

Analisis Klasifikasi Bunga Iris Menggunakan Algoritma C4.5

SKRIPSI



disusun oleh

Feri Irawan

17.11.1060

**PROGRAM SARJANA
PROGRAM STUDI IFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
2021**

Analisis Klasifikasi Bunga Iris Menggunakan Algoritma C4.5

SKRIPSI

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Sistem Informatika



disusun oleh

Feri Irawan

17.11.1060

**PROGRAM SARJANA
PROGRAM STUDI INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
2021**

PERSETUJUAN

SKRIPSI

Analisis Klasifikasi Bunga Iris Menggunakan Algoritma C4.5

yang dipersiapkan dan disusun oleh

Feri Irawan

17.11.1060

telah disetujui oleh Dosen Pembimbing Skripsi

8 Desember 2020

Dosen Pembimbing,

Norhikmah, M.Kom

NIK:190302245

PENGESAHAN

SKRIPSI

Analisis Klasifikasi Bunga Iris Menggunakan Algoritma C4.5

yang dipersiapkan dan disusun oleh

Feri Irawan

17.11.1060

telah dipertahankan di depan Dewan Penguji
16 Desember 2020

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Eli Pujastuti, M.Kom
NIK. 190302227

Anna Baita, M.Kom
NIK. 190302290

Norhikmah, M.Kom
NIK. 190302245

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
16 Desember 2020

DEKAN FAKULTAS ILMU KOMPUTER

Krisnawati, S.Si., M.T.
NIK. 190302038

PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

Yogyakarta, 8 Desember 2020



Feri Irawan

NIM. 17.11.1060

MOTTO

Selama Ada Niat dan Keyakinan Semua Akan Jadi Mungkin.



PERSEMBAHAN

Alhamdulillah, Segala puji bagi Allah SWT, kita memuji-Nya, dan meminta pertolongan, pengampunan serta petunjuk kepada-Nya. Kita berlandung kepada Allah dari kejahatan diri kita dan keburukan amal kita. Barang siapa mendapat dari petunjuk Allah, maka tidak akan ada yang menyesatkannya dan barang siapa yang sesat maka tidak ada pemberi petunjuk baginya. Aku bersaksi bahwa tidak ada tuhan selain Allah dan bahwa dan bahwa nabi Muhammad adalah hamba dan Rasul-Nya. Semoga doa, shalawat tercurah pada junjungan dari suri tauladan kita Nabi Muhammad SAW, keluarganya, dan sahabat serta siapa saja yang mendapat hingga hari kiamat. Aamiin.

Persembahan skripsi ini dan rasa terima kasih aku ucapkan untuk :

1. Keluarga tercinta, kedua orang tuaku, kakak - kakakku, serta adikku yang telah memberikan kasih sayang, do'a, dukungan serta motivasi baik secara moril maupun materil untuk selalu terikat dengan hukum syara' dan menjadi orang yang bahagia di dunia maupun di akhirat.
2. Teman – temanku kelas 17-Informatika-03 Universitas Amikom Yogyakarta yang senantiasa selalu membantu dan memberikan semangat dalam menjalani belajar.
3. Bapak / Ibu Dosen Universitas Amikom Yogyakarta yang memberikan pengajaran kepada saya agar selalu disiplin dalam melakukan segala hal.
4. Almamaterku tercinta Universitas Amikom Yogyakarta.

KATA PENGANTAR

Segala puji dan syukur kehadiran Allah SWT atas berkah, rahamat dan hidayah-Nya yang senantiasa dilimpahkan kepada penulis, sehingga bisa menyelesaikan skripsi dengan judul “ Analisis Klasifikasi Bunga Iris Menggunakan Algoritma C4.5” sebagai syarat untuk menyelesaikan Program Sarjana (S1) pada Program Sarjana Fakultas Ilmu Komputer dan Jurusan Informatika Universitas Amikom Yogyakarta.

