

**PENERAPAN ALGORITMA *DECISION TREE* UNTUK PENENTUAN
KONSENTRASI MAHASISWA PROGRAM STUDI MANAJEMEN
UNISMUH MAKASSAR**

SKRIPSI

Diajukan Sebagai Salah Satu Syarat Untuk Mendapatkan Gelar
Sarjana Komputer (S.Kom) Program Studi Informatika



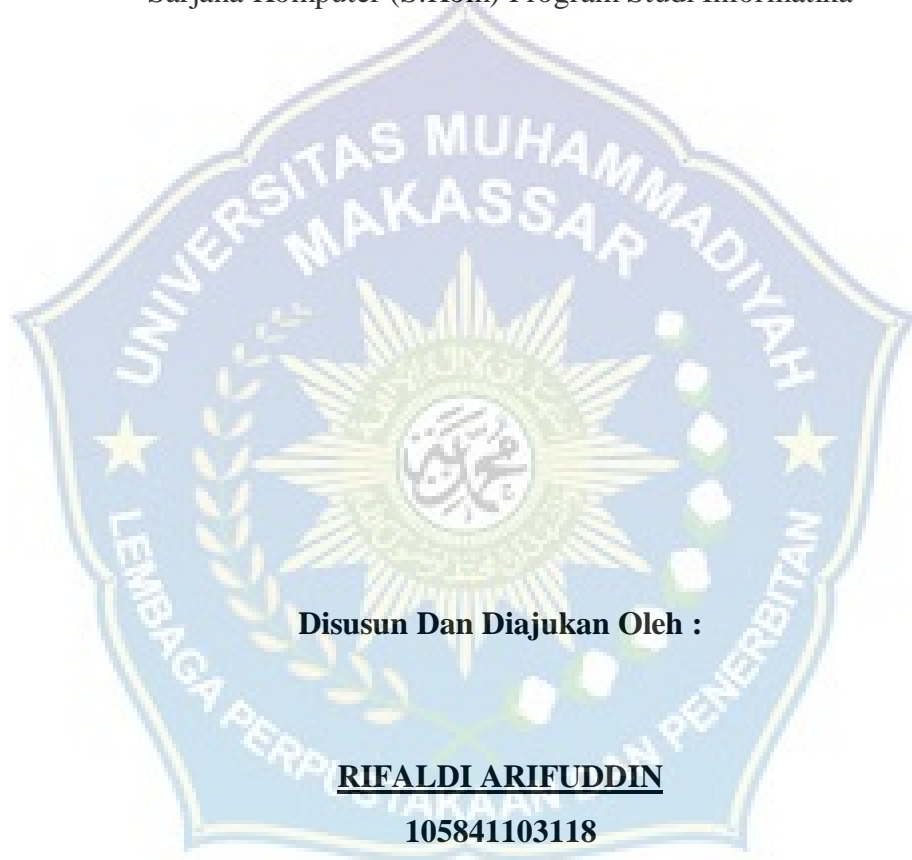
RIFALDI ARIFUDDIN

105841103118

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAKASSAR**

**PENERAPAN ALGORITMA *DECISION TREE* UNTUK PENENTUAN
KONSENTRASI MAHASISWA PROGRAM STUDI MANAJEMEN
UNISMUH MAKASSAR**

Diajukan Sebagai Salah Satu Syarat Untuk Mendapatkan Gelar
Sarjana Komputer (S.Kom) Program Studi Informatika



**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MAKASSAR**



UNIVERSITAS MUHAMMADIYAH MAKASSAR

FAKULTAS TEKNIK

GEDUNG MENARA IQRA LT. 3

Jl. Sultan Alauddin No. 259 Telp. (0411) 866 972 Fax (0411) 865 588 Makassar 90221

Website: www.unismuh.ac.id, e_mail: unismuh@gmail.com

Website: <http://teknik.unismuh.makassar.ac.id>

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

PENGESAHAN

Skripsi atas nama **Rifaldi Arifuddin** dengan nomor induk Mahasiswa **105 84 11031 18**, dinyatakan diterima dan disahkan oleh Panitia Ujian Tugas Akhir/Skripsi sesuai dengan Surat Keputusan Dekan Fakultas Teknik Universitas Muhammadiyah Makassar Nomor : 001/05/A.5-II/1/45/2024, sebagai salah satu syarat guna memperoleh gelar Sarjana Komputer pada Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar pada hari Sabtu tanggal 31 Januari 2024.

