 dengan criterion Gain Ratio dan dievaluasi menggunakan 10-fold Cross Validation. Berdasarkan hasil evaluasi, model menghasilkan confusion matrix: 43 siswa Layak diprediksi benar (True Positive), 47 siswa Tidak Layak diprediksi benar (True Negative), 5 siswa Tidak Layak diprediksi sebagai Layak (False Positive), dan 4 siswa Layak diprediksi sebagai Tidak Layak (False Negative). Hasil evaluasi akurasi ditampilkan pada Gambar 2, dan rangkuman metrik disajikan pada Tabel 4.



accuracy: 90.89% +/- 9.96% (micro average: 90.91%)

	true Layak	true Tidak Layak	class precision
pred. Layak	43	5	89.58%
pred. Tidak Layak	4	47	92.16%
class recall	91.49%	90.38%	

Gambar 2. Hasil Akurasi Model Decision Tree (10-fold Cross Validation)

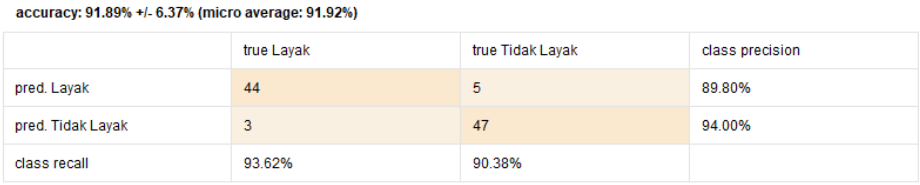
Tabel 4. Hasil Evaluasi Model Decision Tree

Metrik	Kelas Layak	Kelas Tidak Layak	Weighted
Presisi	89,58%	92,16%	91,99%
Recall	91,49%	90,38%	90,25%
F1-Score*	90,52%	91,26%	90,91%
Akurasi	-	-	90,89%

*F1-score dihitung secara manual menggunakan persamaan (6), karena metrik tersebut tidak tersedia langsung pada operator Performance (Classification) di Altair AI Studio 2026.0.5.

3.2.2 Random Forest

Model Random Forest dibangun dengan 100 pohon keputusan melalui mekanisme bootstrap sampling, pemilihan subset fitur secara acak pada tiap percabangan, serta agregasi hasil melalui majority voting. Berdasarkan hasil evaluasi, model menghasilkan confusion matrix: 44 True Positive, 47 True Negative, 5 False Positive, dan 3 False Negative. Hasil evaluasi akurasi ditampilkan pada Gambar 3, dan rangkuman metrik disajikan pada Tabel 5.



accuracy: 91.89% +/- 6.37% (micro average: 91.92%)

	true Layak	true Tidak Layak	class precision
pred. Layak	44	5	89.80%
pred. Tidak Layak	3	47	94.00%
class recall	93.62%	90.38%	

Gambar 3. Hasil Akurasi Model Random Forest (10-fold Cross Validation)

Tabel 5. Hasil Evaluasi Model Random Forest

Metrik	Kelas Layak	Kelas Tidak Layak	Weighted
Presisi	89,80%	94,00%	92,79%
Recall	93,62%	90,38%	91,92%
F1-Score*	91,67%	92,15%	91,93%
Akurasi	-	-	91,89%

*F1-score dihitung secara manual menggunakan persamaan (6), karena metrik tersebut tidak tersedia langsung pada operator Performance (Classification) di Altair AI Studio 2026.0.5.

3.3 Perbandingan Performa Algoritma

Perbandingan performa kedua model disajikan secara ringkas pada Tabel 6.