Dalam penyusunan skripsi ini banyak hambatan serta rintangan yang penulis hadapi namun pada akhirnya dapat melaluinya berkat adanya bimbingan dan bantuan dari berbagai pihak baik secara moral maupun spiritual. Untuk itu pada kesempatan ini penulis menyampaikan ucapan terimakasih kepada:

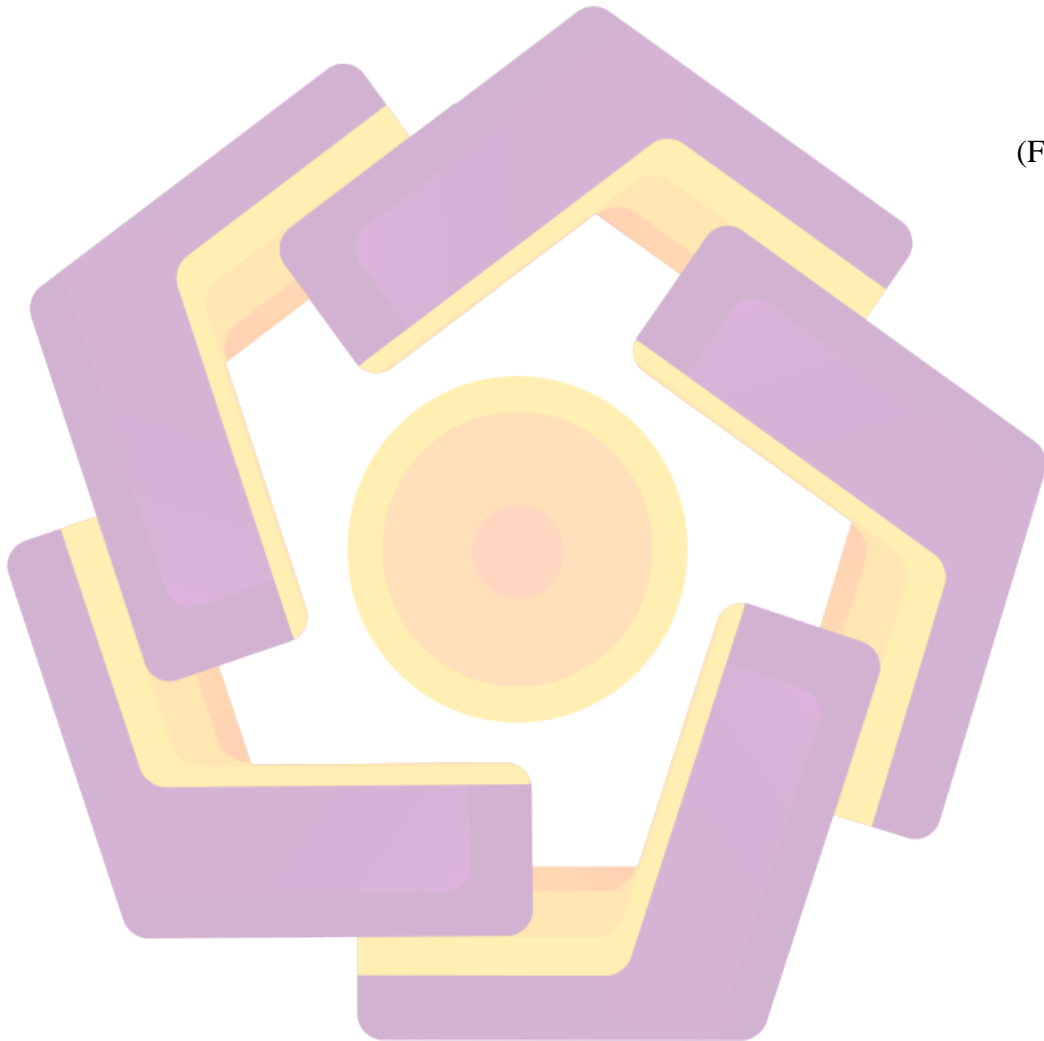
1. Kedua orang tua, ayahanda Wajiyo dan ibunda Surtinah yang senantiasa memberikan kasih sayang dan dukungan serta doa yang tiada henti-hentinya kepada penulis.
2. kakaku tercinta kakak Rohmad Irawan dan kakak Yaya Rogayah Ilhami serta adik tercinta Rizky Ahmad Vauzi yang telah menyemangati.
3. Segenap Keluarga dan teman yang telah menyemangati dan membantu penyelesaian skripsi.
4. Ibu Norhikmah, M.Kom selaku dosen Pembimbing Skripsi yang telah bersedia membimbing, memberikan tambahan ilmu, memberi solusi pada setiap permasalahan atas kesulitan dalam penulisan skripsi ini.
5. Seluruh Bapak/Ibu dosen Universitas Amikom Yogyakarta yang telah memberikan pengetahuan yang sangat bermanfaat.
6. Seluruh teman – teman seangkatan terutama Kelas 17 – Informatika – 03 yang selalu menemani perjuangan dimasa kuliah.
7. Dan semua pihak yang tidak dapat penulis sebut satu persatu yang telah membantu selama penyelesaian skripsi ini.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna dikarenakan terbatasnya pengalaman dan pengetahuan yang dimiliki penulis. Oleh karena itu, penulis mengharapkan segala bentuk saran serta masukan bahkan kritik yang membangun dari berbagai pihak. Semoga skripsi ini dapat bermanfaat bagi para pembaca dan semua pihak khususnya dalam bidang Informatika.