Panitia Ujian :

Makassar, 19 Rajab 1445 H
3 Januari 2024 M

1. Pengawas Umum

a. Rektor Universitas Muhammadiyah Makassar

Prof. Dr. H. AMBO ASSE, M.Ag

b. Dekan Fakultas Teknik Universitas Hasanuddin

Prof. Dr. Eng. MUHAMMAD ISRAN RAMLI, ST., MT

2. Penguji

a. Ketua : Dr. Ir. Zahir Zuhuddin, M.Sc

b. Sekortaris : Asyraful Insan Asry, S.Kom., MT

3. Anggota

1. Rizki Kusliana Bakti, ST., MT.

2. Muhyiddin A M Hayat, S.Kom., M.T.

3. Lukman Anas, S.Kom., M.T.

Mengetahui :

Pembimbing I

Titin Wahyuni S.Pd., MT.

Pembimbing II

Fahrin Irfhamna Rachman, S.Kom., MT



Dekan Fakultas Teknik

Dr. Hj. Hurnawaty, ST., MT., IPM

NBM : 795 108



UNIVERSITAS MUHAMMADIYAH MAKASSAR
FAKULTAS TEKNIK

GEDUNG MENARA IQRA LT. 3

Jl. Sultan Alauddin No. 259 Telp. (0411) 866 972 Fax (0411) 865 588 Makassar 90221

Website: www.unismuh.ac.id, e_mail: unismuh@gmail.com

Website: <http://teknik.unismuh.makassar.ac.id>

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

HALAMAN PENGESAHAN

Tugas Akhir ini diajukan untuk memenuhi syarat ujian guna memperoleh gelar Sarjana Komputer (S.Kom) Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar.

Judul Skripsi : **PENERAPAN ALGORITMA DECISION TREE UNTUK PENENTUAN KONSENTRASI MAHASISWA PROGRAM STUDI MANAJEMEN UNISMUH MAKASSAR.**

Nama : RIFALDI ARIFUDDIN

Stambuk : 105841103118

Makassar, 31 Januari 2024

Telah Diperiksa dan Disetujui
Oleh Dosen Pembimbing;

Pembimbing I

Pembimbing II


Titin Wahyuni S.Pd., MT.


Fahrir Irhamna Rachman, S.Kom., MT



Mengetahui,

Ketua Prodi Informatika


Muhyidin A. M. Hayat, S.kom., MT

NBM : -

Abstrak

Pemilihan konsentrasi adalah tindakan penting bagi mahasiswa semester enam keatas pada program studi manajemen Universitas Muhammadiyah Makassar, Mata kuliah yang ditawarkan oleh prodi manajemen menjurus kebeberapa bidang yaitu Sdm, Pemasaran, Dan Keuangan. Pemilihan yang tepat dapat membantu mahasiswa dalam menentukan mata kuliah yang sesuai minat dan nilai akademiknya. Proses pengklasifikasian yang dilakukan sangat sederhana yaitu dengan menjumlahkan nilai yang telah ditentukan dan selanjutnya hasil dari penjumlahan tersebut dilakukan pelabelan konsentrasi, perhitungan yang dilakukan sangat sederhana sehingga dalam pembagiannya rentan terjadi kekeliruan. Untuk membantu proses pembagian konsentrasi penulis menggunakan proses Data Mining yaitu metode klasifikasi dengan Algoritma Classification and Regression Tree (CART) pada penelitian ini menggunakan data siswa angkatan 2018, 2019, 2020 yang berjumlah 20 data set nilai akurasi dari Algoritma CART adalah sebesar 100% berdasarkan hasil yang diperoleh Algoritma CART ini bisa digunakan untuk proses pengklasifikasian konsentrasi berdasarkan minat dan nilai akademik mahasiswa.

Kata kunci: *Konsentrasi, Manajemen, Klasifikasi, Algoritma CART*

Abstract

Choosing a concentration is an important action for students in the sixth semester and above in the management study program at Muhammadiyah University of Makassar. The courses offered by the management study program focus on several fields, namely Human Resources, Marketing and Finance. The right choice can help students determine courses that suit their interests and academic values. The classification process carried out is very simple, namely by adding up the values that have been determined and then the results of the addition are labeled with concentration. The calculations carried out are very simple so that in the division it is prone to errors. To help with the process of dividing concentration, the author uses a Data Mining process, namely a classification method with the Classification and Regression Tree (CART) Algorithm. In this study, data from students from the 2018, 2019, 2020 classes, totaling 20 data sets, the accuracy value of the CART Algorithm is 100% based on the results. The results obtained by the CART algorithm can be used for the process of classifying concentrations based on students' interests and academic values.