Tabel 6. Perbandingan Performa Decision Tree dan Random Forest

Metrik	Decision Tree	Random Forest
Akurasi	90,89%	91,89%
Presisi (Weighted)	91,99%	92,79%
Recall (Weighted)	90,25%	91,92%
F1-Score (Weighted)	90,91%	91,93%
Standar Deviasi Akurasi	±9,96%	±6,37%

Random Forest menghasilkan nilai yang lebih tinggi pada seluruh metrik evaluasi dibandingkan Decision Tree, dengan selisih akurasi sebesar 1,00 poin persentase (91,89% dibandingkan 90,89%). Selain itu, Random

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 162
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Information System Research (JOSH)
ISSN 2686-228X (media online)
Volume 8, No. 1, Oktober 2026, pp 157−166
DOI 10.47065/josh.v8i1.11037
https://ejurnal.seminar-id.com/index.php/josh/

Forest memiliki standar deviasi akurasi yang lebih rendah (±6,37%) dibandingkan Decision Tree (±9,96%) berdasarkan evaluasi 10-fold Cross Validation. Hasil ini mengindikasikan bahwa performa Random Forest relatif lebih konsisten antar-fold. Pada dataset yang terbatas, yaitu 99 sampel, konsistensi tersebut menjadi pertimbangan penting dalam mengevaluasi kinerja model, meskipun belum cukup untuk memastikan keandalannya pada data lain. Hasil confusion matrix juga menunjukkan bahwa Random Forest menghasilkan False Negative lebih sedikit dibandingkan Decision Tree, yaitu 3 berbanding 4, sedangkan jumlah False Positive kedua model sama, yaitu 5 siswa. Dalam konteks seleksi kandidat lomba robotik, False Negative perlu mendapat perhatian karena dapat menyebabkan siswa yang sebenarnya memenuhi kriteria kelayakan tidak terpilih sebagai kandidat. Sejalan dengan temuan tersebut, Random Forest memiliki recall kelas Layak yang lebih tinggi, yaitu 93,62%, dibandingkan Decision Tree sebesar 91,49%. Temuan ini menunjukkan bahwa Random Forest berpotensi menjadi alternatif untuk mendukung proses seleksi kandidat di PT. Sonnic Robotic. Namun, selisih kinerja yang relatif kecil dan penggunaan dataset dari satu lembaga membatasi generalisasi hasil, sehingga penerapan model tetap perlu disertai pertimbangan profesional pengajar dan pengujian pada data yang lebih luas.

3.4 Analisis Feature Importance

Analisis feature importance dilakukan pada kedua model untuk mengidentifikasi variabel yang paling berpengaruh dalam proses klasifikasi. Hasil analisis pada model Decision Tree (Tabel 7) menunjukkan bahwa hanya lima dari sembilan variabel yang digunakan sebagai node pemisah, dengan kreativitas menempati posisi dominan (weight 0,466), sementara persentase kehadiran, nilai bulan 1, bulan 3, dan bulan 4 sama sekali tidak digunakan (weight 0).

Tabel 7. Feature Importance Decision Tree

Variabel	Weight
Kreativitas	0,466
Nilai Bulan 2	0,168
Kemandirian	0,162
Pemahaman Materi	0,126
Nilai Bulan 5	0,077
% Kehadiran	0
Nilai Bulan 1, 3, 4	0

Berdasarkan Tabel 7, variabel Kreativitas memiliki tingkat kepentingan (feature importance) tertinggi dalam model Decision Tree dengan bobot sebesar 0,466. Hal ini menunjukkan bahwa kreativitas menjadi variabel yang paling dominan dalam proses klasifikasi kelayakan kandidat peserta lomba robotik. Selanjutnya, Nilai Bulan 2 memiliki bobot sebesar 0,168, diikuti oleh Kemandirian sebesar 0,162 dan Pemahaman Materi sebesar 0,126. Sementara itu, Nilai Bulan 5 memiliki kontribusi sebesar 0,077. Variabel % Kehadiran serta Nilai Bulan 1, 3, dan 4 memiliki nilai feature importance sebesar 0, yang menunjukkan bahwa variabel tersebut tidak memberikan kontribusi terhadap pembentukan keputusan pada model Decision Tree. Dengan demikian, model lebih banyak memanfaatkan aspek kreativitas, capaian nilai pada bulan tertentu, kemandirian, dan pemahaman materi dalam menentukan klasifikasi kelayakan kandidat. Berbeda dengan Decision Tree, model Random Forest (Tabel 8) mengaktifkan seluruh sembilan variabel input. Kreativitas tetap menjadi variabel dengan weight tertinggi (0,235), namun distribusi kepentingan antar variabel jauh lebih proporsional, dengan rentang weight antara 0,068 hingga 0,235.