Yogyakarta, 8 Desember 2020

Penulis,

(Feri Irawan)



DAFTAR ISI

HALAMAN JUDUL.....	2
HALAMAN PERSETUJUAN	3
HALAMAN PENGESAHAN	4
HALAMAN PERNYATAAN.....	5
HALAMAN MOTTO	6
HALAMAN PERSEMBAHAN	7
KATA PENGANTAR	8
DAFTAR ISI	10
DAFTAR TABEL.....	13
DAFTAR GAMBAR	14
DAFTAR ISTILAH	16
INTISARI.....	17
ABSTRACT	18
BAB I PENDAHULUAN	
1.1 Latar Belakang	19
1.2 Rumusan Masalah	20
1.3 Batasan Masalah.....	20
1.4 Maksud dan Tujuan Penelitian	21
1.5 Manfaat Penelitian.....	21
1.6 Metode Penelitian	21
1.6.1 Dasar Penelitian	21
1.6.2 Studi Literatur.....	21
1.6.3 Pengumpulan Data	21
1.6.4 Pengolahan Data	21
1.6.5 Perancangan dan Desain Sistem	22
1.6.6 Implementasi	22
1.6.7 Pengujian	22
1.6.8 Hasil	22
1.7 Sistematika Penulisan	23

BAB II LANDASAN TEORI

2.1 Kajian Pustaka.....	24
2.2 Dasar Teori.....	27
2.2.1 Data Mining.....	27
2.2.2 Algoritma C4.5.....	29
2.2.3 Pengukuran Akurasi Klasifikasi.....	36
2.3 Aplikasi Web.....	37
2.3.1 HTML.....	37
2.3.2 CSS.....	38
2.3.3 PHP.....	38
2.3.4 Basis Data MySQL.....	39
2.4 Rancangan Aplikasi.....	39
2.4.1 UML.....	39
2.4.2 Use Case Diagram.....	39
2.4.3 Activity Diagram.....	41
2.4 Bunga Iris.....	42
BAB III ANALISIS DAN PERANCANGAN	
3.1 Alat dan Bahan.....	44
3.1.1 Alat.....	44
a. Perangkas Keras (Hardware).....	44
b. Perangkat Lunak (Software).....	44
3.1.2 Bahan.....	45
3.2 Alur penelitian.....	46
3.2.1 Dasar Penelitian.....	46
3.2.2 Studi Literatur.....	47
3.2.3 Pengumpulan Data.....	47
a. Sumber Data.....	47
b. Jenis Data.....	48
3.2.4 Pengolahan Data.....	48
a. Perbersihan Data (<i>data cleaning</i>).....	49
b. Integrasi Data (<i>data integration</i>).....	49

c. Seleksi data (data selection)	49
d. Transformasi data (data transformation)	49
e. Penambangan Data (Data Mining)	54
3.2.5 Perancangan dan Desain Sistem	54
a. UML	54
1) Use Case Diagram.....	54
2) Activity Diagram.....	58
b. Perancangan Database	69
c. Perancangan Tampilan Aplikasi.....	71
BAB IV IMPLEMENTASI DAN PEMBAHASAN	
4.1 Implementasi	80
4.1.1 Implementasi Perangkat Keras	80
4.1.2 Implementasi Perangkat Lunak	80
4.1.3 Implementasi Basis Data	80
4.1.4 Implementasi Algoritma C4.5	81
4.1.5 Implementasi Tampilan Aplikasi	82
4.2 Pengujian.....	88
4.2.1 Pengujian Akurasi.....	88
4.2.2 Pengujian Waktu Proses.....	90
4.3 Hasil.....	90
BAB V PENUTUP	
5.1 Kesimpulan	91
5.2 Saran	91
DAFTAR PUSTAKA	92