Keyword: *Concentration, Management, Classification, CART Algorithm*

KATA PENGANTAR

Assalamu'alaikum Warahmatullahi Wabaraktuh

Alhamdulillah, Segala puji kehadiran Allah Subhanahu Wa Ta'ala karena atas berkat rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan laporan skripsi ini tepat pada waktunya sebagai salah satu syarat untuk wisuda. Shalawat beserta salam semoga senantiasa tercurah kepada junjungan kita Baginda Rasulullah Shalallahu „Alaihi Wassalam beserta para keluarga, sahabat, dan para pengikut Beliau hingga akhir zaman.

Dalam pembuatan laporan skripsi ini, penulis banyak dibantu oleh berbagai pihak dengan memberikan banyak masukan dan petunjuk, serta mendukung dan menjadi motivasi bagi penulis. Maka dari itu, ucapan terima kasih penulis haturkan kepada :

1. Kepada **Kedua Orang Tua Penulis Dan Keluarga** yang telah memberikan wejangan dan motivasi baik secara moril maupun materil
2. Ibu **Dr. Ir. Hj Nurnawati, S.T., M.T., I.P.M**, selaku Dekan Fakultas Teknik.
3. Bapak **Muh. Syafaat S Kuba, S.T., M.T**, selaku Wakil Dekan Fakultas Teknik.
4. Bapak **Muhyiddin AM Hayat S.Kom., M.T**, selaku Ketua Prodi Informatika.
5. Ibu **Titin Wahyuni S.Pd., M.T**, selaku Dosen Pembimbing 1 pada skripsi ini.

6. Bapak **Fahrim Irhamna Rachman S.Kom., M.T**, selaku Dosen Pembimbing 2 pada skripsi ini.
7. Dosen dan Staf Fakultas Teknik Universitas Muhammadiyah Makassar.
8. Teman-teman Khususnya Angkatan 2018 Fakultas Teknik, Universitas Muhammadiyah Makassar, terima kasih atas dukungan dan doanya.
9. Teman-teman kelas B angkatan 2018 Program Studi Informatika Universitas Muhammadiyah Makassar
10. Teman-teman Pengurus Kelembagaan Fakultas Teknik, Universitas Muhammadiyah Makassar, terima kasih atas dukungan dan doanya.

Akhir kata, penulis mengharapkan agar skripsi ini dapat bermanfaat, baik bagi penulis pribadi maupun pada pihak-pihak lain. Serta, semoga segala masukan baik berupa kritik maupun saran yang membangun yang ditujukan kepada penulis dapat menjadikan penulis menjadi lebih baik lagi untuk kedepan. Terima kasih.

Billahi fisabililhaq, fastabiqul khairat.

Wassalamualaikum Wr.Wb.

Makassar, 31 Januari 2024

Penulis

Rifaldi Arifuddin

DAFTAR ISI

PENERAPAN ALGORITMA DECISION TREE UNTUK PENENTUAN KONSENTRASI MAHASISWA PROGRAM STUDI MANAJEMEN UNISMUH MAKASSAR.....	i
LEMBAR PENGESAHAN.....	ii
LEMBAR PENGESAHAN.....	iii
ABSTRAK	iv
KATA PENGANTAR	v
DAFTAR ISI	vii
DAFTAR GAMBAR.....	ix
DAFTAR TABEL	x
DAFTAR LAMPIRAN.....	xi
DAFTAR ISTILAH.....	xii
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah.....	1
B. Rumusan Masalah.....	4
C. Tujuan Penelitian	4
D. Manfaat Penelitian.....	4
E. Ruang Lingkup Penelitian	6
F. Sistematika Penelitian	6
BAB II TINJAUAN PUSTAKA	8
A. Landasan Teori.....	8
1. Konsentrasi.....	8
2. Program Studi Manajemen Unismuh Makassar.....	9
3. Mahasiswa.....	10
4. Data Mining	11
5. Decesion Tree.....	14
6. Classification and Regression Tree (CART)	15
7. Pengujian White Box.....	18
8. Daftar Symbol.....	20
9. Visual Studio Code.....	24