Tabel 8. Feature Importance Random Forest

Variabel	Weight
Kreativitas	0,235
Pemahaman Materi	0,127
% Kehadiran	0,123
Kemandirian	0,118
Nilai Bulan 1	0,11
Nilai Bulan 4	0,074
Nilai Bulan 3	0,073
Nilai Bulan 2	0,073
Nilai Bulan 5	0,068

Berdasarkan Tabel 8, variabel Kreativitas memiliki tingkat kepentingan (feature importance) tertinggi pada model Random Forest dengan bobot sebesar 0,235. Hal ini menunjukkan bahwa kreativitas merupakan variabel yang paling dominan dalam menentukan klasifikasi kelayakan kandidat peserta lomba robotik. Selanjutnya, Pemahaman Materi memiliki bobot sebesar 0,127, diikuti oleh % Kehadiran sebesar 0,123, Kemandirian sebesar 0,118, dan Nilai Bulan 1 sebesar 0,110. Sementara itu, Nilai Bulan 4, Nilai Bulan 3, dan Nilai Bulan 2 masing-masing memiliki bobot sebesar 0,074, 0,073, dan 0,073, sedangkan Nilai Bulan 5 memiliki kontribusi terendah

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 163
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Information System Research (JOSH)
ISSN 2686-228X (media online)
Volume 8, No. 1, Oktober 2026, pp 157−166
DOI 10.47065/josh.v8i1.11037
https://ejurnal.seminar-id.com/index.php/josh/

dengan bobot sebesar 0,068. Berbeda dengan Decision Tree, seluruh sembilan variabel memiliki nilai feature importance pada Random Forest. Hal ini menunjukkan bahwa Random Forest mempertimbangkan seluruh variabel input dalam proses klasifikasi, meskipun dengan tingkat kontribusi yang berbeda. Secara keseluruhan, aspek kreativitas, pemahaman materi, kehadiran, dan kemandirian menjadi variabel yang relatif lebih dominan dibandingkan variabel nilai bulanan dalam pembentukan prediksi kelayakan kandidat.

3.5 Pembahasan

Hasil penelitian menunjukkan bahwa algoritma Random Forest memiliki kinerja yang relatif lebih baik dibandingkan Decision Tree dalam mengklasifikasikan kelayakan kandidat peserta lomba robotik di PT. Sonnic Robotic. Random Forest memperoleh akurasi sebesar 91,89%, weighted precision 92,79%, weighted recall 91,92%, dan weighted F1-score 91,93%. Hasil ini sejalan dengan penelitian Dana et al [9], Adrian et al. [10] pada tahun 2025, dan Setiono [11] di tahun 2022, yang menunjukkan bahwa Random Forest menghasilkan kinerja klasifikasi yang lebih baik dibandingkan algoritma pembanding pada masing-masing penelitian. Sementara itu, Decision Tree memperoleh akurasi sebesar 90,89%, weighted precision 91,99%, weighted recall 90,25%, dan weighted F1-score 90,91%. Selisih akurasi sebesar 1 poin persentase menunjukkan bahwa perbedaan kinerja kedua algoritma relatif kecil. Meskipun demikian, Random Forest menghasilkan nilai yang lebih tinggi pada seluruh metrik evaluasi. Namun, perbedaan domain penelitian, karakteristik dataset, dan konfigurasi pemodelan menyebabkan hasil akurasi antarpenelitian tidak dapat dibandingkan secara langsung.

Temuan ini memperkuat hasil penelitian terdahulu mengenai potensi Random Forest dalam menyelesaikan permasalahan klasifikasi sekaligus memperluas penerapannya pada konteks seleksi kandidat peserta lomba robotik. Dalam penelitian ini, Random Forest memanfaatkan seluruh sembilan variabel input, termasuk kehadiran dan nilai bulanan dari lima periode pengamatan, sedangkan Decision Tree menggunakan lima variabel dalam model yang terbentuk. Perbedaan tersebut menunjukkan bahwa Random Forest memanfaatkan lebih banyak variabel dalam proses klasifikasi pada dataset yang digunakan. Oleh karena itu, Random Forest dapat dipertimbangkan sebagai model pendukung keputusan untuk membantu PT. Sonnic Robotic mengembangkan proses seleksi yang lebih terstruktur dan berbasis data. Meskipun demikian, hasil klasifikasi tidak dimaksudkan untuk menggantikan penilaian profesional pengajar, melainkan sebagai informasi pendukung dalam menentukan kelayakan kandidat.

