

**IMPLEMENTASI CRISP-DM MODEL MENGGUNAKAN
METODE *DECISION TREE* DENGAN ALGORITMA CART
PADA PREDIKSI KEJADIAN BANJIR**



TUGAS AKHIR

**Disusun Untuk Melengkapi Salah Satu Syarat Menyelesaikan Pendidikan
Sarjana Terapan Pada Jurusan Teknik Elektro
Program Studi Teknik Telekomunikasi
Politeknik Negeri Sriwijaya**

Oleh :

**MSY AULIA HASANAH
061740351484**

**POLITEKNIK NEGERI SRIWIJAYA
PALEMBANG
2021**

TUGAS AKHIR
IMPLEMENTASI CRISP-DM MODEL MENGGUNAKAN
METODE *DECISION TREE* DENGAN ALGORITMA CART
PADA PREDIKSI KEJADIAN BANJIR



Disusun Untuk Memenuhi Syarat Menyelesaikan Pendidikan Sarjana
Terapan Pada Jurusan Teknik Elektro Program Studi Teknik
Telekomunikasi Politeknik Negeri Sriwijaya

Oleh :

Nama : Msy Aulia Hasanah (061740351484)

Dosen Pembimbing I : Sopian Soim, S.T., M.T.

Dosen Pembimbing II : Dr. Ade Silvia Handayani, S.T., M.T

POLITEKNIK NEGERI SRIWIJAYA
PALEMBANG
2021

**IMPLEMENTASI CRISP-DM MODEL MENGGUNAKAN
METODE *DECISION TREE* DENGAN ALGORITMA CART
PADA PREDIKSI KEJADIAN BANJIR**



TUGAS AKHIR

**Disusun Untuk Memenuhi Syarat Menyelesaikan Pendidikan Sarjana
Terapan Pada Jurusan Teknik Elektro Program Studi Teknik
Telekomunikasi Politeknik Negeri Sriwijaya**

Oleh :

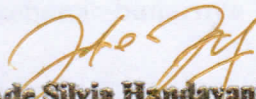
**MSY AULIA HASANAH
061740351484**

Palembang, Agustus 2021

Pembimbing I

Pembimbing II


Sopian Soim, S.T., M.T
NIP. 197103142001121001


Dr. Ade Silvia Handayani, S.T., M.T.
NIP. 197609302000032002

Mengetahui,

**Ketua Jurusan
Teknik Elektro**

**Ketua Program Studi
Teknik Telekomunikasi**


Ir. Iskandar Lutfi, M.T.
NIP. 196501291991031002


Lindawati, S.T., M.T.I
NIP. 197105282006042001

SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini menyatakan:

Nama : Msy Aulia Hasanah
Jenis Kelamin : Perempuan
Tempat, Tanggal Lahir : Palembang, 29 Oktober 1999
Alamat : Jl. K.H Wahid Hasyim Lr. Tajur NO. 823, Kec. Seberang Ulu 1, Kel. 5 Ulu. Kota Palembang
NPM : 061740351484
Program Studi : Teknik Telekomunikasi
Jurusan : Teknik Elektro
Judul Skripsi/Laporan Akhir* : ***Implementasi CRISP-DM Model Menggunakan Metode Decision Tree dengan Algoritma CART Pada Prediksi Kejadian Banjir***

Menyatakan dengan sesungguhnya bahwa:

1. Skripsi/Laporan Akhir* ini adalah hasil karya saya sendiri serta bebas dari tindakan plagiasi, dan semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar.
2. Dapat menyelesaikan segala urusan terkait pengumpulan revisi Skripsi/Laporan Akhir yang sudah disetujui oleh dewan penguji paling lama 1 bulan setelah ujian Skripsi/Laporan Akhir.
3. Dapat menyelesaikan segala urusan peminjaman/penggantian alat/buku dan lainnya paling lama 1 bulan setelah ujian Skripsi/Laporan Akhir.

Apabila dikemudian hari diketahui ada pernyataan yang terbukti tidak benar dan tidak dapat dipenuhi, maka saya siap bertanggung jawab dan menerima sanksi tidak diikutsertakan dalam prosesi wisuda serta dimasukkan dalam daftar hitam oleh Jurusan Teknik Elektro sehingga berdampak tertundanya pengambilan Ijazah & Transkrip (ASLI & COPY). Demikian surat pernyataan ini dibuat dengan sebenar-benarnya dan dalam keadaan sadar tanpa paksaan.