B. Penelitian Terkait	25
C. Kerangka Pikir	28
BAB III METODE PENELITIAN	29
A. Jenis Penelitian.....	29
B. Data dan Sumber Data.....	29
C. Perancangan Sistem.....	30
D. Teknik Pengujian Sistem	31
E. Teknik Analisis Data	32
1. Reduksi Data	33
2. Display Data.....	34
3. Penarikan Kesimpulan	34
BAB IV HASIL DAN PEMBAHASAN	35
A. Deskripsi Dataset	35
B. Analisis Data Mentah	35
C. Data Preprocessing	37
D. Pelabelan Data Mahasiswa	40
1. Identifikasi Mata Kuliah Konsentrasi.....	40
2. Identifikasi Pelabelan Konsentrasi	42
E. Perhitungan Algoritma CART	46
F. Proses Machine Learning Perhitungan Algoritma CART	52
G. Proses Machine Learning	57
H. Teknik Pengujian Sistem	63
BAB V PENUTUP	65
A. KESIMPULAN	65
B. SARAN.....	65
DAFTAR PUSTAKA	66
LAMPIRAN	67

DAFTAR GAMBAR

Gambar 1. Knowlegde Discoveryin Database.....	13
Gambar 2. Decision Tree	14
Gambar 3. Diagram Kerangka Pikir	28
Gambar 4. Diagram Alir Algoritma CART	32
Gambar 5. Diagram Alir Algoritma CART Lanjutan.....	33
Gambar 6. Pohon Keputusan CART.....	50
Gambar 7. Output Source Code.....	68
Gambar 8. Surat Pengambilan Data.....	69
Gambar 9. Surat Pengambilan Data.....	70
Gambar 10. Surat Penelitian.....	71



DAFTAR TABEL

Tabel 1. Simbol Data Flow Diagram (DFD).....	21
Tabel 2. Simbol Flow Map.....	22
Tabel 3. Simbol Entity Relationship Diagram	23
Tabel 4. Dataset Mahasiswa Manajemen 2018	35
Tabel 5. Dataset Mahasiswa Manajemen 2019.....	36
Tabel 6. Dataset Mahasiswa Manajemen 2020	37
Tabel 7. Dataset Nilai Mahasiswa Manajemen 2018	38
Tabel 8. Dataset Nilai Mahasiswa Manajemen 2019	38
Tabel 9. Dataset Nilai Mahasiswa Manajemen 2020	39
Tabel 10. Dataset Mata Kuliah Konsentrasi SDM	40
Tabel 11. Dataset Mata Kuliah Konsentrasi Pemasaran	41
Tabel 12. Dataset Mata Kuliah Konsentrasi Keuangan	41
Tabel 13. Dataset Pelabelan Konsentrasi Mahasiswa.....	42
Tabel 14. Dataset Penjumlahan Untuk Konsentrasi SDM	43
Tabel 15. Dataset Penjumlahan Untuk Konsentrasi Pemasaran.....	44
Tabel 16. Dataset Penjumlahan Untuk Konsentrasi Keuangan.....	45
Tabel 17. Calon Cabang Mutakhir.....	46
Tabel 18. Jumlah Calon Cabang	48
Tabel 19. Nilai Kandidate Split	48
Tabel 20. Perhitungan Nilai Q (S t)	49
Tabel 21. Hasil Pengujian Secara Manual dan Decision Tree	64

DAFTAR LAMPIRAN

Lampiran 1. Source Code.....	33
Lampiran 2. Outpur Source Code.....	50
Lampiran 3. Surat Pengambilan Data	33
Lampiran 4. Surat Penelitian	50



