

Analisis Prediksi Gempa Bumi Menggunakan Algoritma Decision Tree C4.5: Studi Literatur

Ahmad Heriyansyah¹, Dr. Hetty Rohayani. AH, ST, M.Kom²

Program Studi Informatika, Fakultas Sains dan Teknologi, Universitas Muhammadiyah Jambi^{1,2}

Email: ahmadheriyansyah22@gmail.com¹, hettyrohayani@email.com²

ABSTRAK

Gempa bumi merupakan bencana alam yang memiliki dampak besar terhadap kehidupan manusia dan infrastruktur. Penelitian ini bertujuan untuk menganalisis dan mengevaluasi efektivitas algoritma pohon keputusan C4.5 dalam memprediksi kejadian gempa bumi berdasarkan data seismik historis. Dengan pendekatan kuantitatif dan kajian literatur sistematis, penelitian ini mengumpulkan informasi dari berbagai studi terdahulu terkait model prediksi gempa bumi. Hasil studi menunjukkan bahwa model C4.5 mampu memberikan akurasi prediksi yang cukup tinggi pada wilayah rawan gempa tertentu, sehingga berpotensi digunakan dalam sistem peringatan dini.

Kata Kunci: C4.5, Prediksi Gempa Bumi, Data Seismik, Pohon Keputusan, Sistem Peringatan Dini

ABSTRACT

Earthquakes are natural disasters that can severely affect human lives and infrastructure. This study aims to analyze and evaluate the effectiveness of the C4.5 decision tree algorithm in predicting earthquakes based on historical seismic data. Using a quantitative approach and systematic literature review, the study compiles information from previous research related to earthquake prediction models. The results suggest that the C4.5 model provides reasonably high prediction accuracy in earthquake-prone areas and could be integrated into early warning systems.

Keywords: C4.5, Earthquake Prediction, Seismic Data, Decision Tree, Early Warning System

PENDAHULUAN

Gempa bumi merupakan salah satu bencana alam yang paling mematikan dan tidak terduga. Wilayah Indonesia secara geologis terletak di pertemuan tiga lempeng tektonik utama dunia, yaitu Lempeng Eurasia, Lempeng Indo-Australia, dan Lempeng Pasifik [1]. Konvergensi ini menjadikan Indonesia sangat rawan terhadap aktivitas seismik [2]. Dalam beberapa dekade terakhir, berbagai gempa besar telah menyebabkan ribuan korban jiwa serta kerusakan infrastruktur yang sangat besar, seperti yang terjadi di Aceh (2004), Yogyakarta (2006), Palu (2018), dan Cianjur (2022) [3].

Prediksi gempa bumi telah menjadi tantangan besar dalam bidang geofisika dan ilmu data karena sifatnya yang kompleks dan nonlinear [4]. Beberapa pendekatan tradisional seperti model probabilistik, analisis statistik spasio-temporal, dan metode geofisika murni, seringkali belum mampu memberikan hasil prediksi yang akurat atau tepat waktu [5]. Oleh karena itu, muncul kebutuhan untuk mengembangkan pendekatan yang lebih adaptif dan berbasis data besar (big data), salah satunya adalah dengan memanfaatkan algoritma kecerdasan buatan seperti Decision Tree C4.5 [6].

Algoritma C4.5 merupakan pengembangan dari ID3 yang dirancang oleh Quinlan [1]. Algoritma ini banyak digunakan dalam klasifikasi data karena kemampuannya menangani atribut numerik dan kategorikal secara bersamaan [7][26]. C4.5 menghasilkan model dalam bentuk

struktur pohon yang terdiri dari keputusan-keputusan logis yang mudah dimengerti oleh manusia [6]. Ini menjadikannya sangat cocok untuk digunakan dalam sistem prediksi dan peringatan dini bencana, karena hasilnya bisa langsung diterjemahkan ke dalam kebijakan mitigasi atau notifikasi risiko [8].

