

**Prediksi Curah Hujan Wilayah Kota Palembang Melalui Penerapan
Machine Learning Dengan Metoda *Decision Tree***

SKRIPSI

**Sebagai salah satu syarat untuk memperoleh gelar
Sarjana Sains Program Studi Fisika**



Oleh:

RIDHO DERRI SAFUTRA

NIM. 08021281823044

JURUSAN FISIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS SRIWIJAYA

2022

LEMBAR PENGESAHAN

**PREDIKSI CURAH HUJAN WILAYAH KOTA PALEMBANG MELALUI
PENERAPAN *MACHINE LEARNING* DENGAN METODA *DECISION TREE***

SKRIPSI

Sebagai salah satu syarat memperoleh gelar
Sarjana Sains Bidang Studi Fisika

Oleh :

RIDHO DERRI SAFUTRA

NIM. 08021281823044

Indralaya, 18 November 2022

Menyetujui

Pembimbing I



Dr. Azhar Kholiq Affandi, M.S.
NIP. 196109151989031003

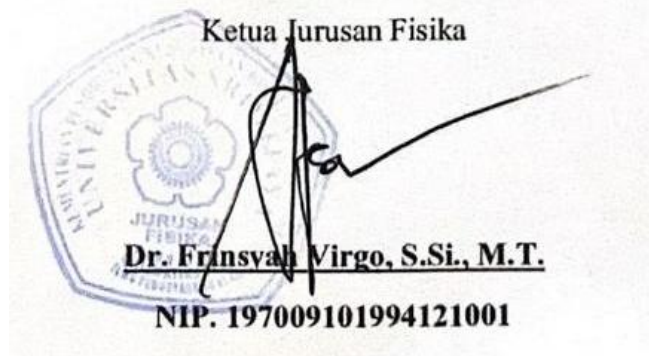
Pembimbing II



Erni S.Si., M.Si
NIP. 197606092003122002

Mengetahui,

Ketua Jurusan Fisika



Dr. Fransyah Virgo, S.Si., M.T.
NIP. 197009101994121001

PERNYATAAN ORISINALITAS

Saya yang bertanda tangan di bawah ini, Mahasiswa Jurusan Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sriwijaya:

Nama : Ridho Derri Safutra

NIM : 08021281823044

Judul TA : Prediksi Curah Hujan Wilayah Kota Palembang Melalui Penerapan
Machine Learning Dengan Metoda *Decision Tree*

Dengan ini menyatakan bahwa skripsi yang saya susun dengan judul tersebut adalah asli atau orisinalitas dan mengikuti etika penulisan karya tulis ilmiah sampai pada waktu skripsi ini diselesaikan, sebagai salah satu syarat untuk memperoleh gelar sarjana sains di program studi Fisika Universitas Sriwijaya.

Demikian surat pernyataan ini dibuat dengan sebenar-benarnya tanpa ada paksaan dari pihak manapun. Apabila di kemudian hari terdapat kesalahan ataupun keterangan palsu dalam surat pernyataan ini, maka saya siap bertanggung jawab secara akademik dan bersedia menjalani proses hukum yang telah ditetapkan.

Indralaya, 18 November 2022

Yang menyatakan



Ridho Derri Safutra

NIM. 08021281823044

KATA PENGANTAR

Segala puji bagi Allah, Tuhan Semesta Alam Yang Maha Pengasih lagi Maha Penyayang. Ucapan syukur penulis haturkan kehadirat Allah SWT karena atas rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan penelitian Tugas Akhir yang berjudul **“Prediksi Curah Hujan Wilayah Kota Palembang Melalui Penerapan *Machine Learning* Dengan Metoda *Decision Tree*”**. Tulisan ini dibuat dalam rangka memenuhi persyaratan untuk memperoleh gelar sarjana sains (S.Si) di Jurusan Fisika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Sriwijaya.

