

**APLIKASI *WEB* PREDIKSI DAMPAK GEMPA DI INDONESIA
MENGUNAKAN METODE *DECISION TREE* DENGAN ALGORITMA C4.5**

SKRIPSI

Diajukan untuk memenuhi sebagai syarat memperoleh gelar
Sarjana Matematika



Oleh

Diory Pribadi Sinaga

1905589

**PROGRAM STUDI MATEMATIKA
DEPARTEMEN PENDIDIKAN MATEMATIKA
FAKULTAS PENDIDIKAN MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS PENDIDIKAN INDONESIA**

2023

LEMBAR HAK CIPTA

APLIKASI *WEB* PREDIKSI DAMPAK GEMPA DI INDONESIA MENGUNAKAN METODE *DECISION TREE* DENGAN ALGORITMA C4.5

Oleh

Diory Pribadi Sinaga

Diajukan sebagai syarat untuk memperoleh gelar Sarjana
Matematika pada program studi Matematika

©Diory Pribadi Sinaga

Universitas Pendidikan Indonesia

April 2023

Hak Cipta dilindungi undang-undang.

Skripsi ini tidak boleh diperbanyak seluruhnya atau sebagian dengan dicetak ulang, difotokopi, atau cara lainnya tanpa izin dari penulis.

LEMBAR PENGESAHAN

DIORY PRIBADI SINAGA

**APLIKASI *WEB* PREDIKSI DAMPAK GEMPA DI INDONESIA
MENGUNAKAN METODE *DECISION TREE* DENGAN ALGORITMA C4.5**

Disetujui dan disahkan oleh pembimbing

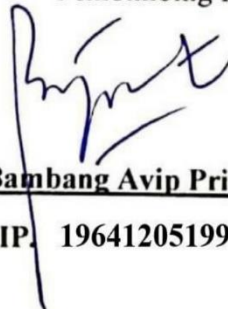
Pembimbing I



Dra. Hj. Rini Marwati, M.S.

NIP. 196606251990012001

Pembimbing II



Dr. Bambang Avip Priatna, M.Si

NIP. 19641205199031001

Mengetahui,

Ketua Departemen Pendidikan Matematika



Dr. H. Dadang Juandi, M.Si.

NIP. 196401171992021001

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa skripsi dengan judul “Aplikasi Web Prediksi Dampak Gempa di Indonesia Menggunakan Metode *Decision Tree* dengan Algoritma C4.5”, ini beserta isinya adalah benar-benar karya saya sendiri. Saya tidak melakukan penjiplakan atau pengutipan dengan cara-cara yang tidak sesuai dengan etika ilmu yang berlaku dalam masyarakat keilmuan. Atas pernyataan ini saya siap menanggung risiko/sanksi yang dijatuhkan kepada saya apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya ini, atau ada klaim dari pihak lain terhadap keaslian karya saya ini.

Bandung, 26 April 2023

Yang membuat pernyataan



Diory Pribadi Sinaga

1905589

KATA PENGANTAR

Salam dari Penulis

Puji syukur penulis panjatkan kepada Tuhan Yang Maha Esa yang telah memberikan kasih dan karuniaNya, sehingga penulis dapat menyelesaikan penyusunan skripsi yang berjudul “Aplikasi *Web* Prediksi Dampak Gempa di Indonesia Menggunakan Metode *Decision Tree* dengan Algoritma C4.5”. Skripsi ini diajukan untuk memenuhi syarat memperoleh gelar sarjana Matematika di Universitas Pendidikan Indonesia.

Skripsi ini terdiri dari lima Bab. Bab I pendahuluan berisi mengenai latar belakang, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan. Bab II berisi teori-teori yang berkaitan dengan penelitian ini. Bab III berisi metodologi penelitian berupa data penelitian, variabel penelitian, dan langkah-langkah penelitian. Bab IV berisi hasil dan pembahasan penelitian. Terakhir Bab V berisi kesimpulan dan saran penelitian.

Penulis menyadari masih ada kekurangan pada skripsi ini yang disebabkan oleh keterbatasan kemampuan penulis. Oleh karena itu, penulis menerima saran dan kritik dari pihak manapun yang dapat menyempurnakan skripsi ini. Demikian skripsi ini penulis buat, semoga bermanfaat bagi banyak pihak yang membaca. Terima kasih...