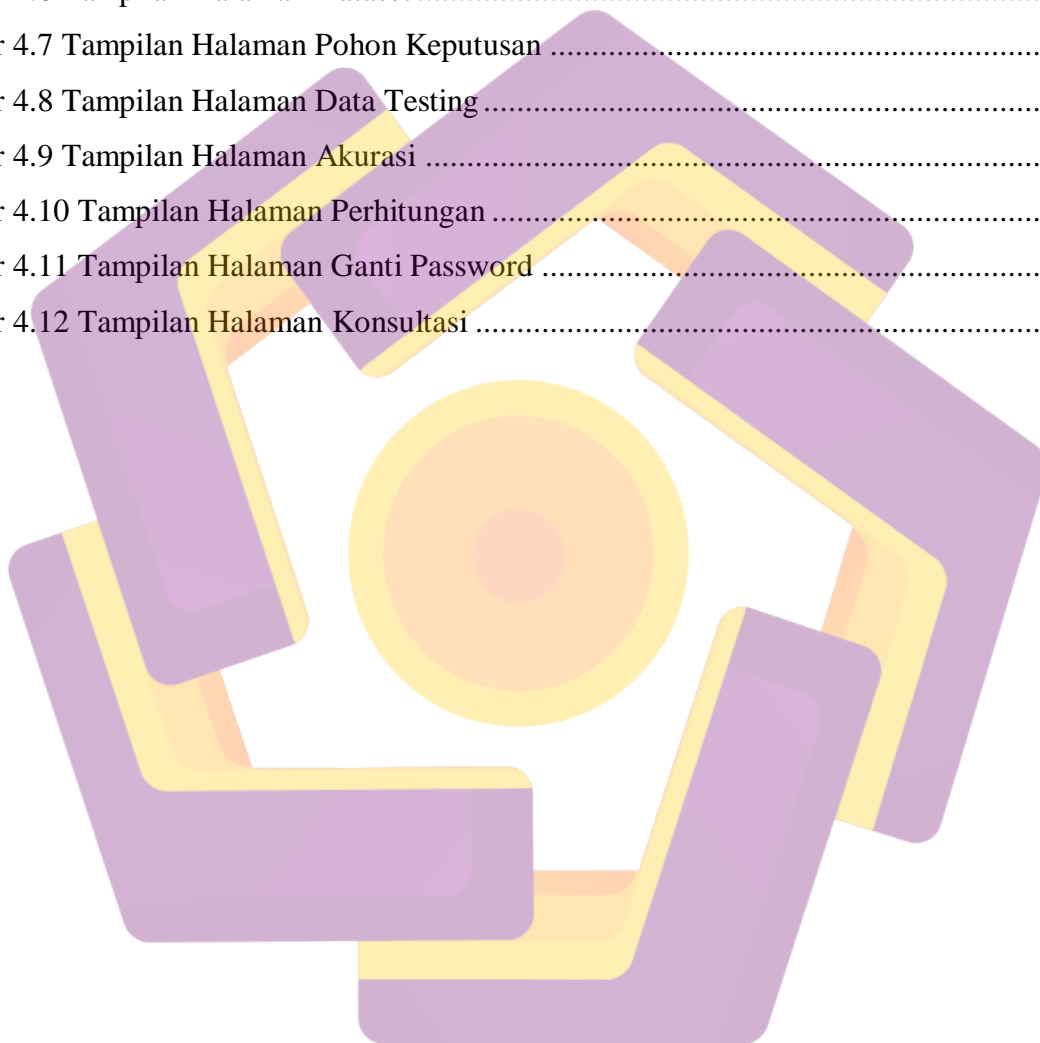
DAFTAR TABEL

Tabel 2.1 Kajian Pustaka.....	24
Tabel 2.1 Keputusan Bermain Tenis.....	30
Tabel 2.1 Perhitungan Node 1	31
Tabel 2.4 Perhitungan Node 1.1	33
Tabel 2.5 Perhitungan Node 1.1.2	36
Tabel 2.6 <i>confusion matrix</i>	37
Tabel 2.7 Simbol-Simbol Use Case Diagram.....	39
Tabel 2.8 Simbol-Simbol Activity Diagram	41
Tabel 3.1 Data Atribut Iris	47
Tabel 3.2 Atribut Awal Data Iris	50
Tabel 3.3 Atribut Data Iris Setelah Transformasi.....	50
Tabel 3.4 Data Sampel Untuk Nilai Mean	51
Tabel 3.5 Data Sampel Diskritisasi.....	53
Tabel 3.6 Deskripsi Use Case Diagram	55
Tabel 3.7 Rancangan Tabel Admins	69
Tabel 3.8 Rancangan Tabel Atribut.....	69
Tabel 3.9 Rancangan Tabel Nilai	70
Tabel 3.10 Rancangan Tabel Dataset.....	70
Tabel 3.11 Rancangan Tabel Testing.....	70
Tabel 4.1 Hasil Akurasi.....	90

DAFTAR GAMBAR

Gambar 2.1 Tahapan Data Mining	27
Gambar 2.2 Beberapa Metode Data Mining	28
Gambar 2.3 Pohon Keputusan Hasil Perhitungan Node 1	33
Gambar 2.4 Pohon Keputusan Hasil Perhitungan Node 1.1	34
Gambar 2.5 Pohon Keputusan Hasil Perhitungan Node 1.1.2	36
Gambar 2.6 Tiga Jenis Bunga Iris	42
Gambar 2.7 Keterangan Bunga Iris	43
Gambar 3.1 Diagram Alur Proses Penelitian	46
Gambar 3.2 Diagram Use Case Diagram	55
Gambar 3.3 Activity Diagram Login Admin / Peneliti	58
Gambar 3.4 Activity Diagram Mengelola Atribut	59
Gambar 3.5 Activity Diagram Mengelola Nilai Atribut	60