Selama proses penulisan & penelitian ini, tak henti-hentinya penulis mendapat berbagai bantuan dan dukungan, serta saran guna penyempurnaan tugas akhir ini. Oleh Karena itu, dalam kesempatan ini Penulis juga ingin menyampaikan rasa terima kasih yang sebesar-besarnya kepada :

1. Allah SWT, karena atas rahmat dan izin-Nya sehingga Tugas Akhir ini dapat terselesaikan, *Alhamdulillahirabbil'alamiin*.
2. Kedua orang tuaku, Bapak Drs. saugani dan Mama Patimah, yang senantiasa mendoakan, mendukung, dan membantu, serta memberikan semangat dan motivasi pada penulis untuk terus berjuang menyelesaikan studi.
3. Bapak Dr. Frinsyah Virgo, S.Si., M.T selaku Ketua Jurusan Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam.
4. Bapak Dr. Azhar Kholiq Affandi, M.S selaku dosen Pembimbing I Tugas Akhir yang telah memberikan arahan, ilmu, dan masukan, serta dukungan moral yang tiada henti-hentinya.
5. Ibu Erni, S.Si., M.Si selaku dosen Pembimbing II Tugas Akhir dan juga dosen Pembimbing Akademik yang telah mengarahkan dan memberi saran baik selama penyelesaian tugas akhir maupun selama berkuliah di Jurusan Fisika.
6. Bapak Drs. Arsali, M.Sc dan Menik Ariani, M.Si selaku dosen Penguji Tugas Akhir yang telah memberikan ilmu, saran, dan kritikan dalam penyusunan tugas akhir ini.
7. Bapak Ibu dosen Jurusan Fisika yang telah mentransfer ilmu pengetahuan dan pengalaman luar biasa selama perkuliahan.

8. Jajaran staf Jurusan Fisika, Kak David dan Babe, yang telah membantu dalam proses administrasi dan surat-menyurat.
9. Rekan-rekan satu penelitian *Machine Learning* dan diskusi ; Edi Iskara, Riko Fadla, & Qoriatu Khoirunnisa yang turut memberikan bantuan dan berbagi suka duka penelitian.
10. Teman-teman korps satuan asisten Laboratorium Fisika Komputasi yang telah mengajarkan dasar-dasar programan dan kerja sama tim.
11. Teman dekat tim *feeder*; Jimi, Hadi, Ihsan, Amar, Syaugi, Yogi, Bagas, Falen, Khoiril, Agung, & Wansya, yang telah membagi semangat dan suka duka selama tahun terakhir perkuliahan
12. Teman-teman satu angkatan, Fisika 2018, AMF18I, yang tak bisa penulis sebutkan namanya satu per satu.
13. Seluruh Web Penyedia Programan ; *Stack Overflow*, *Kaggle*, *Jupyter Notebook*, *Github* yang telah memberikan batuan berupa *source code* dalam menghadapi masalah *error* program.
14. Penulis buku Dios Kurniawan, M.Sc dengan judul “Pengenalan *Machine Learning* Python” yang digunakan penulis sebagai pedoman pemahaman dasar-dasar *coding*.
15. Utamanya diri sendiri, yang berusaha memperlebar zona nyaman dan mempertahankan idealisme, guna penyempurnaan tugas akhir.

Indralaya, 18 November 2022

Penulis



Ridho Derri Safutra

NIM. 08021281823044

**PREDIKSI CURAH HUJAN WILAYAH KOTA PALEMBANG MELALUI
PENERAPAN *MACHINE LEARNING* DENGAN METODA *DECISION TREE***

Oleh:

Ridho Derri Safutra

NIM. 08021281283044

ABSTRAK

Kota Palembang merupakan salah satu kota yang rentan terhadap bencana banjir dan erosi karena memiliki intensitas curah hujan yang sangat tinggi setiap tahunnya. Oleh sebab itu untuk mengantisipasi bencana tersebut dapat dilakukan melalui prediksi curah hujan. Penelitian ini bertujuan untuk klasifikasi cuaca (curah hujan) berdasarkan kategori hujan dengan menghasilkan hujan atau tidak hujan pada kota Palembang Sumatera Selatan periode Januari 2011-Desember 2021 menggunakan data sekunder dari website NASA POWER. Metoda decision tree yang digunakan adalah algoritma CART (*Classification And Regression Tree*) dengan *Confusion Matrix* sebagai pengujian model dan pohon keputusan sebagai visualisasi data *training*. Berdasarkan Evaluasi model menunjukkan algoritma CART mempunyai performa yang cukup bagus dengan akurasi senilai 81,74%, presisi 83,25% dan *recall* 94,05%.