DAFTAR ISTILAH

SK	Surat Keputusan
DIKTI	Direktorat Jenderal Pendidikan Tinggi (Ditjen Dikti) merupakan salah satu unit utama di Kementerian Pendidikan dan Kebudayaan.
CART	Classification and Regression Tree merupakan salah satu metode atau algoritma dari salah satu teknik pohon keputusan.
Descriptive	Yaitu proses untuk menemukan karakteristik penting dari data dalam satu basis data
Predictive	Proses untuk menemukan pola dari data dengan menggunakan beberapa variable ain di masa depan
Data Selection	Pemilihan data
Pre-processing	Proses pembersihan melibatkan penghapusan data duplikat, memeriksa data yang bertentangan dan memperbaiki kesalahan dalam data seperti kesalahan ketik
Transformation	Proses integrasi data yang dipilih untuk membuat data cocok untuk proses penambangan data
White-box	Merupakan sebuah teknik pengujian perangkat lunak yang melibatkan pemeriksaan dan pengujian struktur internal perangkat lunak.
Impurity	Adalah ukuran homogenitas data. Data dikatakan murni atau homogen jika hanya memuat satu kelas saja
Gini Impurity	Adalah kemungkinan salah mengklasifikasikan elemen yang dipilih secara acak dalam kumpulan data jika elemen tersebut diberi label acak sesuai dengan distribusi kelas dalam kumpulan data
Root Node	Node paling atas dari pohon keputusan. Itu tidak memiliki simpul induk. Ini mewakili seluruh populasi atau sampel
Internal Node	Adalah Simpul akar mewakili fitur pada dataset

Decision Node

Untuk membuat keputusan berdasarkan fitur dari dataset yang diberikan. Node ini memiliki simpul cabang, bisa berupa decision node atau leaf node.

Leaf Node

Digunakan untuk mewakili keluaran hasil keputusan dari simpul induknya

Indeks Gini

Sebagai fungsi biaya yang digunakan untuk mengevaluasi pemisahan dalam kumpulan data



BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Universitas Muhammadiyah Makassar merupakan salah satu perguruan tinggi Muhammadiyah yang merupakan amal usaha Muhammadiyah dalam pengembangan pendidikan, khususnya pada jenjang pendidikan tinggi. Sejak berdiri hingga saat ini telah memiliki 2 program doktor, 8 program pascasarjana, 8 fakultas, 50 program studi tingkat satu dan 3 diploma. Salah satunya adalah program studi manajemen. Manajemen didirikan dengan surat keputusan pendirian program studi nomor : 028/DIKTI/1983 tanggal 11 september 1983 yang ditanda tangani oleh Menteri Pendidikan dan Kebudayaan Prof. Dr. Nugroho Notosusanto. Penyelenggaraan program studi manajemen dimulai pada bulan juli tahun 1991 dengan SK izin operasional nomor : 1712/D/T/K-IX/2009 tanggal 31 maret 2009. Tujuan utama program studi manajemen adalah mewujudkan program studi yang memiliki daya saing tinggi dan dipercaya untuk mengembangkan ilmu pendidikan dalam bidang Manajemen yang berkarakter Islami di kawasan timur Indonesia melalui penguatan kegiatan pendidikan dan pengajaran, penelitian, pengabdian pada masyarakat serta kemuhammadiyah.

Pemilihan konsentrasi pilihan adalah tindakan penting bagi mahasiswa semester enam keatas pada Program Studi Manajemen Universitas Muhammadiyah Makassar, Mata kuliah yang ditawarkan oleh Prodi Manajemen menjurus kebeberapa bidang, pemilihan yang tepat dapat membantu mahasiswa dalam menentukan mata kuliah yang sesuai minat dan nilai akademiknya. Konsentrasi

mata kuliah ini digunakan mahasiswa dalam memilih mata kuliah yang diminati, menjadi bahan untuk mengambil tema skripsi dan menentukan profesi untuk dunia pekerjaan ketika sudah lulus nanti. Banyak mahasiswa yang memilih mata kuliah konsentrasi ini dengan pertimbangan yang sangat subjektif seperti hanya karena teman sekelas memilih bidang keahlian tersebut. Banyak mahasiswa yang masih merasa kebingungan dengan pemilihan mata kuliah itu sendiri karena semester sebelumnya belum ada mata kuliah pilihan tersebut. konsultasi dengan dosen pembimbing akademik yang kurang optimal, Adapun faktor lain karena kurangnya sosialisasi mengenai bidang keahlian yang ditawarkan oleh fakultas, serta kurang inisiatif untuk bergaul dan bertanya kepada kakak tingkat yang secara waktu telah lebih dulu melewatinya.

Dari permasalahan yang telah dijelaskan ada beberapa kondisi yang harus diselesaikan dengan pengambilan keputusan dalam kepastian, memperkecil tingkat resiko dari pilihan mata kuliah yang ditawarkan prodi, serta mengurangi pengambilan keputusan dalam kondisi ketidakpastian. Pengambilan Keputusan merupakan kegiatan memilih suatu strategi atau tindakan dalam pemecahan masalah tersebut. Tindakan memilih strategi atau aksi yang diyakini akan memberikan solusi terbaik. Namun kurangnya informasi tentang alternatif pilihan membuat proses pengambilan keputusan yang tepat sulit dilakukan.

