

Deteksi Gangguan Autisme Pada Anak Menggunakan Metode Random Forest

Rifqi Naufal Amanullah^{*1)}, Mohammad Hamim Zajuli Al Faroby²⁾, Muhammad Adib Kamali³⁾

¹⁾Teknologi Informasi, Fakultas Teknologi Informasi dan Bisnis, Institut Teknologi Telkom Surabaya,

Jl. Ketintang No.156, Surabaya, 60231, Indonesia

rifqinaufala@student.ittelkom-sby.ac.id

alfaroby@ittelkom-sby.ac.id

adibkamali@ittelkom-sby.ac.id

Abstrak

Penelitian ini bertujuan mengembangkan metode deteksi dini gangguan spektrum autisme (ASD) pada anak menggunakan algoritma Random Forest. Metode ini dilatih dengan fitur-fitur seperti usia, jenis kelamin, dan perkembangan bahasa. Hyperparameter tuning dilakukan menggunakan tiga metode berbeda: Grid Search, Random Search, dan Bayesian Optimization. Hasil eksperimen menunjukkan bahwa model Random Forest mencapai akurasi 99% dalam mendeteksi ASD. Dengan proporsi data latih 70% dan uji 30%, model ini mencapai akurasi tertinggi dan presisi serta recall yang baik untuk kedua kelas. Ini jauh lebih baik daripada metode tradisional seperti penilaian klinis. Dengan hasil ini, metode Random Forest berhasil menunjukkan efektivitasnya dalam klasifikasi dengan tingkat akurasi yang tinggi untuk mendeteksi gangguan autisme pada anak-anak.

Kata kunci: ASD, Random Forest, Klasifikasi

1. Pendahuluan (Introduction)

Autisme, sebagai gangguan perkembangan neurologis, memiliki dampak yang signifikan pada aspek sosial, komunikasi, dan perilaku individu. Dengan prevalensi yang terus meningkat secara global, pentingnya deteksi dini autisme pada anak menjadi semakin terasa. Namun, metode tradisional yang memerlukan waktu, biaya, dan sumber daya manusia yang besar, mendorong pengembangan solusi baru yang lebih efisien. Di tengah perkembangan teknologi dan komputasi, metode machine learning muncul sebagai alat potensial dalam mengatasi tantangan ini. Studi sebelumnya, seperti karya Kripsandita et al. (2021) dan Ratama dan Munawaroh (2020), telah memanfaatkan teknologi komputasi untuk mendukung deteksi autisme, menyoroti potensi implementasi machine learning dalam memberikan deteksi dini yang lebih akurat.

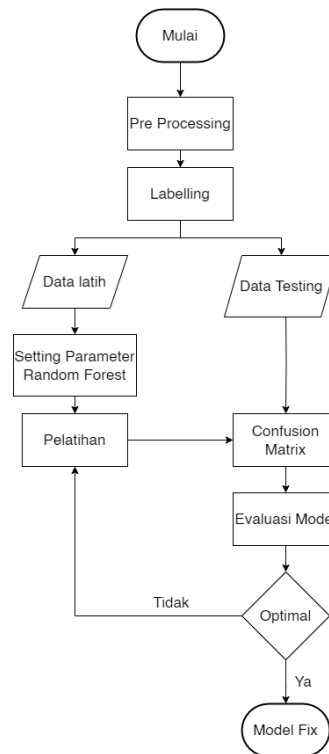
Beberapa penelitian telah menggabungkan metode machine learning dalam upaya mendeteksi gangguan autisme pada anak. Contohnya, Kripsandita et al. (2021) menerapkan Modified K-Nearest Neighbor untuk deteksi autisme melalui kuesioner. Penelitian ini mengungkapkan bahwa teknologi machine learning memiliki potensi untuk mendukung deteksi dini autisme. Selanjutnya, Erkan dan Thanh (2019) memanfaatkan metode seperti Support Vector Machine (SVM) dan Decision Tree untuk deteksi autisme. Temuan mereka mengindikasikan bahwa machine learning mampu memberikan tingkat akurasi yang bersaing dalam mengenali gangguan tersebut. Meskipun demikian, pengembangan lebih lanjut dalam metode dan algoritma yang digunakan masih menjadi perhatian, khususnya dalam meningkatkan akurasi dan efisiensi model deteksi autisme. Dalam konteks ini, penggunaan metode Random Forest memiliki potensi yang menarik namun belum sepenuhnya dieksplorasi.