Kata Kunci: CART (*Classification And Regression Tree*), *decision tree*, *machine learning*, Kota Palembang, prediksi curah hujan

Indralaya, 18 November 2022

Menyetujui

Pembimbing I



Dr. Azhar Kholiq Affandi, M.S.
NIP. 196109151989031003

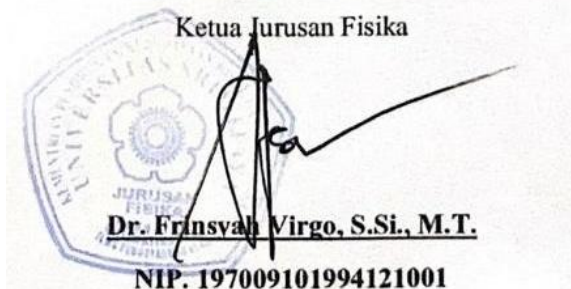
Pembimbing II



Erni S.Si., M.Si
NIP. 197606092003122002

Mengetahui,

Ketua Jurusan Fisika



Dr. Fransyah Virgo, S.Si., M.T.
NIP. 197009101994121001

**RAINFALL FORECASTING IN PALEMBANG USING MACHINE LEARNING
WITH THE DECISION TREE METHOD**

Oleh:

Ridho Derri Safutra

NIM. 08021281283044

ABSTRACT

Palembang City is one of those cities vulnerable to flood and erosion because it has high rainfall intensity in each year. One of possible method to anticipate this problem is using rainfall forecasting. This research is present a prediction of weather (Rainfall) based on the rain category with the result that rain and not rain in Palembang City on the period January 2011-December 2021 using data from website NASA POWER. The decision tree method used to be a algorithm CART (Classification And Regression Tree) with Confusion Matrix test parameters and “pohon keputusan” as visualization from training data. Based on the results of the model evaluation, it shows that the CART algorithm has a fairly good performance in classifying with an accuracy value of 81.74%, precision 83,25%, and recall 94,05%.

Keyword: CART (Classification And Regression Tree), decision tree, machine learning, Kota Palembang, rainfall forecasting

Indralaya, 18 November 2022

Menyetujui

Pembimbing I



Dr. Azhar Kholiq Affandi, M.S.
NIP. 196109151989031003

Pembimbing II



Erni S.Si., M.Si
NIP. 197606092003122002

Mengetahui,

Ketua Jurusan Fisika



Dr. Frinsyah Virgo, S.Si., M.T.
NIP. 197009101994121001

DAFTAR ISI

LEMBAR PENGESAHAN	II
PERNYATAAN ORISINALITAS	III
KATA PENGANTAR	II
ABSTRAK	V
ABSTRACT	VI
DAFTAR ISI	VII
DAFTAR GAMBAR	IX
DAFTAR TABEL	X
DAFTAR LAMPIRAN	XI
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	2
1.3. Batasan Masalah	2
1.4. Tujuan Penelitian	2
1.5. Manfaat Penelitian	3
BAB II TINJAUAN PUSTAKA	4
2.1. Prediksi Curah Hujan	4
2.2. <i>Machine Learning</i>	4
2.2.1. Proses <i>Machine Learning</i>	6
2.2.2. <i>Training Dataset</i>	6
2.2.3. Mengukur Kinerja Model	7
2.3. <i>Supervised Learning</i>	8
2.4. <i>Decision Tree</i>	9
2.4.1. <i>Decision Tree Clasification</i> dengan <i>Scikit-learn</i>	10
2.4.2. Metode CART	11
2.4.2. Contoh Pembuatan Pohon Keputusan	13
BAB III METODE PENELITIAN	16
3.1. Waktu dan Tempat Pengerjaan	16
3.2. Deskripsi Lokasi Penelitian	16
3.3. Alat Bantu Penelitian	17
3.4. Tahapan-Tahapan Penelitian	17
3.5. Analisa Hasil Model <i>Machine Learning</i>	20
BAB IV HASIL DAN PEMBAHASAN	22
4.1. Pemahaman Data	22