Palembang, Agustus 2021
Yang Menyatakan,

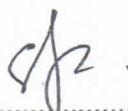
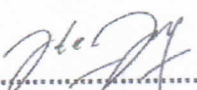


Msy Aulia Hasanah

Mengetahui,

Pembimbing I Sopian Soim, S.T., M.T.

Pembimbing II Dr. Ade Silvia Handayani, S.T., M.T.


.....

.....

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

Nama : Msy Aulia Hasanah

NIM : 061740351484

Judul : **Implementasi CRISP-DM Model Menggunakan Metode *Decision Tree* Dengan Algoritma CART Pada Prediksi Kejadian Banjir.**

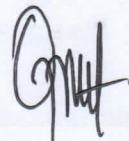
Menyatakan bahwa laporan tugas akhir saya merupakan hasil kerja saya sendiri dan bukan hasil penjiplakan/plagiat. Apabila ditemukan unsur penjiplakan/plagiat dalam tugas akhir ini kecuali yang telah disebutkan sumbernya, maka saya bersedia menerima sanksi akademik dari Politeknik Negeri Sriwijaya.

Demikianlah, Pernyataan ini saya buat dalam keadaan sadar dan tidak dipaksakan.



Palembang, Agustus 2021

Penulis,



Msy Aulia Hasanah

MOTTO DAN PERSEMBAHAN

الرَّحْمَنِ الرَّحِيمِ بِسْمِ

Motto :

- Ssesungguhnya sesudah kesulitan itu ada kemudahan. Maka apabila kamu telah selesai (dari sesuatu urusan). Kerjakanlah dengan sungguh-sungguh (urusan) yang lain. Dan hanya kepada tuhanmulah hendaknya kamu berharap. (Q.S Alam Nasrah : 6,7,8).
- Kesabaran dan ketekunan adalah kunci keberhasilan
- Saya datang, saya pulang, saya bimbingan, saya ujian, saya revisi dan sayamaju untuk menang
- Hidup memberi pilihan dan hidup juga memberi kesempatan, maka yang pantas untuk dilakukan adalah sebuah kepastian dalam memilih dan memanfaatkan kesempatan yang ada dengan sebaik-baiknya

Kami persembahkan kepada :

- ALLAH SWT
- Baginda Nabi Muhammad SAW
- Kedua orang tua yang senantiasa mendoakan keberhasilan kami
- Seluruh keluarga yang kami sayangi
- Teman-teman kami di Politeknik Negeri Sriwijaya yang selalu bekerja sama dalam menuntut ilmu dan orang spesial yang selalu memberikandukungan skripsi ini hingga selesai
- Almamater kami yang senantiasa mendampingi kami selama kuliah

ABSTRAK

Implementasi CRISP-DM Model Menggunakan Metode *Decision Tree* Dengan Algoritma CART Pada Prediksi Kejadian Banjir.

(2021 : x v + 69 halaman + 39 gambar + 13 tabel + 11 lampiran)

MSY AULIA HASANAH

061740351484

JURUSAN TEKNIK ELEKTRO

PROGRAM STUDI TEKNIK TELEKOMUNIKASI

POLITEKNIK NEGERI SRIWIJAYA

Banjir merupakan salah satu bencana yang sangat berpengaruh dan merugikan baik itu dari segi ekonomi, lingkungan maupun keselamatan masyarakat. Faktor utama penyebab banjir ialah curah hujan dan debit air sungai. Tingginya intensitas curah hujan dapat mempengaruhi jumlah volume debit air yang mengalir pada saluran sungai yang melebihi kapasitas alirannya sehingga terjadilah banjir. Untuk mengatasi masalah tersebut dibutuhkan teknologi dan informasi untuk mengelola data dengan tujuan memprediksi dan mengestimasi tentang apa yang terjadi kedepan, melakukan analisis dengan membaca serta mengetahui pola pendekatan tersembunyi dari sekumpulan data melalui proses teknik statistik, matematik, kecerdasan buatan (*Artificial Intelegent*) dan *Machine Learning* untuk mengestrak pengetahuan atau menemukan pola dari suatu data yang besar yang disebut dengan *data mining*. Metode yang diterapkan dalam proses *data mining* yaitu dengan menggunakan teknik klasifikasi pohon keputusan atau *Decision Tree* dengan algoritma CART (*Classification And Regression Tree*) untuk menghasilkan suatu pola pengetahuan dengan mengklasifikasi data iklim, data curah hujan, dan data debit air sungai. Tahap analisis menerapkan pendekatan *data mining* dengan proses standarisasi CRISP-DM (*Cross Industry Standard Process for Data Mining*) kemudian hasil klasifikasi dievaluasi dengan uji parameter *Confusion Matrix* untuk mengetahui tingkat akurasi.