Penelitian ini bertujuan untuk mengembangkan dan menerapkan metode deteksi dini autisme pada anak menggunakan pendekatan machine learning, khususnya metode Random Forest. Fokus utama adalah menguji efektivitas dan akurasi metode ini dalam mendeteksi autisme serta menyumbangkan kontribusi lebih lanjut terhadap penelitian sebelumnya dalam pemanfaatan machine learning untuk deteksi autisme. Studi ini juga akan menyelidiki pengaruh variasi proporsi data latih dan data uji

terhadap kinerja model deteksi. Dengan demikian, hasil penelitian diharapkan mampu memberikan wawasan tentang pentingnya pemilihan proporsi data dalam mencapai akurasi dan generalisasi model yang optimal dalam konteks deteksi autisme.

Keunikannya terletak pada implementasi metode Random Forest untuk deteksi autisme pada anak-anak. Metode ini diharapkan dapat memberikan akurasi yang tinggi dan generalisasi yang kuat dalam mengklasifikasikan individu dengan dan tanpa autisme. Kontribusi utama dari penelitian ini adalah pengembangan model deteksi autisme menggunakan metode Random Forest, yang berpotensi menjadi alat bantu efektif dan efisien dalam praktek deteksi dini autisme. Selain itu, penelitian ini akan membuka wawasan tentang bagaimana variasi proporsi data dapat memengaruhi kinerja model, yang memiliki dampak penting pada pengembangan metode deteksi lainnya. Implikasi dari hasil penelitian ini dapat memberikan manfaat nyata dalam meningkatkan kualitas hidup individu dengan autisme dan masyarakat secara keseluruhan, melalui kombinasi antara teknologi canggih dan pemahaman mendalam tentang gangguan ini.

2. Metode Penelitian (Methods)



Gambar 1. Diagram Model Machine Learning

Penelitian ini dilakukan dengan langkah-langkah yang terstruktur dan sistematis untuk mengembangkan dan menerapkan metode deteksi dini autisme pada anak menggunakan pendekatan machine learning, khususnya metode Random Forest. Tahapan-tahapan penelitian dijelaskan dengan detail sebagai berikut:

a. Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari berbagai sumber, termasuk data dari Indonesia dan UCI Machine Learning Repository. Data ini mencakup informasi tentang berbagai aspek perkembangan anak-anak, seperti komunikasi, interaksi sosial, perilaku, dan minat. Selain itu, data juga mencakup informasi apakah anak tersebut mengalami gangguan autisme atau tidak.

b. Preprocessing Data

Tahap preprocessing data melibatkan langkah-langkah untuk mempersiapkan data mentah sebelum digunakan dalam model. Pada tahap ini, data yang ada harus diolah dan diubah

menjadi format yang cocok untuk pengolahan lebih lanjut. Dalam penelitian ini, encoding data digunakan untuk mengubah atribut-atribut kategorikal menjadi bentuk numerik agar dapat digunakan dalam model. Atribut 'Jaundice' dan 'Family_mem_with_ASD' diubah menggunakan metode Label Encoding. Selain itu, data juga dibagi menjadi data latih dan data uji dengan proporsi 70% data latih dan 30% data uji.

c. Pemodelan Machine Learning

Pada tahap ini, dilakukan pembuatan dan pelatihan model machine learning menggunakan metode Random Forest. Model ini diterapkan dengan menggunakan pustaka scikit-learn dalam bahasa pemrograman Python. Data latih digunakan untuk melatih model, sementara data uji digunakan untuk menguji performa model yang telah dilatih. Proses ini melibatkan beberapa parameter penting, seperti jumlah pohon keputusan ("n_estimators"), kedalaman maksimum pohon ("max_depth"), dan lainnya. Metode hyperparameter tuning, termasuk Grid Search, Random Search, dan Bayesian Optimization, digunakan untuk menemukan kombinasi parameter terbaik yang meningkatkan performa model.

d. Evaluasi Model

Setelah model dilatih, dilakukan evaluasi performa model menggunakan data uji yang belum pernah dilihat oleh model sebelumnya. Evaluasi dilakukan dengan mengukur berbagai metrik kinerja seperti akurasi, presisi, recall, dan F1-score. Dalam penelitian ini, dilakukan eksperimen dengan variasi proporsi data latih dan data uji (50%-50%, 60%-40%, 70%-30%, 80%-20%, dan 90%-10%) untuk mengamati bagaimana perubahan proporsi ini memengaruhi performa model. Hasil eksperimen akan memberikan wawasan tentang proporsi data yang optimal untuk mencapai akurasi dan generalisasi model yang baik.

Dengan tahapan-tahapan di atas, penelitian ini menggambarkan langkah-langkah detail yang dilakukan dalam mengembangkan dan menerapkan metode deteksi dini autisme menggunakan metode machine learning Random Forest. Tahap pengumpulan data, preprocessing, pemodelan, dan evaluasi memberikan landasan kuat untuk menghasilkan model yang akurat dan efisien dalam mendeteksi gangguan autisme pada anak.