4.2. Persiapan Data	23
4.3. Pembuatan Model	25
4.4. Pengujian Hasil Model	27
BAB V PENUTUP	30
5.1. Kesimpulan	30
5.2. Saran	30
DAFTAR PUSTAKA	31
LAMPIRAN	33

DAFTAR GAMBAR

Gambar 2.1. Ruang lingkup <i>Artificial Intelligence</i> dan <i>Machine Learning</i>	5
Gambar 2.2. Perbedaan pemrograman klasik dan pemrograman <i>Machine Learning</i>	5
Gambar 2.3. <i>Training dataset</i> dan <i>testing dataset</i>	8
Gambar 2.4. <i>Supervised learning</i>	9
Gambar 2.5. <i>Decision tree</i>	10
Gambar 2.6. Hasil <i>splitting</i> data pada setiap parameter.....	14
Gambar 2.7. Pohon keputusan.....	15
Gambar 3.1. Peta lokasi penelitian.....	16
Gambar 3.2. Bagan alir tahapan-tahapan penelitian.....	19
Gambar 4.1. Contoh model pohon keputusan (kedalaman tiga pohon).....	26
Gambar 4.2. Contoh empat leaf node model pohon keputusan.....	26
Gambar 4.3. Hasil <i>Confusion Matrix</i> model.....	27
Gambar 4.4. Perbandingan nilai <i>training score</i> dan <i>testing score</i>	29

DAFTAR TABEL

Tabel 2.1. Contoh <i>structured data</i>	7
Tabel 2.2. Penggunaan <i>scikit-learn</i>	11
Tabel 2.3. Data pembuatan pohon keputusan.....	13
Tabel 2.4. Hasil perhitungan <i>gini average</i> dan <i>information gain</i>	14
Tabel 3.1. <i>Confusion Matrix</i>	20
Tabel 4.1. <i>Dataset</i> NASA POWER.....	22
Tabel 4.2. Kategori hujan BMKG.....	23
Tabel 4.3. Proses <i>filter</i> data.....	23
Tabel 4.4. Proses transformasi data.....	24
Tabel 4.5. Parameter <i>dataset</i> NASA POWER.....	25
Tabel 4.6. Evaluasi model.....	28

DAFTAR LAMPIRAN

Lampiran 1 <i>Script</i> Program.....	34
Lampiran 2 Gambar <i>Website</i> NASA POWER.....	38
Lampiran 3 Contoh Tabel <i>Dataset</i> NASA POWER dengan 39 data awal.....	39
Lampiran 4 Gambar Keseluruhan Pohon Keputusan.....	40
Lampiran 5 Tabel Hasil Prediksi Percobaan Data Baru.....	41
Lampiran 6 Tabel Perbandingan Data Aktual Dan Prediksi Pada <i>Testing Dataset</i>	42

BAB I

PENDAHULUAN

1.1. Latar Belakang

Iklim adalah kondisi cuaca rata-rata untuk suatu daerah atau zona dalam jangka waktu yang lama yang diukur dengan beberapa parameter cuaca. Iklim di Indonesia adalah iklim tropis yang relatif hangat dan lembab di sepanjang tahun. Dengan iklim ini, Indonesia hanya memiliki dua musim yang berbeda yaitu musim kemarau dan musim hujan. Pada musim hujan, Provinsi Sumatera Selatan memiliki intensitas curah hujan yang sangat tinggi terutama di kota Palembang sebesar 588 mm pada bulan Desember (Badan Pusat Statiska Kota Palembang, 2021). Akibat curah hujan tinggi beberapa daerah di Kota Palembang mengalami bencana erosi dan banjir, ini diperkuat dengan letak geografis Kota Palembang yang berada di dataran rendah yang terdiri dari sungai, rawa, dan gambut.