Berbagai studi telah mencoba menerapkan C4.5 dalam konteks prediksi gempa, dengan menggunakan dataset dari USGS [4], BMKG [5], dan sumber seismik lain. Variabel yang umum digunakan meliputi magnitudo, kedalaman, lokasi episentrum, waktu kejadian, serta frekuensi aktivitas seismik sebelumnya [9]. Hasil dari model ini kemudian dikategorikan menjadi klasifikasi seperti “potensi gempa ringan”, “sedang”, atau “berisiko tinggi” [10].

Melalui penelitian ini, penulis ingin mengulas literatur secara sistematis tentang penerapan algoritma C4.5 dalam prediksi gempa bumi, mengevaluasi akurasi dan efektivitasnya, serta mengeksplorasi potensi pengembangan metode ini dalam konteks sistem peringatan dini di Indonesia [11]. Penelitian ini juga dimaksudkan untuk memberikan kontribusi konseptual dan teknis bagi pengembangan model prediksi bencana berbasis kecerdasan buatan yang adaptif dan skalabel [12].

METODOLOGI PENELITIAN

1. *Review Literatur Sistematis:* Peneliti mengumpulkan dan mengulas artikel dari jurnal ilmiah dan prosiding yang membahas prediksi gempa

- menggunakan algoritma C4.5 [13][27].
2. *Analisis Data:* Studi-studi terpilih dianalisis untuk mengidentifikasi variabel yang umum digunakan dalam prediksi, seperti magnitudo, kedalaman, koordinat episenter, dan waktu kejadian [14].
 3. *Evaluasi Akurasi dan Implikasi:* Akurasi model dalam studi dibandingkan dan dikaitkan dengan efektivitas implementasi dalam sistem peringatan gempa [15].
 4. *Klasifikasi Referensi:* Sebanyak 25 referensi dipilih berdasarkan relevansi dengan topik dan kredibilitas ilmiahnya, Berikut ini, tinjauan pustaka tentang Analisis prediksi Gempa bumi menggunakan algoritma C4.5 ini akan di sajikan ke dalam sebuah tabel tinjauan literatur pada Tabel 1:

N O	PENULIS	TAH UN	JUDUL
1	Quinlan, J. R.	1993	Programs for Machine Learning
2	Hossain, M. S., Roy, D., & Saha, H. N.	2020	Earthquake prediction using machine learning and data mining

3	Arif, M. & Surya, I. B.	2021	Implementasi algoritma C4.5 dalam prediksi klasifikasi gempa bumi di Sumatera Barat
4	USGS	2020	<i>Earthquake Catalog</i>
5	BMKG	2021	<i>Data Katalog Gempa Indonesia</i>
6	Witten, I. H., Frank, E., & Hall, M. A.	2016	Data Mining: Practical Machine Learning Tools and Techniques
7	Jaiswal, K. S., & Wald, D. J.	2011	Rapid estimation of the economic consequences of global earthquakes
8	Kusumandari, R. D., & Pramudito, D. A.	2020	Penerapan Decision Tree dalam prediksi sebaran gempa bumi di Indonesia
9	Han, J., Kamber, M., & Pei, J.	2011	Data Mining: Concepts and Techniques

			seismic data
2 2	Setiawan, R.	2021	Implementasi sistem peringatan dini berbasis machine learning
2 3	Manullang, S. O., & Handayani, T.	2019	Evaluasi metode pohon keputusan pada data bencana
2 4	Das, A., & Roy, S.	2020	A study on earthquake prediction models using data mining techniques
2 5	BNPB	2022	Annual Disaster Report
2 6	Hetty Rohayani	2021	Penerapan Algoritma C4.5 untuk Prediksi Kelayakan Kredit Nasabah pada Bank XYZ
2 7	Hetty Rohayani, S. H. Lestari	2020	Sistem Pendukung Keputusan Penentuan Penerima Beasiswa Menggunakan

			Metode C4.5
2 8	Hetty Rohayani, A. Fadillah, N. Arifin	2022	Analisis Prediksi Risiko Diabetes Menggunakan Algoritma Decision Tree C4.5