4. KESIMPULAN

Penelitian ini membandingkan kinerja algoritma Decision Tree dan Random Forest dalam mengklasifikasikan kelayakan kandidat peserta lomba robotik di PT. Sonnic Robotic berdasarkan sembilan variabel akademik dan non-akademik yang dinilai menggunakan rubrik terstruktur berskala 1–4. Penetapan bobot variabel dan ambang batas klasifikasi sebesar 75 didasarkan pada expert judgment untuk menyesuaikan proses klasifikasi dengan kebutuhan seleksi di lembaga tersebut. Hasil pengujian menggunakan 10-fold cross-validation menunjukkan bahwa Random Forest memperoleh akurasi sebesar 91,89%, presisi weighted 92,79%, recall weighted 91,92%, dan F1-score weighted 91,93%. Sementara itu, Decision Tree memperoleh akurasi 90,89%, presisi 91,99%, recall 90,25%, dan F1-score 90,91%. Random Forest juga menunjukkan variasi akurasi antarfold yang lebih rendah, dengan standar deviasi ±6,37%, dibandingkan Decision Tree sebesar ±9,96%, serta memanfaatkan seluruh sembilan variabel input dalam proses pemodelannya. Hasil tersebut menunjukkan bahwa Random Forest memiliki kinerja yang relatif lebih baik dan lebih stabil dibandingkan Decision Tree pada dataset yang digunakan. Namun, keunggulan ini belum dapat digeneralisasikan pada populasi siswa atau lembaga pendidikan lain karena penelitian hanya menggunakan data dari satu lembaga dengan cakupan yang terbatas. Secara praktis, hasil penelitian ini dapat menjadi bahan pertimbangan bagi PT. Sonnic Robotic dalam mengembangkan proses seleksi kandidat yang lebih terstruktur dan berbasis data. Meskipun demikian, hasil klasifikasi tidak dimaksudkan untuk menggantikan penilaian pengajar, melainkan sebagai informasi pendukung dalam pengambilan keputusan. Secara teoretis, penelitian ini memberikan bukti empiris terbatas mengenai perbandingan kinerja Decision Tree dan Random Forest pada konteks seleksi peserta lomba robotik. Penelitian ini memiliki keterbatasan berupa penggunaan dataset dari satu lembaga, cakupan data yang terbatas pada siswa sekolah dasar, serta belum diterapkannya hyperparameter tuning pada Random Forest. Oleh karena itu, penelitian selanjutnya disarankan untuk menggunakan dataset yang lebih besar dan berasal dari beberapa lembaga, memperluas cakupan jenjang pendidikan, menerapkan optimasi hiperparameter, serta membandingkan kinerja algoritma lain, seperti Support Vector Machine dan Gradient Boosting. PT. Sonnic Robotic juga disarankan untuk menerapkan rubrik penilaian yang konsisten, mengevaluasi hasil klasifikasi secara berkala, dan memperbarui model seiring bertambahnya data agar kinerja serta relevansinya terhadap kebutuhan seleksi dapat terus ditinjau. Sebagai kontribusi akhir, penelitian ini memberikan landasan empiris awal bagi pengembangan sistem pendukung keputusan berbasis klasifikasi untuk seleksi kandidat peserta lomba robotik. Hasil komparasi Decision Tree dan Random Forest dapat dimanfaatkan sebagai referensi dalam menentukan pendekatan klasifikasi yang sesuai dengan karakteristik data dan kebutuhan seleksi di PT. Sonnic Robotic dengan tetap menempatkan penilaian profesional pengajar sebagai bagian penting dalam pengambilan keputusan.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 164
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Information System Research (JOSH)
ISSN 2686-228X (media online)
Volume 8, No. 1, Oktober 2026, pp 157−166
DOI 10.47065/josh.v8i1.11037
https://ejurnal.seminar-id.com/index.php/josh/