Sebagai bentuk antisipasi dari beberapa bencana alam yang terjadi dapat dilakukan dengan memprediksi curah hujan. Dengan memakai data cuaca beberapa waktu lalu, maka dapat diprediksi banyaknya curah hujan yang akan terjadi dalam masa mendatang. Ada banyaknya metode dalam memprediksi curah hujan di suatu tempat, salah satunya dengan *machine learning*.

Machine learning adalah studi mengenai algoritma mesin yang dapat mengetahui pola pola didalam data, dengan maksud untuk mengubah berbagai jenis data menjadi sebuah perbuatan yang nyata dengan campur tangan manusia sesedikit mungkin. Pembelajaran *machine learning*, memungkinkan kita membuat mesin (komputer) yang “belajar” dari data yang sudah ada, kemudian mesin tersebut bisa membuat keputusan secara sendiri tanpa memprogram ulang (Kurniawan, 2021). Pada penelitian ini, data yang diberikan berupa kumpulan data cuaca beserta beberapa parameter yang mengindikasinya.

Berbagai penelitian yang telah dilakukan dengan beberapa metode *machine learning* diantaranya oleh Mulyani et al., 2019 yang bertema “Prediksi Curah Hujan di Kabupaten Majalengka dengan Menggunakan Algoritma Regresi” dan Hasanah et al., 2021 yang bertema “Implementasi CRISP-DM Model Menggunakan Metode *Decision Tree* dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir”, pada kedua penelitian tersebut memiliki tingkat akurasi yang tinggi akan tetapi output yang

dihasilkan berbeda dimana di metode regresi hanya berupa nilai numerik sedangkan di metode *decision tree* sampai menghasilkan output kategorikal berupa hujan atau tidak hujan sehingga dapat mudah dimengerti oleh peneliti lain, hal ini sesuai dengan kelebihan utama dari model *decision tree* yaitu mudah diinterpretasikan dan sangat cocok untuk visualisasi data. Maka dari itu, penelitian ini akan berfokus menggunakan metode *decision tree* dalam prediksi curah hujan. Dengan bahasa pemrograman yang digunakan adalah python akan dihasilkan model *machine learning* teroptimal dalam menghadapi kasus. Setelah mendapatkan model teroptimal dibandingkan dengan data aktual yang diperoleh dari pengukuran manual pada *website* NASA POWER, serta dapat memprediksi curah hujan pada tahun berikutnya.

1.2. Rumusan Masalah

1. Bagaimana perancangan model *machine learning* untuk memprediksi curah hujan menggunakan metode *decision tree*.
2. Bagaimana perbandingan antara data aktual dan data kalkulasi yang dihasilkan dari model.

1.3. Batasan Masalah

1. Data yang digunakan adalah data cuaca harian pada tahun 2011-2021 dengan parameter antara lain curah hujan, kelembaban udara, tekanan udara, temperatur minimum, maksimum dan kecepatan angin di Kota Palembang.
2. Implementasi *Machine Learning* prediksi curah hujan dengan menggunakan metode *decision tree*.
3. Dataset yang digunakan berasal dari *website* NASA POWER (*National Aeronautics and Space Administration* dan *The Prediction Of Worldwide Energy Resources*).

1.4. Tujuan Penelitian

1. Memperoleh model *machine learning* untuk memprediksi curah hujan menggunakan metode *decision tree*.
2. Mendapatkan implentasi model dari *decision tree* berdasarkan data cuaca di Kota Palembang Sumatera Selatan.

1.5. Manfaat Penelitian

1. Dapat menghasilkan estimasi curah hujan dengan nilai akurasi yang besar sehingga dapat diterima sebagai aplikasi prediksi curah hujan di Kota Palembang.
2. Dapat memprediksi curah hujan berdasarkan parameter-parameter cuaca.