Kata Kunci : CRISP-DM, CART, *Confusion Matrix*, *Decision Tree*, *Data Mining*

ABSTRACT

Implementasi CRISP-DM Model Menggunakan Metode *Decision Tree* Dengan Algoritma CART Pada Prediksi Kejadian Banjir.

(2021 : xv + 69 pages + 39 pictures + 13 tables + 1 1 appendix

MSY AULIA HASANAH

061740351484

ELECTRICAL ENGINEERING DEPARTMENT

**PROGRAM OF STUDY IN APPLIED GRADUATION OF THE
TELECOMMUNICATION ENGINEERING
STATE POLYTECHNIC OF SRIWIJAYA**

Floods are one of the most influential and detrimental disasters in terms of economy, environment and public safety. The main factors causing flooding are rainfall and river water discharge. The high intensity of rainfall can affect the volume of water discharge flowing in river channels that exceeds its flow capacity, resulting in flooding. To overcome this problem, technology and information are needed to manage data with the aim of predicting and estimating what will happen in the future, conducting analysis by reading and knowing hidden approaches patterns from a set of data through statistical, mathematical, artificial intelligence and machine learning techniques. to extract knowledge or find patterns from a large data called data mining. The method applied in the data mining process is by using a decision tree classification technique or Decision Tree with the CART (Classification And Regression Tree) algorithm to generate a pattern of knowledge by classifying climate data, rainfall data, and river water discharge data. The analysis phase applies a data mining approach with the CRISP-DM (Cross Industry Standard Process for Data Mining) standardization process, then the classification results are evaluated by using the Confusion Matrix parameter test to determine the level of accuracy.

Kata Kunci : CRISP-DM, CART, *Confusion Matrix*, *Decision Tree*, *Data Mining*

KATA PENGANTAR

Puji dan syukur penulis panjatkan kepada Tuhan Yang Maha Esa atas berkat dan kasih karunia-Nya, sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul **“IMPLEMENTASI CRISP-DM MODEL MENGGUNAKAN METODE DECISION TREE DENGAN ALGORITMA CART PADA PREDIKSI KEJADIAN BANJIR”**. Tugas Akhir ini dibuat untuk memenuhi salah satu kurikulum di Jurusan Teknik Elektro Program Studi Teknik Telekomunikasi Politeknik Negeri Sriwijaya Palembang.

Dalam pelaksanaan penulisan Tugas Akhir ini, penulis mengucapkan terima kasih kepada Bapak **Sopian Soim, S.T., M.T** dan Ibu **Dr. Ade Silvia Handayani, S.T., M.T** selaku dosen pembimbing yang telah memberikan banyak bimbingan dan masukan yang membantu penulis dalam menyelesaikan tugas akhir ini. Selain itu penulis juga mengucapkan terima kasih kepada :

1. Allah SWT atas nikmatnya yang telah diberikan kepada saya, sehingga dapat menyelesaikan laporan Tugas Akhir dalam keadaan yang sehat dan tanpa kekurangan apapun.
2. Bapak Dr. Ing. Ahmad Taqwa, M.T., selaku Direktur Politeknik Negeri Sriwijaya;
3. Bapak Ir. Iskandar Lutfi, M.T., selaku Ketua Jurusan Teknik Elektro Politeknik Negeri Sriwijaya;
4. Bapak Destra Andika Pratama S.T., M.T, selaku Sekretaris Jurusan Teknik Elektro Politeknik Negeri Sriwijaya;
5. Ibu Lindawati, S.T., M.T., selaku Ketua Program Studi DIV Teknik Telekomunikasi Politeknik Negeri Sriwijaya;
6. Bapak / Ibu Dosen Program Studi DIV Teknik Telekomunikasi Politeknik Negeri Sriwijaya;
7. Orang Tua dan saudaraku yang tak henti-hentinya memberikan doa dan dorongan semangat;
8. Dwi Indra Saputra yang telah memberikan motivasi dan ikut membantu penulis untuk menyelesaikan Tugas Akhir ini.