HASIL DAN PEMBAHASAN

Hasil telaah literatur menunjukkan bahwa algoritma Decision Tree C4.5 memberikan hasil yang menjanjikan dalam klasifikasi potensi gempa bumi [17]. Penggunaan algoritma ini dalam studi prediksi bencana sangat penting karena mampu membentuk pohon keputusan dari data historis aktivitas seismik dengan struktur aturan yang mudah diinterpretasikan [18]. Dalam penelitian yang dilakukan oleh beberapa peneliti di Jepang, Turki, India, dan Indonesia, C4.5 digunakan untuk mengidentifikasi pola dari dataset gempa bumi dengan atribut utama seperti magnitudo, kedalaman, waktu kejadian, jarak dari patahan aktif, dan frekuensi aktivitas seismik lokal [19].

Secara teknis, algoritma C4.5 bekerja dengan menghitung information gain dari setiap atribut dalam dataset untuk menentukan atribut mana yang paling efektif dalam memisahkan data ke dalam kelas-kelas gempa [1], [6]. Hal ini memungkinkan sistem untuk membentuk cabang-cabang pohon berdasarkan nilai-nilai ambang tertentu, seperti: Jika magnitudo > 5.0 dan kedalaman < 70 km, serta frekuensi aktivitas di lokasi > 3 kali dalam 14 hari terakhir, maka klasifikasi: Risiko Tinggi [20]. Model

prediktif semacam ini sangat berguna karena dapat dijadikan dasar dalam mengembangkan sistem peringatan dini (Early Warning System) yang dapat memberikan notifikasi ke masyarakat berdasarkan parameter yang diukur sensor-sensor seismik [13], [21].

Beberapa studi kasus yang ditinjau menunjukkan performa model yang cukup baik. Misalnya, studi oleh Hossain dkk. (2020) di Nepal mencatat akurasi 88,3% dalam mengklasifikasikan potensi gempa dengan menggunakan dataset dari USGS [2], [4]. Penelitian lain oleh Arif dan Surya (2021) di daerah Sumatera Barat menemukan bahwa C4.5 mampu memberikan hasil klasifikasi yang valid dengan presisi 81%, terutama dalam membedakan kejadian gempa dangkal dan gempa dalam [3].

C4.5 juga dinilai unggul dibanding metode lain seperti regresi logistik atau Naïve Bayes karena fleksibilitasnya dalam mengelola data yang memiliki kombinasi nilai numerik dan kategorikal [11], [12], [22]. Selain itu, hasil klasifikasinya tidak hanya berupa angka atau skor, melainkan berbentuk aturan yang sangat membantu pengambil keputusan dalam memahami dasar klasifikasi model [23][28].

Selain dari sisi teknis, C4.5 juga mudah diimplementasikan dalam perangkat lunak terbuka seperti WEKA, RapidMiner, atau diintegrasikan dalam framework Python menggunakan pustaka seperti scikit-learn [6]. Hal ini menjadikan algoritma ini tidak hanya akurat, tapi juga ekonomis dan mudah diterapkan dalam berbagai institusi, termasuk di negara berkembang [24].

Namun, model ini bukan tanpa keterbatasan. Salah satu tantangan utama adalah potensi overfitting ketika model terlalu dalam dan terlalu banyak cabangnya [9]. Untuk itu, proses pruning sangat penting agar model tetap optimal. Selain itu, C4.5 tidak menangkap korelasi spasial antar lokasi gempa, sehingga masih terbatas dalam aspek geografi. Untuk mengatasi hal ini, beberapa penelitian terbaru mulai menggabungkan C4.5 dengan algoritma spasial seperti K-Means clustering atau Geographic Information System (GIS) untuk menyempurnakan analisis wilayah rawan [19], [25].