Sistem Penentuan Keputusan merupakan Sistem informasi interaktif yang menyediakan informasi, pemodelan, dan pemanipulasian data. Dengan Sistem ini akan membantu mahasiswa dalam menentukan konsentrasi mata kuliah pilihan yang sesuai dengan bakat dan nilai akademik mereka ataupun mahasiswa yang masih merasa bingung dengan pemilihan mata kuliah pilihan yang ditawarkan oleh program studi manajemen Universitas Muhammadiyah Makassar.

Berdasarkan hal tersebut, Maka dalam Penelitian ini berbeda dari penelitian pendukung sebelumnya dalam hal manfaat dari pengelompokan jenis konsentrasi. Penelitian ini tidak hanya berfungsi memberikan penggambaran tentang Pemilihan konsentrasi terhadap mahasiswa namun juga memberikan penggambaran penentuan kebijakan bagi kampus. Sehingga dengan adanya pengelompokan data ini pihak kampus juga dapat mengetahui jenis konsentrasi yang banyak di minati oleh mahasiswa Program Studi Manajemen Universitas Muhammadiyah Makassar.

Berdasarkan latar belakang tersebut maka penulis tertarik untuk melakukan penelitian dengan judul “ PENERAPAN ALGORITMA DECISION TREE UNTUK PENENTUAN KONSENTRASI MAHASISWA PROGRAM STUDI MANAJEMEN UNISMUH MAKASSAR “

B. Rumusan Masalah

Berdasarkan latar belakang penelitian di atas, peneliti dapat merumuskan masalah sebagai berikut:

1. Bagaimana membangun suatu aplikasi sistem penentuan pemilihan konsentrasi mahasiswa Program Studi Manajemen ?
2. Bagaimana mengimplementasikan Algoritma *Decision Tree* pada sistem penentuan pemilihan konsentrasi mahasiswa Program Studi Manajemen ?

C. Tujuan Penelitian

Berdasarkan latar belakang dan rumusan masalah penelitian di atas, peneliti dapat memberikan tujuan penelitian sebagai berikut:

1. Untuk menghasilkan sebuah sistem penentuan pemilihan konsentrasi mahasiswa Program Studi Manajemen
2. Untuk mengetahui penggunaan Algoritma *Decision Tree* dalam penentuan konsentrasi mata kuliah mahasiswa Program Studi Manajemen

D. Manfaat Penelitian

Adapun manfaat penelitian yang akan didapatkan dari penelitian adalah :

1. Manfaat bagi Peneliti
 - a. Penelitian ini dapat membantu peneliti untuk meningkatkan pemahaman dan pengetahuan dalam menggunakan algoritma *Decision Tree* dalam penentuan konsentrasi mahasiswa Program Studi Manajemen
 - b. Penelitian ini dapat membantu memecahkan masalah yang berkaitan dengan masalah penentuan konsentrasi mahasiswa. Dalam penelitian ini, peneliti

dapat mengidentifikasi faktor-faktor apa saja yang berkontribusi pada masalah penentuan konsentrasi mahasiswa.

2. Manfaat Bagi Dosen

- a. Manfaat penelitian bagi dosen yaitu dapat memberikan system pendukung keputusan untuk mahasiswa yang mengalami kendala dalam pemilihan konsentrasi
- b. Dosen dapat mengetahui pilihan konsentrasi yang paling banyak diminati oleh mahasiswa
- c. Dosen dapat menggunakan hasil penelitian untuk memperbaiki proses pengawasan akademik, seperti mengidentifikasi mahasiswa yang berpotensi gagal dan memberikan bantuan untuk meningkatkan performa akademik mereka.

3. Manfaat Bagi Mahasiswa

- a. Penelitian ini dapat bermanfaat bagi mahasiswa yaitu dapat membantu mahasiswa dalam menentukan konsentrasi yang akan dipilih.
- b. Akan membantu mahasiswa dalam meningkatkan kualitas dan kompetensi mahasiswa dalam konsentrasi yang dipilih sehingga lebih ahli dan terampil.