5. PERNYATAAN PENULIS

5.1 Kontribusi Penulis

Konseptualisasi: Aura Nurshifa, Entin Sutinah, dan Nani Agustina;
Metodologi: Aura Nurshifa dan Entin Sutinah;
Perangkat Lunak: Aura Nurshifa;
Validasi: Aura Nurshifa, Entin Sutinah, dan Nani Agustina;
Analisis Formal: Aura Nurshifa dan Entin Sutinah;
Investigasi: Aura Nurshifa dan Entin Sutinah;
Sumber Daya: Aura Nurshifa, Entin Sutinah, dan Nani Agustina;
Kurasi Data: Entin Sutinah dan Nani Agustina;
Penulisan Draf Awal: Aura Nurshifa dan Entin Sutinah;
Penulisan, Peninjauan, dan Penyuntingan: Aura Nurshifa, Entin Sutinah, dan Nani Agustina;
Visualisasi: Aura Nurshifa dan Entin Sutinah;
Seluruh penulis telah membaca dan menyetujui versi manuskrip yang telah diterbitkan.

5.2 Ketersediaan Data

Data yang digunakan dalam penelitian ini tersedia berdasarkan permintaan yang wajar kepada penulis korespondensi, dengan tetap memperhatikan kerahasiaan data peserta dan kebijakan PT. Sonnic Robotic.

5.3 Pendanaan

Penelitian ini tidak menerima pendanaan khusus dari lembaga pendanaan publik, komersial, maupun organisasi nirlaba. Seluruh biaya yang terkait dengan pelaksanaan penelitian dan penulisan manuskrip ini ditanggung secara mandiri oleh para penulis.

5.4 Pernyataan Dewan Peninjau Institusional

Penelitian ini telah memperoleh persetujuan dari institusi terkait sebelum pelaksanaan penelitian dan pengumpulan data. Pelaksanaan penelitian dilakukan sesuai dengan ketentuan dan prosedur yang berlaku di PT. Sonnic Robotic.

5.5 Pernyataan Persetujuan Berdasarkan Informasi

Persetujuan berdasarkan informasi telah diperoleh dari narasumber sebelum pelaksanaan wawancara. Narasumber telah diberi penjelasan mengenai tujuan penelitian dan penggunaan informasi yang diberikan untuk kepentingan penelitian dan publikasi ilmiah.

5.6 Pernyataan Konflik Kepentingan

Para penulis menyatakan bahwa mereka tidak memiliki kepentingan keuangan yang bersaing atau hubungan pribadi yang diketahui yang dapat dianggap memengaruhi pekerjaan yang dilaporkan dalam artikel ini.

5.7 Penggunaan AI Generatif dan Teknologi Berbantuan AI dalam Proses Penulisan

Para penulis menggunakan AI Generatif untuk meningkatkan kejelasan penulisan dalam artikel ini. Para penulis telah meninjau dan menyunting konten yang dihasilkan dengan bantuan AI serta bertanggung jawab penuh atas versi akhir publikasi.