Di Indonesia, potensi implementasi model C4.5 sangat besar, terutama karena negara ini memiliki akses ke dataset gempa dari BMKG yang telah terdokumentasi dengan baik [5]. Pemerintah dapat memanfaatkan model ini dalam sistem SIG berbasis web yang terhubung dengan data real-time dari sensor seismik di lapangan [21]. Misalnya, ketika sensor mencatat adanya gempa susulan dengan pola mirip sebelumnya, sistem dapat secara otomatis memunculkan notifikasi berdasarkan aturan model C4.5 [18]. Lebih jauh, penerapan model prediksi ini juga bisa ditingkatkan melalui kolaborasi antar lembaga seperti BMKG, BPPT, LAPAN, serta universitas-universitas yang memiliki keahlian di bidang geofisika dan ilmu data [16].

KESIMPULAN

Penelitian ini menegaskan bahwa algoritma C4.5 memiliki potensi besar dalam bidang prediksi gempa bumi berbasis data historis seismik [1], [6]. Dengan struktur pohon keputusan yang dihasilkan, model ini tidak hanya mampu mengklasifikasikan risiko gempa dengan cukup akurat, tetapi juga menghasilkan aturan logika yang dapat dengan mudah dimengerti dan diterapkan oleh praktisi kebencanaan [17].

Penggunaan C4.5 dalam berbagai studi menunjukkan bahwa akurasi model berkisar antara 75% hingga 90% tergantung pada kualitas dan kompleksitas dataset [2], [3], [12]. Kemampuan algoritma ini untuk mengolah data numerik dan kategorikal secara bersamaan menjadikannya unggul dibandingkan beberapa algoritma lainnya dalam konteks interpretabilitas dan kecepatan [22].

Namun, untuk meningkatkan akurasi dan skalabilitas model dalam konteks nyata, penelitian di masa depan disarankan untuk mengeksplorasi pendekatan hybrid seperti integrasi C4.5 dengan SVM, Random Forest, atau algoritma deep learning [11], [15]. Selain itu, pengumpulan data yang lebih rinci dari sensor seismik serta pengembangan model spasio-temporal akan menjadi kunci dalam menciptakan sistem prediksi gempa yang lebih akurat dan dapat diandalkan [18], [19].

Implementasi teknologi ini harus didukung oleh kolaborasi lintas disiplin, mulai dari ilmuwan data, ahli geologi, pemerintah daerah, hingga masyarakat umum [16],

[25]. Pendekatan prediktif semacam ini tidak hanya penting untuk peringatan dini, tetapi juga untuk perencanaan kota, mitigasi risiko bencana, serta pengembangan infrastruktur tangguh bencana di masa depan.

DAFTAR PUSTAKA

- [1] J. R. Quinlan, *C4.5: Programs for Machine Learning*. San Francisco, CA: Morgan Kaufmann, 1993.
- [2] M. S. Hossain, D. Roy, dan H. N. Saha, "Earthquake prediction using machine learning and data mining," *Int. J. Comput. Appl.*, vol. 176, no. 38, pp. 25–30, 2020.
- [3] M. Arif dan I. B. Surya, "Implementasi algoritma C4.5 dalam prediksi klasifikasi gempa bumi di Sumatera Barat," *J. Informatika dan Sistem Informasi*, vol. 7, no. 2, pp. 101–108, 2021.
- [4] USGS, "Earthquake Catalog," [Online]. Available: <https://earthquake.usgs.gov/>, 2020.
- [5] BMKG, "Data Katalog Gempa Indonesia," [Online]. Available: <https://www.bmkg.go.id/>, 2021.
- [6] I. H. Witten, E. Frank, dan M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed. San Francisco, CA: Morgan Kaufmann, 2016.
- [7] K. S. Jaiswal dan D. J. Wald, "Rapid estimation of the economic consequences of global earthquakes," *Earthquake Spectra*, vol. 27, no. 1, pp. 1–29, 2011.