Bandung, 26 April 2023

Yang membuat pernyataan



Diory Pribadi Sinaga

1905589

UCAPAN TERIMA KASIH

Penulis menyadari bahwa penyelesaian skripsi ini tidak terlepas dari bimbingan, dukungan, bantuan, dan doa dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Ibu Dra. Hj. Rini Marwati, M.S., selaku Dosen Pembimbing I yang telah memberikan waktunya untuk memberikan bimbingan, kritik, dan saran selama penulis mengerjakan skripsi ini mulai dari awal hingga selesai.
2. Bapak Dr. Bambang Avip Priatna, M.Si, selaku Dosen Pembimbing II yang telah memberikan waktunya untuk memberikan bimbingan, kritik, dan saran selama penulis mengerjakan skripsi ini mulai dari awal hingga selesai.
3. Ibu Fitriani Agustina, S.Si., M.Si., selaku Dosen Pembimbing akademik yang telah memberikan waktunya untuk membantu pendampingan akademik perkuliahan.
4. Bapak Dr. Cece Kustiawan, M.Si, selaku Ketua Program Studi Matematika, Universitas Pendidikan Indonesia.
5. Bapak Dr. H. Dadang Juandi, M.Si., selaku Ketua Departemen Pendidikan Matematika, Universitas Pendidikan Indonesia.
6. Seluruh Dosen di lingkungan Departemen Pendidikan Matematika, Universitas Pendidikan Indonesia.
7. Kedua orang tua yang telah memberikan motivasi, dukungan materil, kasih sayang, dan doa kepada penulis sehingga penyusunan skripsi ini berjalan dengan lancar.
8. Alvandi Sinaga, selaku kakak penulis yang telah membiayai dan memberikan fasilitas selama proses pendidikan penulis.

Penulis bersyukur atas dukungan yang dilakukan semua pihak sehingga penulis dapat menyelesaikan skripsi ini, semoga Tuhan Yang Maha Esa memberikan balasan kepada semua pihak di atas.

ABSTRAK

Prediksi merupakan hal yang sangat penting bahkan sangat dibutuhkan karena dapat mengurangi risiko bila diketahui terlebih dahulu. Prediksi dapat dilakukan dengan menggunakan data *mining* berupa klasifikasi ke dalam kelas yang sudah ditentukan. *Decision Tree* merupakan suatu metode klasifikasi yang dapat digunakan untuk melakukan suatu prediksi pada kasus tertentu. Dalam hal ini kasus yang diteliti yaitu berkaitan dengan prediksi dampak gempa yang dirasakan berdasarkan Skala Intensitas Gempabumi BMKG (SIG-BMKG). *Decision Tree* memiliki beberapa algoritma diantaranya ID3, CART, dan C4.5. Pada penelitian ini algoritma yang dipakai adalah algoritma C4.5 karena di samping mudah untuk digunakan, algoritma ini tahan terhadap *missing value* dan dapat menangani variabel kontinu. Penelitian ini mengkonstruksi dan menganalisis performa model *Decision Tree* dengan algoritma C4.5 pada kasus prediksi dampak gempa di Indonesia, serta mengembangkan aplikasi *web* yang dapat digunakan pengguna. Dalam proses pengembangan model, penelitian ini menggunakan metodologi CRISP-DM (*Cross-Industry Standard Process for Data Mining*). Model yang dipakai dalam memprediksi adalah model dengan performa terbaik. Penelitian ini menggunakan teknik validasi *Split Validation* dan *k-fold Cross Validation* dalam menemukan model terbaik. Model yang terbaik diperoleh pada iterasi pertama *10-fold Cross Validation*. Model terbaik tersebut digunakan pada aplikasi *web* untuk memprediksi dampak gempa di Indonesia.

Kata Kunci : Data *mining*, Klasifikasi, *Decision Tree*, C4.5, Model, CRISP-DM

ABSTRACT

Prediction is very important even needed because it can reduce risk if known in advance. Predictions can be made using data mining in the form of classification into predetermined classes. Decision Tree is a classification method that can be used to make a prediction in a particular case. In this case, the case studied is related to the prediction of the impact of the earthquake felt based on the BMKG Earthquake Intensity Scale (SIG-BMKG). Decision Tree has several algorithms including ID3, CART, and C4.5. In this study, the algorithm used is the C4.5 algorithm because besides being easy to use, this algorithm is resistant to missing values and can handle continuous variables. This research constructs and analyzes the performance of the Decision Tree model with the C4.5 algorithm in the case of earthquake impact prediction in Indonesia, as well as developing a web application that can be used by users. In the model development process, this research uses the CRISP-DM (Cross-Industry Standard Process for Data Mining) methodology. The model used in predicting is the model with the best performance. This research uses Split Validation and k-fold Cross Validation validation techniques in finding the best model. The best model was obtained in the first iteration of 10-fold Cross Validation. The best model is used in a web application to predict the impact of earthquakes in Indonesia.