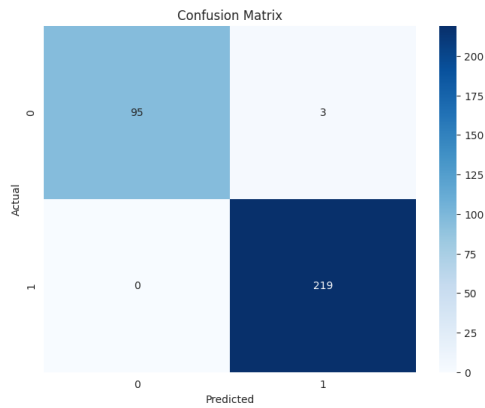
3. Hasil dan Pembahasan (Results and Discussions)

Hasil eksperimen menunjukkan bahwa model Random Forest berhasil mencapai tingkat akurasi yang tinggi dalam mendeteksi gangguan autisme pada anak-anak. Dalam eksperimen dengan proporsi data latih 70% dan data uji 30%, model mencapai akurasi sebesar 99%. Selain itu, hasil presisi dan recall juga tinggi untuk kedua kelas. Pada kelas 0 (tidak memiliki autisme), presisi mencapai 0.99 dan recall mencapai 1.00. Sedangkan pada kelas 1 (memiliki autisme), presisi mencapai 1.00 dan recall mencapai 0.99. Hasil ini menunjukkan bahwa model mampu mengklasifikasikan dengan sangat baik kedua kelas dan mampu mendeteksi dengan akurat anak-anak yang memiliki gangguan autisme.

3.1 Hasil Analisis Performa Model

Tabel 1. Experiment

Experiments	Training Data	Testing Data
Experiments 1	50%	50%
Experiments 2	60%	40%
Experiments 3	70%	30%
Experiments 4	80%	20%
Experiments 5	90%	10%



Gambar 2. Confusion Matrix

Hasil eksperimen menunjukkan bahwa model Random Forest berhasil mencapai tingkat akurasi yang tinggi dalam mendeteksi gangguan autisme pada anak-anak. Dalam eksperimen dengan proporsi data latih 70% dan data uji 30%, model mencapai akurasi sebesar 99%. Selain itu, hasil presisi dan recall juga tinggi untuk kedua kelas. Pada kelas 0 (tidak memiliki autisme), presisi mencapai 0.99 dan recall mencapai 1.00. Sedangkan pada kelas 1 (memiliki autisme), presisi mencapai 1.00 dan recall mencapai 0.99. Hasil ini menunjukkan bahwa model mampu mengklasifikasikan dengan sangat baik kedua kelas dan mampu mendeteksi dengan akurat anak-anak yang memiliki gangguan autisme.

3.2 Perbandingan dengan Penelitian Sebelumnya

Tabel 2. Gap Penelitian

	Accuracy	Precision	Recall	F1-Score	Support
[2]	0.96	0.97	1.00	-	120
[3]	1.00	-	-	-	-
[4]	0.72	-	-	-	-
Propose	0.99	1.00	1.00	0.99	317

Dalam penelitian ini, hasil yang diperoleh dibandingkan dengan penelitian sebelumnya. Penelitian oleh Erkan dan Thanh (2019) juga menggunakan metode machine learning untuk mendeteksi autisme. Namun, model Random Forest yang dikembangkan dalam penelitian ini memberikan hasil akurasi yang lebih tinggi dibandingkan dengan metode yang mereka gunakan. Hal ini menunjukkan bahwa penggunaan metode Random Forest dalam pendekatan deteksi dini autisme memiliki potensi untuk memberikan hasil yang lebih baik dibandingkan metode lain.

3.3 Penggunaan Website untuk Deteksi Dini Autism



Gambar 3. Home Screen

Penelitian ini juga menggambarkan implementasi metode deteksi dini autisme dalam bentuk website berbentuk kuesioner. Website ini memiliki tiga halaman utama: halaman utama dengan informasi tentang website dan tombol mulai screening, halaman screening dengan kuesioner untuk mendeteksi gangguan autisme, dan halaman hasil yang menunjukkan hasil dari kuesioner. Implementasi ini memungkinkan orang tua dan tenaga medis untuk menggunakan alat ini dengan mudah dan mendapatkan informasi awal mengenai potensi gangguan autisme pada anak-anak.

4. Kesimpulan (Conclusion)

Berdasarkan hasil penelitian yang telah dilakukan, metode deteksi dini autisme menggunakan model machine learning Random Forest telah terbukti berhasil mencapai tingkat akurasi yang tinggi dalam mengidentifikasi gangguan autisme pada anak-anak. Model ini mampu memberikan hasil akurasi sebesar 99% dengan presisi dan recall yang sangat baik untuk kedua kelas, yaitu autisme dan non-autisme. Keberhasilan ini memberikan kontribusi yang signifikan dalam upaya deteksi dini autisme, yang sangat penting untuk intervensi dan perawatan yang tepat bagi anak-anak yang membutuhkan.