REFERENCES

[1] M. N. Gul, W. Abbasi, and M. Y. Wani, “Revolutionizing educational decision - making : a robust machine learning mechanism for predicting student performance,” J. Electr. Syst. Inf Technol, vol. 12, no. 1, pp. 2–23, 2025, doi: 10.1186/s43067-025-00230-z.
[2] A. R. Azis, “Analisis Komparasi Algoritma Machine Learning dalam Prediksi Performa Akademik Mahasiswa : Literature Review,” J. Ilmu Komput. dan Inform., vol. 4, no. 2, pp. 143–150, 2024, doi: https://doi.org/10.54082/jiki.212.
[3] K. W. Haryanto, M. F. Roji, and M. R. Lazuardi, “Sistem Pendukung Pengambilan Keputusan Pemilihan Peserta Lomba Kompetensi Siswa Menggunakan Metode SMART,” J. SPIRIT, vol. 15, no. 2, pp. 79–93, 2023, doi: 10.53567/spirit.v15i2.318.
[4] B. O. S. Gultom, H. I. Sinaga, R. Fauji, T. Siahaan, and L. M. Harahap, “Strategi Rekrutmen Transparan Sebagai Solusi Mengurangi Bias Dan Nepotisme Dalam Manajemen Sumber Daya Manusia: Kajian Literatur Sistematis,” J. Organ. DAN Manaj. Indones., vol. 3, no. 3, pp. 239–249, 2025, [Online]. Available: https://ejournal.mediapenamandiri.com/index.php/jomi/article/view/123/71.
[5] N. A. Harahap et al., “Perbandingan Kinerja Algoritma Decision Tree dan Random Forest dalam Prediksi Kelulusan Siswa,” J. Sist. Inf. TGD, vol. 5, no. 1, pp. 124–129, 2026, doi: 10.53513/jursi.v5i1.12469.
[6] A. H. Nasrullah, “Implementasi Algoritma Decision Tree Untuk Klasifikasi Produk Laris,” J. Ilm. Ilmu Komput., vol. 7, no. 2, pp. 45–51, 2021, doi: https://doi.org/10.35329/jiik.v7i2.203.
[7] D. Hartama and N. Amalya, “Perbandingan Algoritma Decision Tree , ID3 , dan Random Forest dalam Klasifikasi Faktor-Faktor yang Mempengaruhi Karier Mahasiswa Ilmu Komputer,” J. Indones. Manaj. Inform. dan Komun., vol. 6, no. 1, pp. 72–80, 2025, doi: https://doi.org/10.35870/jimik.v6i1.1113.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 165
This work is licensed under a Creative Commons Attribution 4.0 International License.

Journal of Information System Research (JOSH)
ISSN 2686-228X (media online)
Volume 8, No. 1, Oktober 2026, pp 157−166
DOI 10.47065/josh.v8i1.11037
https://ejurnal.seminar-id.com/index.php/josh/