9. Seluruh Staff dan Pengajar Teknik Elektro Program Studi Sarjana Terapan Teknik Elektro
10. Teman-temanku Seperjuangan dalam menyelesaikan Tugas Akhir ini.
Terutama Kelas 8 TEA Angkatan 2017
11. Semua pihak yang telah membantu dalam menyelesaikan laporan ini yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa dalam penyusunan tugas akhir ini masih terdapat kekurangan dan kekeliruan, baik mengenai isi maupun cara penulisan. Untuk itu penulis sangat mengharapkan saran dan kritik yang bersifat membangun. Penulis berharap semoga Tugas akhir ini dapat dilanjutkan menjadi Tugas Akhir yang bermanfaat bagi kita semua, umumnya para pembaca dan khususnya penulis serta bagi mahasiswa Politeknik Negeri Sriwijaya Teknik Elektro Program Studi Teknik Telekomunikasi.

Palembang, 21 Agustus 2021

Penulis

DAFTAR ISI

	Halaman
HALAMAN SAMPUL.....	i
HALAMAN JUDUL	ii
LEMBAR PENGESAHAN.....	iii
LEMBAR PERNYATAAN	iv
MOTTO DAN PERSEMBAHAN	v
ABSTRAK.....	vi
ABSTRACK.....	vii
KATA PENGANTAR.....	viii
DAFTAR ISI	x
DAFTAR GAMBAR.....	xiii
DAFTAR TABEL	xiv
DAFTAR LAMPIRAN	xv
 BAB I PENDAHULUAN	 1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	4
1.3 Ruang Lingkup dan Batasan Masalah	4
1.4 Tujuan	5
1.5 Manfaat	5
1.6 Metode Penulisan	6
1.7 Sistematika Penulisan	7
 BAB II TINJAUAN PUSTAKA.....	 8
2.1 Data Mining	8
2.2 Proses Data Mining.....	8
2.2.1 Proses Data Mining.....	8
2.2.2 Metode Data Mining	9
2.2.3 Pengetahuan.....	10
2.2.4 Evaluation.....	11
2.3 Teknik Klasifikasi Pada Data Mining	11
2.4 Decision Tree.....	12
2.5 Algoritma CART (Classification And Regression Tree).....	14
2.6 CRISP-DM (Cross Industry Standart Process for data mining).....	16
2.7 Confusion Matrix	18
2.8 Penelitian Sebelumnya	19
 BAB III METODOLOGI PENELITIAN	 21
3.1 Metode Penelitian	21
3.2 Prinsip dan Konsep Analisis.....	21
3.2.1 Objek Penelitian.....	21
3.2.2 Metode Pengumpulan Data	22
3.2.3 Perangkat yang Digunakan.....	22

3.3	Kerangka Penelitian	23
3.3.1	Studi Literatur	24
3.3.2	Pengumpulan Data	24
3.3.3	Pengelolaan Data Mining	25
3.3.4	Analisis Hasil	25
3.3.5	Laporan Tugas Akhir	25
3.4	Pengolahan Data Mining CRISP-DM	25
3.4.1	Business Understanding	26
3.4.2	Data Understanding	26
3.4.3	Data Preparation	26
3.4.4	Modelling	28
3.4.5	Evaluation	28
3.4.6	Deployment	28
BAB IV HASIL DAN PEMBAHASAN.....		29
4.1	Hasil Pengumpulan Data	29
4.2	Implementasi CRISP-DM Menggunakan Algoritma CART dalam Klasifikasi Data Iklim untuk Prediksi Curah Hujan Berpotensi Banjir	33
4.2.1	Business Understanding	33
4.2.2	Data Understanding	33
4.2.3	Data Preparation	35
4.2.3.1	Data cleaning	35
4.2.3.2	Data Selection	37
4.2.3.3	Data Integration	39
4.2.3.4	Data Transformation	39
4.2.4	Modelling	41
4.2.4.1	Proses Data Mining Menggunakan tools pycharm	42
4.2.5	Evaluation	47
4.2.6	Deployment	49
4.3	Analisis Hasil Klasifikasi Data Iklim untuk Prediksi Curah Hujan Berpotensi Banjir	49
4.4	Implementasi CRISP-DM Menggunakan Algoritma CART dalam Klasifikasi Data Iklim dan Data DAS untuk Prediksi Kejadian yang Berpotensi Banjir	50
4.4.1	Business Understanding	50
4.4.2	Data Understanding	50
4.4.3	Data Preparation	52
4.4.3.1	Data cleaning	52
4.4.3.2	Data Selection	53
4.4.3.3	Data Integration	55
4.4.3.4	Data Transformation	55
4.4.4	Modelling	56
4.4.4.1	Proses Data Mining Menggunakan tools pycharm	57
4.4.5	Evaluation	61
4.4.6	Deployment	63