Gambar 3.6 Activity Diagram Mengelola Dataset	61
Gambar 3.7 Activity Diagram pohon keputusan	62
Gambar 3.8 Activity Diagram Data Testing	63
Gambar 3.9 Activity Diagram Akurasi	64
Gambar 3.10 Activity Diagram Hasil	65
Gambar 3.11 Activity Diagram Ganti Password	66
Gambar 3.12 Activity Diagram Logout	67
Gambar 3.13 Activity Diagram Konsultasi	68
Gambar 3.14 Rancangan Tampilan Halaman Awal	71
Gambar 3.15 Rancangan Tampilan Halaman Login	72
Gambar 3.16 Rancangan Tampilan Halaman Atribut	72
Gambar 3.17 Rancangan Tampilan Halaman Nilai Atribut	73
Gambar 3.18 Rancangan Tampilan Halaman Dataset	74
Gambar 3.19 Rancangan Tampilan Halaman Pohon Keputusan	74
Gambar 3.20 Rancangan Tampilan Halaman Data Testing	75
Gambar 3.21 Rancangan Tampilan Halaman akurasi	75
Gambar 3.22 Rancangan Tampilan Halaman Perhitungan	76
Gambar 3.23 Rancangan Tampilan Halaman Ganti Password	77

Gambar 3.24 Rancangan Tampilan Halaman Konsultasi	78
Gambar 4.1 Hasil Pembuatan Basis Data	81
Gambar 4.2 Tampilan Halaman Awal	82
Gambar 4.3 Tampilan Halaman Login	83
Gambar 4.4 Tampilan Halaman Atribut.....	83
Gambar 4.5 Tampilan Halaman Nilai Atribut.....	84
Gambar 4.6 Tampilan Halaman Dataset	84
Gambar 4.7 Tampilan Halaman Pohon Keputusan	85
Gambar 4.8 Tampilan Halaman Data Testing	85
Gambar 4.9 Tampilan Halaman Akurasi	86
Gambar 4.10 Tampilan Halaman Perhitungan	86
Gambar 4.11 Tampilan Halaman Ganti Password	87
Gambar 4.12 Tampilan Halaman Konsultasi	87