DAFTAR PUSTAKA

- Abijono, H., Santoso, P., & Anggreini, N. L. (2021). *Algoritma Supervised Learning Dan Unsupervised Learning Dalam Pengolahan Data*. Jurnal Teknologi Terapan: G-Tech, 4(2), 315–318.
- Amrizal, V. dan Aini, Q. 2013. *Kecerdasan Buatan*. Jakarta : Halaman Moeka.
- Arista, A. (2022). *Comparison Decision Tree and Logistic Regression Machine Learning Classification Algorithms to determine Covid-19*. Sinkron, 7(1), 59–65.
- Badan Pusat Statistik Kota Palembang. 2021."Curah Hujan Kota Palembang 2009-2021". <https://palembangkota.bps.go.id/> (di akses tanggal 25 Mei 2022).
- Breiman L., Friedman J.H Olshen R.A & Stone C.J. 1984. *Classification AndRegression Tree*. New York, NY: Chapman And Hall
- Chollet, F. 2018. *Deep Learning with Python*. New York : Manning Publication.
- Hadion Wijouo. (2021). *Teknik Pengambilan Keputusan*. Insan Cendikia Mandiri.
- Hasanah, M, A., Soim, S & Handayani, A, S. (2021). *Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART untuk Prediksi Curah Hujan Berpotensi Banjir*. Journal of Applied Informatics and Computing (JAIC), 5(2), 103-108.
- Indrabayu, Harun, A., Pallu , M., S., dan Achmad, A., 2012, *Prediksi Curah Hujan dengan Jaringan Syaraf Tiruan*. Prosiding Hasil Penelitian Fakultas Teknik Elektro, Universitas Hasanuddin : Makasar.
- Khoeri, I., & Iskandar, D. (2021). *Implementasi Machine Learning Dengan Decision Tree Algoritma C4.5 Dalam Penerimaan Karyawan Baru Pada Pt. Gitareksa*.Jurnal Sosial Dan Teknologi (SOSTECH), 1(7), 615–623.
- Kurniawan, D. 2021. *Pengenalan Machine Learning dengan Python*. Jakarta : Elex Media Komputindo.
- Latifah, R., Wulandari, E. S., & Kreshna, P. E. (2019). *Model Decision Tree Untuk Prediksi Jadwal KerjaMenggunakanScikit-Learn*.Jurnal Universitas Muhammadiyah Jakarta, 1–6.

- Lewis, M.D dan Roger, J. 2000. *An Introduction to Classification and Regression Tree(CART) Analysis*. Presented at the 2000 Annual Meeting of Society For Academy Emergency Medicine in San Fransisco, California
- Putra, J. W. G. 2020. *Pengenalan Konsep Pembelajaran Mesin dan Deep Learning*. Tokyo : Tokyo Institut Of Technology.
- Perdana, D, A., Zakaria, A., & Sumiharni. (2015). *Studi Pemodelan Curah hujan sintetik dari beberapa stasiun di wilayah Pringsewu*. JRSDD, 3(1), 45-56.
- Mulyani et al., (2019). *Prediksi Curah Hujan Di Kabupaten Majalengka Dengan Menggunakan Algoritma Regresi*. Jurnal Sistem Informasi Dan Teknologi Informasi, 8(1), 67-77.
- Nafi, N. (2015). *Algoritma Cart Dalam Penentuan Pohon Keputusan Sertifikasi Guru*. Nopember, 7(2), 41–47.
- Negnevitsky, M. 2005. *Artificial Intelligence : A Guide to Intelligent System (2nd ed)*. London : Pearson Education.
- Russel, S. dan Norvig, P. 2010. *Artificial Intelligence : A Modern Approach (3rd ed)*. New Jersey : Pearson.
- Shwart, S.S. and David, S. B. 2014. *Understanding Machine Learning From Theory to Algorithms*. New York : Cambridge Univeristy Press.
- Sodik, F., Dwi, B., & Kharisudin, I. (2020). *Perbandingan Metode Klasifikasi Supervised Learning pada Data Bank Customers Menggunakan Python*. Jurnal Matematika, 3, 689–694.
- Tando , J., Komalig, H., dan Nainggolan , N., 2016, *Prediksi Jumlah penumpang Kapal Laut di Pelabuhan Laut Manado Menggunakan Model ARMA*. Jurnal de Cartesian 5(2) : 95.