- [8] R. D. Kusumandari dan D. A. Pramudito, "Penerapan Decision Tree dalam prediksi sebaran gempa bumi di Indonesia," *J. Geomatika dan Geodesi*, vol. 8, no. 1, pp. 45–55, 2020.
- [9] J. Han, M. Kamber, dan J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. San Francisco, CA: Morgan Kaufmann, 2011.
- [10] D. Kusnadi, "Sistem klasifikasi daerah rawan gempa menggunakan metode C4.5 dan GIS," *J. Teknol. dan Sist. Komput.*, vol. 9, no. 3, pp. 219–228, 2021.
- [11] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [12] M. M. Rahman dan M. S. Hossain, "Comparative analysis of decision tree algorithms for earthquake prediction," *Int. J. Comput. Sci. Issues*, vol. 15, no. 6, pp. 30–38, 2018.
- [13] D. Pratiwi dan R. Rakhmawati, "Sistem peringatan dini gempa bumi berbasis data mining," *J. Tek. Informatika*, vol. 6, no. 1, pp. 55–61, 2020.
- [14] N. Yuliana dan S. Hadi, "Penggunaan algoritma pohon keputusan untuk prediksi gempa dangkal dan dalam," *J. Sains Komput. dan Informatika*, vol. 10, no. 2, pp. 135–144, 2022.
- [15] B. Bhattacharya dan S. Ghosh, "A hybrid approach to earthquake prediction using decision trees and clustering," *Int. J. Disaster Risk Reduct.*, vol. 60, p. 102268, 2021.
- [16] E. Turban dan J. E. Aronson, *Decision Support Systems and Intelligent Systems*, 7th ed. Upper Saddle River, NJ: Pearson, 2005.
- [17] A. Yilmaz dan T. Ince, "Earthquake prediction using hybrid machine learning algorithms in Turkey," *Nat. Hazards*, vol. 106, no. 2, pp. 1041–1062, 2021.
- [18] M. Basu dan S. Ghosh, "Seismic risk assessment using decision trees," *Earthq. Eng. Struct. Dyn.*, vol. 45, no. 5, pp. 675–692, 2016.
- [19] D. Suharto dan M. Sari, "Analisis spasial prediksi gempa bumi dengan integrasi K-Means dan decision tree," *J. Tek. Sipil dan Perencanaan*, vol. 21, no. 1, pp. 34–42, 2019.
- [20] F. Nurdiansyah dan M. Rachmat, "Pemanfaatan data mining untuk mitigasi bencana gempa bumi," *J. Informatika dan Teknologi*, vol. 11, no. 1, pp. 23–31, 2021.
- [21] M. Ibrahim dan B. Baharuddin, "Decision tree-based earthquake prediction using historical seismic data," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 3, pp. 70–76, 2020.
- [22] R. Setiawan, "Implementasi sistem peringatan dini berbasis machine learning," *J. Sist. Informasi*, vol. 13, no. 2, pp. 119–128, 2021.
- [23] S. O. Manullang dan T. Handayani, "Evaluasi metode pohon keputusan pada data bencana," *J. Matematika dan Sains*, vol. 25, no. 2, pp. 89–98, 2019.
- [24] A. Das dan S. Roy, "A study on earthquake prediction models using data mining techniques," *Procedia Comput. Sci.*, vol. 167, pp. 1200–1209, 2020.
- [25] BNPB, "Annual Disaster Report," Badan Nasional Penanggulangan Bencana, Jakarta, Indonesia, 2022.
- [26] H. Rohayani, "Penerapan Algoritma C4.5 untuk Prediksi Kelayakan Kredit Nasabah pada

Bank XYZ,” *J. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 2, pp. 134–142, 2021.

[27] H. Rohayani and S. H. Lestari, “Sistem Pendukung Keputusan Penentuan Penerima Beasiswa Menggunakan Metode C4.5,” *J.*

Informatika: J. Pengemb. IT, vol. 7, no. 1, pp. 15–23, 2020.

[28] H. Rohayani, A. Fadillah, and N. Arifin, “Analisis Prediksi Risiko Diabetes Menggunakan Algoritma Decision Tree C4.5,” *J. Sist. dan Teknol Inf. (JUSTIN)*, vol. 5, no. 3, pp. 201–209, 2022.