DAFTAR ISTILAH



KDD	: <i>Knowledge Discovery from Database</i>
k-NN	: <i>k-Nearest Neighbors Algorithm</i>
TP	: <i>True Positive</i>
FP	: <i>False Positive</i>
FN	: <i>False Negative</i>
TN	: <i>True Negative</i>
HTML	: <i>Hyper Text Markup Language</i>
CSS	: <i>Cascading Style Sheet</i>
PHP	: <i>Hypertext Preprocessor</i>
DBMS	: <i>Relational Database Management System</i>
UML	: <i>Unified Modeling Language</i>
PC	: <i>Personal Computer</i>
AMD	: <i>Advanced Micro Devices</i>
VGA	: <i>Radeon Vega Graphics</i>
GHz	: <i>Gigahertz</i>
SSD	: <i>Solid State Drive</i>
GB	: <i>Gigabyte</i>
XAMPP	: X (tempat sistem operasi apapun), Apache, MySQL, PHP dan Perl

INTISARI

Dalam mengklasifikasi data, terdapat beberapa algoritma yang bisa digunakan untuk klasifikasi, salah satunya yaitu algoritma C4.5, algoritma ini memiliki tingkat akurasi dalam mengklasifikasi relatif tinggi, Klasifikasi merupakan suatu proses dalam mencari suatu himpunan model (fungsi) untuk dapat mendeskripsikan dan membedakan kelas-kelas data atau konsep-konsep, yang memiliki tujuan agar dapat menggunakan model tersebut sehingga bisa memprediksi kelas dari objek yang belum diketahui kelasnya.

Metode yang digunakan dalam pengumpulan data dalam penelitian ini menggunakan metode kualitatif dan dilaksanakan dengan jenis kualitatif Study Pustaka (*Library Research Method*) yaitu metode untuk mendapatkan data dibutuhkan dan mencari informasi mengenai pengolahan data baik dari buku-buku, laporan penelitian maupun jurnal dari internet.

Hasil dari penelitian ini adalah aplikasi Data Mining ini dengan menggunakan algoritma klasifikasi C4.5 dapat memproses data menjadi aturan – aturan yang digunakan sebagai penentu keputusan sehingga aplikasi dapat memprediksi spesies bunga iris dan mendapatkan presentase keakuratan sebesar 72% serta waktu prosesnya 0.18638706207275 detik dan Pengolahan data pada penelitian menggunakan teknik diskritisasi yang mengkategorikan nilai dari atribut menjadi 0 yaitu nilai atribut yang berada dibawah rata-rata sedangkan 1 yaitu nilai yang berada diatas rata-rata, hal tersebut ternyata mempengaruhi nilai akurasi karena nilai dari atribut kurang beragam sehingga nilai akurasi pada algoritma menjadi lebih rendah dibandingkan dengan data awal yaitu tipe data diskrit.

Kata Kunci : Penambangan Data, Klasifikasi, Algoritma C4.5 dan Bunga Iris.

ABSTRACT

In classifying data, there are several algorithms that can be used for classification, one of which is the C4.5 algorithm, this algorithm has a relatively high level of accuracy in classifying, Classification is a process in finding a set of models (functions) to describe and differentiate classes. data class or concepts, which have the aim of being able to use the model so that it can predict the class of objects whose class is not yet known.

The method used in data collection in this study uses qualitative methods and is carried out with a qualitative type of Library Research Method, which is a method to obtain the required data and seek information about data processing from books, research reports and journals from the internet.

The result of this research is that this Data Mining application using the C4.5 classification algorithm can process data into rules that are used as decision makers so that the application can predict the iris species and get an accuracy percentage of 72% and the processing time is 0.18638706207275 seconds and data processing. In research using discretization techniques that categorize the value of the attribute into 0, namely the attribute value that is below the average while 1 is the value that is above the average, this turns out to affect the accuracy value because the value of the attribute is less diverse so that the accuracy value in the algorithm becomes lower than the initial data, namely the discrete data type

Keyword: Data Mining, Classification, C4.5 Algorithm and Iris.