[8] E. R. B. Sebayang, Y. H. Chrisnanto, and M. V, "Klasifikasi Data Kesehatan Mental di Industri Teknologi Menggunakan Algoritma Random Forest," IJESPG J., vol. 1, no. 3, pp. 237–253, 2023, [Online]. Available: http://ijespgjournal.org.
[9] A. R. Dana, R. V. Kristananda, M. Bagas, S. Wibowo, and D. A. Prasetya, "Perbandingan Algoritma Decision Tree dan Random Forest dengan Hyperparameter Tuning dalam Mendeteksi Penyakit Stroke," Semin. Nas. Inform. Bela Negara, vol. 4, pp. 66–75, 2024, [Online]. Available: https://santika.upnjatim.ac.id/submissions/index.php/santika/article/view/349.
[10] A. Adrian and I. Verawati, "Analisis Performa Logistic Regression dan Random Forest dalam Klasifikasi Kelayakan Penerimaan Kredit," vol. 4, no. 2, pp. 148–158, 2025, doi: https://doi.org/10.59095/ijcsr.v4i2.205.
[11] M. H. Setiono, "Komparasi Algoritma Decision Tree, Random Forest, Svm dan K-Nn Dalam Klasifikasi Kepuasan Penumpang Maskapai Penerbangan," Inti Nusa Mandiri, vol. 17, no. 1, pp. 32–39, 2022, doi: https://doi.org/10.33480/inti.v17i1.3420.
[12] D. Rahmallia, "Identifikasi Minat dan Bakat Siswa di SMA melalui Sistem Pendukung Keputusan Berbasis Machine Learning," J. Pendiidkan Indones. Teor. Penelit. dan Inov. ISSN, vol. 4, no. 3, pp. 32–29, 2024, doi: 10.59818/jpi.v4i3.804.
[13] T. Susilawati and A. B. Trisnawan, "Pemanfaatan Machine Learning untuk Peningkatan Akurasi Sistem Pendukung Keputusan Prediktif," J. Unitek, vol. 18, no. 2, pp. 2580–2582, 2025, doi: https://doi.org/10.52072/unitek.v18i2.1702.
[14] E. Desvazulinda, M. Iqba, and M. S. Novelan, "Analysis of the Decision Tree (C4 . 5) and Random Forest Algorithms to Determine Student Eligibility for Final Project Assignments Based on Academic Requirements," J. Sist. Komput. dan Inform., vol. 7, no. 4, pp. 1667–1675, 2026, doi: 10.30865/json.v7i4.9947.
[15] R. Raya et al., Menguasai Teknik Sampling: Dasar-Dasar untuk Analisis Data. PT Penerbit Qriset Indonesia, 2025.
[16] A. Ramdani, C. D. Sofyan, F. Ramdani, M. F. A. T. Tama, M. Rachmatsyah, and Angga, "Algoritma Klasifikasi Data Mining Untuk Memprediksi Masyarakat Dalam Menerima Bantuan Sosial," J. Ilm. Sist. Inf., vol. 1, no. 2, pp. 39–47, 2022, doi: https://doi.org/10.51903/juisi.v1i2.363.
[17] M. A. Firman, S. A. F. Djamalilleil, W. Zega, L. Efrizoni, and Rahmaddeni, "Model Klasifikasi IPK Mahasiswa Menggunakan Algoritma Decision Tree dan Random Forest Berbasis Feature Engineering," Techno.Com J. Teknol. Inf., vol. 24, no. 2, pp. 391–404, 2025, doi: 10.62411/tc.v24i2.12384.
[18] L. Savitri and R. Nursalim, "Klasifikasi Kualitas Air Minum menggunakan Penerapan Algoritma Machine Learning dengan Pendekatan Supervised Learning," Diophantine J. Math. Its Appl., vol. 02, no. 01, 2023, doi: 10.33369/diophantine.v2i01.28260.
[19] E. Jannah, V. Sihombing, and M. Masriza, "Penerapan Data Mining Klasifikasi Kepuasan Pelanggan Transportasi Online Menggunakan Algoritma C4 . 5," MEANS (Media Inf. Anal. dan Sist., vol. 8, no. 1, pp. 1–7, 2023, doi: https://doi.org/10.54367/means.v8i1.
[20] G. Yoga and A. Priatdana, "Decision Tree Model for Classifying University Students Eligible for UKT Waivers," JSIKTI J. Sist. Inf. dan Komput. Terap. Indones., vol. 7, no. September, pp. 285–294, 2024, doi: http://doi.org/10.26594/register.v8i2.2541.
[21] K. G. Santiyuda, "Random Forest for Precise Predictions of Customer Experience at Restaurant X," JSIKTI J. Sist. Inf. dan Komput. Terap. Indones., vol. 6, no. June, pp. 235–244, 2024, doi: 10.33173/jsikti.240.
[22] A. Riansah, O. Nurdiawan, and R. Herdiana, "Penerapan Algoritma Random Forest Dan Decision Tree Untuk Meningkatkan Akurasi Klasifikasi Penjualan Pada Toko Bangunan," JATI (Jurnal Mhs. Tek. Inform., vol. 9, no. 3, pp. 4242–4249, 2025, doi: 10.36040/jati.v9i3.13622.
[23] F. A. S. Wibowo et al., "Impact Of Feature Selection On Decision Tree And Random Forest For Classifying Student Study Success," BAREKENG J. Math. Its Appl., vol. 19, no. 3, pp. 2083–2096, 2025, doi: https://doi.org/10.30598/barekengvol19iss3pp2083-209.
[24] N. G. Ramadhan and A. Khoirunnisa, "Klasifikasi Data Malaria Menggunakan Metode Support Vector Machine," J. Media Inform. Budidarma, vol. 5, no. 4, pp. 1580–1584, 2021, doi: 10.30865/mib.v5i4.3347.
[25] Y. Nurohman and R. Sanjaya, "Sistem Pendukung Keputusan Siswa Berprestasi Menggunakan Metode Fuzzy-Ahp Di SMK PGRI Cibaribis," E-Prosiding Tek. Inform., vol. 6, no. 2, pp. 234–246, 2025, [Online]. Available: https://eprosiding.ars.ac.id/index.php/pti/article/view/68/74.

Copyright © 2026 The Author(s). Published by *Forum Kerjasama Pendidikan Tinggi*. Page 166
This work is licensed under a Creative Commons Attribution 4.0 International License.