4.5 Analisis Hasil Klasifikasi Data Iklim dan Data DAS untuk Prediksi Kejadian yang Berpotensi Banjir	63
BAB V KESIMPULAN DAN SARAN	65
5.1 Kesimpulan.....	65
5.2 Saran	66
DAFTAR PUSTAKA.....	67
LAMPIRAN	

DAFTAR GAMBAR

Gambar 2.1	Proses <i>Data Mining</i>	8
Gambar 2.2	Metode <i>Data Mining</i>	9
Gambar 2.3	Contoh Model <i>Decision Tree</i>	10
Gambar 2.4	Konsep Pohon Keputusan	12
Gambar 2.5	Konsep Dasar Pohon Keputusan	13
Gambar 2.6	Proses CRISP-DM.....	16
Gambar 3.1	Kerangka Penelitian.....	24
Gambar 3.2	Metodologi data maining CRISP-DM.....	26
Gambar 3.3	Proses Data Preparation	27
Gambar 3.4	Pemodelan	28
Gambar 4.1	Suhu Minimum	29
Gambar 4.2	Suhu Maksimum.....	30
Gambar 4.3	Suhu Rata-rata	30
Gambar 4.4	Kelembaban Rata-rata.....	31
Gambar 4.5	Lama Penyinaran Matahari	31
Gambar 4.6	Curah Hujan	32
Gambar 4.7	Debit Air Sungai.....	32
Gambar 4.8	Kejanggalan data iklim dan keterangan	36
Gambar 4.9	Dataset dengan Format <i>xlsx</i>	40
Gambar 4.10	Dataset dengan Format <i>csv</i>	41
Gambar 4.11	Import Dataset Iklim.....	42
Gambar 4.12	Source Coding Membaca Data.....	43
Gambar 4.13	Data Terbaca.....	43
Gambar 4.14	Sebagian Source Coding Klasifikasi Algoritma CART	44
Gambar 4.15	Rule yang terbentuk dari Decision Tree	45
Gambar 4.16	Classification Tree dalam Prediksi Curah Hujan.....	46
Gambar 4.17	Source Coding Confusion Matrix.....	48
Gambar 4.18	Confusion Matrix pada Algoritma CART.....	48
Gambar 4.19	Data Missing Value	53
Gambar 4.20	Dataset dengan Format <i>xlsx</i>	55
Gambar 4.21	Dataset dengan Format <i>csv</i>	56
Gambar 4.22	Import Dataset	57
Gambar 4.23	Source Coding Membaca Data.....	58
Gambar 4.24	Data Berhasil Terbaca	58
Gambar 4.25	Sebagian Source Coding Klasifikasi Algoritma CART	59
Gambar 4.26	Rule Decision Tree	60
Gambar 4.27	Classification Tree dalam Prediksi Potensi Banjir	60
Gambar 4.28	Source Coding Confusion Matrix.....	62
Gambar 4.29	Confusion Matrix pada Algoritma CART.....	62

DAFTAR TABEL

Tabel 2.1	Frekuensi Penggunaan Algoritma Pohon Keputusan	14
Tabel 2.2	Confution Matrix	18
Tabel 2.3	Penelitian Sebelumnya.....	19
Tabel 3.1	Perangkat Keras	23
Tabel 3.2	Perangkat Lunak.....	23
Tabel 4.1	Data IklimBMKG.....	33
Tabel 4.2	Keterangan Data	34
Tabel 4.3	Kategori Intensitas Hujan BMKG	35
Tabel 4.4	Data Perubahan Value curah Hujan.....	36
Tabel 4.5	Data Selection	38
Tabel 4.6	Data Curah Hujan dan Debit Air Sungai	50
Tabel 4.7	Keterangan Data	52
Tabel 4.8	Data Selection	54

DAFTAR LAMPIRAN

- Lampiran 1** Daftar Riwayat Hidup
- Lampiran 2** Lembar Kesepakatan Bimbingan TA Pembimbing I
- Lampiran 3** Lembar Kesepakatan Bimbingan TA Pembimbing II
- Lampiran 4** Lembar Konsultasi Pembimbing I
- Lampiran 5** Lembar Konsultasi Pembimbing I
- Lampiran 6** Lembar Rekomendasi
- Lampiran 7** Lembar Pelaksanaan Revisi
- Lampiran 8** *Letter of Acceptance*
- Lampiran 9** *Submitted Paper*
- Lampiran 10** *Source Coding*
- Lampiran 11** *Dataset*