Keywords : *Data mining, Classification, Decision Tree, C4.5, Model ,CRISP-DM*

DAFTAR ISI

LEMBAR PENGESAHAN	i
LEMBAR PERNYATAAN	ii
KATA PENGANTAR	iii
UCAPAN TERIMA KASIH.....	iv
ABSTRAK	v
ABSTRACT	vi
DAFTAR ISI.....	vii
DAFTAR GAMBAR	x
DAFTAR TABEL.....	xi
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	3
1.3. Batasan Masalah.....	3
1.4. Tujuan Penelitian.....	4
1.5. Manfaat Penelitian.....	4
1.6. Sistematika Penulisan.....	4
BAB II KAJIAN TEORI.....	6
2.1. Gempa Bumi.....	6
2.2. Skala MMI (<i>Modified Mercalli Intensity</i>).....	7
2.3. Skala Intensitas Gempabumi BMKG (SIG-BMKG).....	7
2.4. <i>Data Mining</i>	9
2.5. Klasifikasi.....	10
2.6. <i>Clustering</i>	11
2.6.1. Metode Clustering.....	12
2.7. <i>k-Means Clustering</i>	14
2.8. <i>Sampling</i>	14
2.9. <i>Simple Random Sampling</i>	16
2.10. <i>Systematic Random Sampling</i>	16
2.11. Pohon Keputusan (<i>Decision Tree</i>)	17
2.12. Algoritma ID3 (<i>Iterative Dichotomiser 3</i>)	19
2.13. Algoritma C4.5 (<i>Classifier Version 4.5</i>)	20

2.14. Validasi Model	20
2.14.1 <i>Split Validation</i>	21
2.14.2 <i>k-fold Cross Validation</i>	22
2.15. <i>Confusion Matrix</i>	22
2.16 <i>Javascript</i>	24
2.17. <i>NodeJs</i>	25
2.17. <i>ExpressJs</i>	26
BAB III METODOLOGI PENELITIAN	27
3.1. Jenis dan Sumber Data Penelitian	27
3.2. Variabel Penelitian	28
3.3. Langkah-langkah Penelitian	28
3.4. Langkah-langkah Penerapan Metode <i>Decision Tree</i> (C4.5)	30
3.5. Langkah-langkah Evaluasi Model <i>Decision Tree</i> (C4.5)	30
3.6. Rancangan Aplikasi Web	30
3.6.1 Rancangan Proses Jalannya Aplikasi.....	30
3.6.2 Rancangan Tampilan	31
BAB IV HASIL DAN PEMBAHASAN	32
4.1. Analisis Data Gempa.....	32
4.2. Persiapan Data (<i>Data Preparation</i>).....	33
4.2.1. Implementasi <i>Systematic Random Sampling</i>	33
4.2.2. Implementasi <i>Euclidean Distance</i>	34
4.2.3. Implementasi <i>k-means Clustering</i>	35
4.3. Model <i>Decision Tree</i> (C4.5).....	37
4.4. <i>Confusion Matrix</i>	42
4.5. <i>Best Model</i> (Model Terbaik)	45
4.5.1 <i>Split Validation</i>	45
4.5.2 <i>k-fold Cross Validation</i>	47
4.6. <i>Deployment</i> (Penyebaran)	50
4.7. Program Aplikasi Web Prediksi Dampak Gempa.....	51
4.7.1. Tampilan Utama Aplikasi	51
4.7.2 Tampilan Halaman Statistik Data Gempa	51
4.7.3 Tampilan Visual Kategori.....	52
4.7.4. Tampilan Halaman Prediksi.....	52