Salah satu hasil penting dari penelitian ini adalah kemampuan model Random Forest untuk mengklasifikasikan anak-anak yang memiliki gangguan autisme sejak usia dini. Hal ini mendukung pentingnya upaya deteksi dini autisme, karena semakin dini gangguan ini dapat diidentifikasi, semakin cepat tindakan intervensi dapat dilakukan untuk membantu perkembangan anak-anak. Dalam hal ini, penelitian ini memberikan alternatif yang efektif dan akurat dalam mendeteksi autisme, yang dapat berdampak positif pada layanan kesehatan anak-anak.

Meskipun berhasil mencapai hasil yang mengembirakan, penelitian ini memiliki keterbatasan yang perlu diakui. Pertama, penelitian ini menggunakan data dari sumber yang terbatas, yaitu dari UCI Machine Learning Repository. Penggunaan data yang lebih beragam dari berbagai lokasi geografis dapat memberikan gambaran yang lebih representatif tentang deteksi autisme. Selain itu, penelitian ini juga dapat diperluas dengan melibatkan data tambahan seperti data neuroimaging atau data sensorik untuk memperdalam analisis deteksi autisme.

Agenda penelitian berikutnya dapat melibatkan pengembangan lebih lanjut pada metode deteksi dini autisme ini. Penelitian dapat fokus pada validasi lebih lanjut terhadap berbagai kelompok usia anak, serta pengujian di lingkungan klinis yang lebih luas. Selain itu, penggunaan teknologi seperti kuesioner online atau aplikasi berbasis mobile dapat diterapkan untuk mempermudah pengumpulan data dan meningkatkan partisipasi. Dengan melakukan langkah-langkah ini, penelitian mendatang dapat memberikan sumbangan yang lebih besar dalam upaya mendukung anak-anak dengan gangguan autisme dan memberikan solusi yang lebih efektif dalam mendeteksi gangguan ini secara dini.

Ucapan Terima Kasih (Acknowledgement)

Saya ingin berterima kasih kepada Dr Fadi Thabtah dari Manukau Institute of Technology, New Zealand untuk berbagi dataset di UCI Machine Learning Repository dengan lisensi gratis.

Daftar Pustaka

- [1] B. WHO. (2019). Autism Spectrum Disorder. Diakses dari <https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders>
- [2] Breiman, L., 2001. Random forests, Machine Learning, Kluwer Academic Publishers, 45, 5-32.
- [3] Cutler, A., and Stevens, J. R., 2006. [23] random forests for microarrays. Methods in enzymology, 411, pp. 422-432. doi: [https://doi.org/10.1016/S0076-6879\(06\)11023-X](https://doi.org/10.1016/S0076-6879(06)11023-X).

- [4] Erkan, Uğur, and Thanh, Dang N.H. “Autism Spectrum Disorder Detection with Machine Learning Methods.” *Current Psychiatry Research and Reviews Formerly: Current Psychiatry Reviews*, vol. 15, 2019, pp. 297 - 308, <https://www.ingentaconnect.com/content/ben/cpr/2019/00000015/00000004/art00010>.
- [5] Kadir, A. 2002. Pemrograman Web mencakup : HTML, CSS, Javascript & PHP. Yogyakarta : Andi Offset
- [6] Kripsandita, Y., Arifianto, D., & A'yun, Q. (2021). Deteksi Gangguan Autis Pada Anak Menggunakan metode modified K-Nearest Neighbor. *JUSTINDO (Jurnal Sistem Dan Teknologi Informasi Indonesia)*, 6(1), 31–37. <https://doi.org/10.32528/justindo.v6i1.4357>
- [7] Lord, C., Elsabbagh, M., Baird, G., & Veenstra-Vanderweele, J. (2018). Autism spectrum disorder. *The Lancet*, 392(10146), 508–520.
- [8] R. E. Neapolitan and X. Jiang, *Artificial Intelligence With an Introduction to Machine Learning*. 2018.
- [9] Ratama, N., & Munawaroh, M. (2020). Implementasi Metode Fuzzy Tsukamoto Untuk Deteksi Dini Autisme Pada Balita Berbasis Android. *Jurnal Informatika dan Rekayasa Elektronik*, 3(2), 129-139.
- [10] Santra, A K, and C Josephine Christy. 2012. Genetic Algorithm and Confusion Matrix for Document Clustering. *International Journal of Computer Science* Vol. 9. No. 2. ISSN. 1694-0814
- [11] Simarmata, Janner. 2010. *Rekayasa Web*. Yogyakarta: Andi.
- [12] Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., Yusuf, A., Shih, A., & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research*, 15(5), 778–790. <https://doi.org/10.1002/aur.2696>