4.7.5. Tampilan Validasi (<i>Split Validation</i>) dan Evaluasi Model.....	53
4.7.6. Tampilan validasi (<i>k-fold Cross Validation</i>) dan Evaluasi Model	53
4.7.7. Tampilan Perbandingan Performa Model <i>Split Validation</i>	54
4.7.8. Tampilan Perbandingan Performa Model <i>k-fold Cross Validation</i>	54
BAB V KESIMPULAN DAN SARAN.....	56
5.1. Kesimpulan.....	56
5.2. Saran	56
DAFTAR PUSTAKA	58
LAMPIRAN.....	60

DAFTAR GAMBAR

Gambar 2.1 Lapisan Bumi	6
Gambar 2.2 <i>Hierararchical Methods</i>	13
Gambar 2.3 Algoritma <i>k-means</i>	14
Gambar 2.4 Metode Penarikan Sampel (<i>sampling</i>)	15
Gambar 2.5 <i>Systematic Random Sampling</i>	16
Gambar 2.6 Contoh Struktur <i>Decision Tree</i>	17
Gambar 2.7 Algoritma <i>Decision Tree</i>	18
Gambar 2.8 <i>Split Validation</i>	21
Gambar 2.9 <i>k-fold Cross Validation</i>	22
Gambar 2.10 Posisi Urutan Bahasa Pemrograman di Dunia	25
Gambar 2.11 <i>framework Express</i>	26
Gambar 3.1 Data Gempa (BMKG)	27
Gambar 3. 2 CRISP-DM	28
Gambar 3.3 Rancangan Tampilan Halaman Utama Aplikasi	31
Gambar 3.4 Tampilan Halaman Melakukan Prediksi	31
Gambar 4.1 Kelas Data Gempa.....	32
Gambar 4.2 Pohon Keputusan.....	42
Gambar 4.3 Perbandingan Performa Model <i>Split Validation</i>	47
Gambar 4.4 Perbandingan Performa <i>10-fold Cross Validation</i>	49
Gambar 4.5 Pohon Keputusan Model Terbaik.....	50
Gambar 4.6 Halaman Utama Aplikasi	51
Gambar 4.7 Tampilan Statistik Data Gempa	51
Gambar 4.8 <i>Box Plot</i> Data Gempa	52
Gambar 4.9 Tampilan Visual Kategori	52
Gambar 4.10 Halaman Prediksi	53
Gambar 4.11 Tampilan <i>Split Validation</i>	53
Gambar 4.12 Tampilan Halaman <i>k-fold Cross Validation</i>	54
Gambar 4.13 Tampilan Halaman Performa <i>Split Validation</i>	54
Gambar 4.14 Tampilan Halaman Performa Model <i>k-fold Cross Validation</i>	55
Gambar 4.15 Tampilan Halaman Performa Model <i>k-fold Cross Validation</i>	55

DAFTAR TABEL

Tabel 2.1 Skala Intensitas Gempabumi BMKG.....	8
Tabel 2.2 <i>Confusion Matrix</i>	23
Tabel 4.1 Statistik Data Gempa	32
Tabel 4.2 <i>Confusion Matrix</i>	44
Tabel 4.3 <i>Confusion Matrix Split 6:4</i>	46
Tabel 4.4 <i>Confusion Matrix Split 7:3</i>	46
Tabel 4.5 <i>Confusion Matrix Split 8:2</i>	46
Tabel 4.6 <i>Confusion Matrix Split 9:1</i>	46
Tabel 4.7 Performa Model <i>Split Validation</i>	47
Tabel 4.8 Performa Model <i>10-fold Cross Validation</i>	48

DAFTAR PUSTAKA

- Anam, C., & Santoso, H. B. (2018). Perbandingan Kinerja Algoritma C4.5 dan Naive Bayes untuk Klasifikasi Penerima Beasiswa. *Energy - Jurnal Ilmiah Ilmu-Ilmu Teknik*, 8(1).
- Anggara, M., Sujiani, H., & Nasution, H. (2016). Pemilihan distance measure Pada K-Means clustering Untuk Pengelompokan member di Alvaro fitness. *Jurnal Sistem dan Teknologi Informasi (JUSTIN)*, 1(1).
- Buulolo, E., Silalahi, N., Fadlina, & Rahim, R. (2017). C4.5 Algorithm To Predict the Impact of the Earthquake. *International Journal of Engineering Research & Technology (IJERT)*, 6(2).
- Dandy Bayu Bramasta. (2022, November 26). *UPDATE Gempa Susulan di Cianjur: Terjadi 259 Kali, Terbesar M 4,2 dan Terkecil 1,2*. KOMPAS.
- Dr. Eng. Yeri Sutopo, M. P. M. T., Prof. Dr. Achmad Slamet, M. S., & Offset, A. (t.t.). *Statistik Inferensial*. Penerbit Andi.
<https://books.google.co.id/books?id=jVJLDwAAQBAJ>
- Dunhan, M. (2003). *Data Mining: Introductory and Advanced Topics*. Prentice Hall. *Engineering*.
- Haekal, B. V., Ernawati, I., & Chamidah, N. (2021). Klasifikasi Kepuasan Pengguna Layanan Aplikasi Shopee Menggunakan Metode Decision Tree C4.5. *Informatik : Jurnal Ilmu Komputer*, 17(3). <https://doi.org/10.52958/iftk.v17i3.3648>
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining: Concepts and Techniques*. Dalam *Data Mining: Concepts and Techniques*. <https://doi.org/10.1016/C2009-0-61819-5>
- Hendro D Situmorang. (2022, Desember 30). *BMKG: Indonesia Diguncang 10.792 Kali Gempa Sepanjang 2022*. beritasatu.
- HSSINA, B., MERBOUHA, A., EZZIKOURI, H., & ERRITALI, M. (2014). A comparative study of decision tree ID3 and C4.5. *International Journal of Advanced Computer Science and Applications*, 4(2).
<https://doi.org/10.14569/specialissue.2014.040203>
- Khadafy, A. R., & Wahono, R. S. (2015). Penerapan Naïve Bayes Untuk Mengurangi Data Noise Pada Klasifikasi Multi Kelas Dengan Decision Tree. *Journal of Intelligent Systems*, 1(2).
- Lakshmi, B. N., Indumathi, T. S., & Ravi, N. (2016). A Study on C.5 Decision Tree Classification Algorithm for Risk Predictions During Pregnancy. *Procedia Technology*, 24. <https://doi.org/10.1016/j.protcy.2016.05.128>
- Larose, D. T., & Larose, C. D. (2014). *Discovering knowledge in data: an introduction to data mining* (Vol. 4). John Wiley & Sons.

- Law, A., & Kelton, D. (1999). *Simulation Modeling and Analysis*. McGraw Hill.
- Mukminin, A., & Riana, D. (2017). Komparasi Algoritma C4 . 5 , Naïve Bayes Dan Neural Network Untuk Klasifikasi Tanah. *Jurnal Informatika Universitas Bina Sarana Informatika*, 4(1).
- Rosandy, T. (2016). Perbandingan Metode Naive Bayes Classifier Dengan Metode Decision Tree (C4. 5) Untuk Menganalisa Kelancaran Pembiayaan (Study Kasus: KSPPS/BMT Al-Fadhila. *Jurnal Teknologi Informasi Magister*, 2(01), 52–62.
- Scheafer, richard L., III, W. M., Itt, R. L., & Gerow, K. G. (2017). Elementary survey sampling seventg edition. Dalam *Brooks/Cole* (Vol. 91).
- Setio, P. B. N., Saputro, D. R. S., & Bowo Winarno. (2020). Klasifikasi Dengan Pohon Keputusan Berbasis Algoritme C4.5. *PRISMA, Prosiding Seminar Nasional Matematika*, 3.
- Singh, S. and Gupta, P. (2014). Comparative study ID3, cart and C4 . 5 Decision tree algorithm: a survey. *International Journal of Advanced Information Science and Technology (IJAIST)*, 27(27).
- Suherman, & Muzaky, I. (2019). Analisis Penjualan Barang Laris Dan Kurang Laris Terhadap Percetakan Awfa Digitl Printing Menggunakan Metode Decision Tree Dengan Optimasi Algoritma Genetika. *Jurnal Teknologi Pelita Bangsa*, 10(1).
- Suryana. (2012). Metodologi Penelitian : Metodologi Penelitian Model Praktis Penelitian Kuantitatif dan Kualitatif. *Universitas Pendidikan Indonesia*.
- Tommy, P. (2011). *ANALISA RISIKO KERUSAKAN BANGUNAN RUMAH TINGGAL TIPE 142 AKIBAT GEMPA (STUDI KASUS: RUMAH TINGGAL DI SEBUAH PERUMAHAN DI KOTA DEPOK)*. Universitas Indonesia.