E. Ruang Lingkup Penelitian

Ruang lingkup penelitian adalah batasan atau ruang lingkup yang dijadikan fokus penelitian. Dengan bantuan daerah penelitian, wilayah atau bidang yang akan dipelajari ditentukan dan batas-batas objek atau fenomena yang diamati ditentukan.

1. Penelitian ini membutuhkan data mahasiswa yang sementara berjalan pada semester satu sampai lima yang akan melakukan perencanaan pemilihan konsentrasi di semester enam, Data yang dimasukkan terdiri atas :

- a. Nama Mahasiswa
- b. Data Nilai Mahasiswa semester satu hingga semester lima
- c. Data Matakuliah semester satu hingga semester lima
- d. Data mahasiswa yang diambil sebanyak 3 angkatan yaitu angkatan 2018,2019,2020.

2. Dalam penelitian ini, akan digunakan teknik *data mining* untuk mengolah data yang telah dikumpulkan. Algoritma yang digunakan adalah Algoritma *Decision Tree*

F. Sistematika Penulisan

Sistematika penulisan dalam penyusunan laporan proposal penelitian ini adalah sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menguraikan tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, ruang lingkup penelitian dan sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Pada bab ini menguraikan mengenai studi pustaka terhadap teori-teori yang digunakan dalam penelitian serta aspek-aspek yang mendukung dalam penelitian.

BAB III METODOLOGI PENELITIAN

Bab ini berisi penjelasan tentang langkah-langkah yang digunakan dalam penelitian, dimulai dari pengidentifikasian masalah, pengumpulan data, analisa algoritma, dan alat ukur yang digunakan dalam penelitian.

BAB IV HASIL DAN PEMBAHASAN

Bab ini berisi penjelasan pembahasan mencakup rangkuman temuan utama dari penelitian dan analisis yang telah dilakukan. Ini memberikan gambaran komprehensif tentang apa yang telah dicapai dalam penelitian tersebut.

BAB V PENUTUP

Bab ini berisi penjelasan Penelitian ini menghasilkan temuan yang memberikan kontribusi pada pemahaman di bidang ini. Meskipun ada batasan, hasil penelitian ini memiliki implikasi praktis yang dapat diterapkan dalam konteks tertentu. Sebagai saran untuk penelitian selanjutnya, pertimbangkan eksplorasi lebih lanjut pada aspek tertentu dan perbaikan pada metode penelitian

BAB II

TINJAUAN PUSTAKA

A. Landasan Teori

1. Konsentrasi

Dalam konteks ini, konsentrasi didefinisikan sebagai suatu kemampuan yang melibatkan upaya aktif untuk memusatkan pikiran, perasaan, kemauan, dan seluruh panca indera pada satu objek dalam suatu aktivitas tertentu. Dengan jelas, konsentrasi bukan sekadar menunjukkan fokus perhatian, melainkan juga melibatkan usaha sadar untuk mengabaikan segala objek atau gangguan yang tidak terkait dengan aktivitas yang sedang dilakukan. Dengan demikian, esensi dari konsentrasi adalah kemampuan untuk mengisolasi dan memfokuskan seluruh daya kognitif dan emosional pada suatu tujuan atau tugas, yang pada gilirannya meningkatkan efisiensi dan efektivitas pelaksanaan aktivitas tersebut, Parnawi (2019),.

Konsentrasi adalah memusatkan fungsi jiwa (perhatian, pikiran, serta perbuatan) pada suatu objek yang sedang dipelajari dengan mengabaikan hal yang tidak berkaitan dengan objek yang sedang dipelajari, (Ratih, dkk, 2019:4).

konsentrasi belajar adalah fenomena di mana pikiran seseorang sepenuhnya terfokus pada proses pembelajaran secara langsung, tanpa adanya kegiatan lain yang mengganggu. Pada dasarnya, konsentrasi belajar melibatkan pemusatan pikiran dan perhatian siswa secara eksklusif pada isi bahan pelajaran atau proses pembelajaran yang tengah berlangsung. Signifikansinya terletak pada kemampuan

siswa untuk menikmati dan sepenuhnya terlibat dalam proses belajar. Dalam konteks ini, kurangnya konsentrasi belajar dapat disebabkan oleh ketidakmampuan siswa untuk menikmati atau merasakan kepuasan dalam menjalani proses pembelajaran. Oleh karena itu, mencapai konsentrasi belajar yang optimal menjadi kunci untuk meningkatkan efektivitas dan pengalaman belajar siswa, Sici, dkk. (2021:21),.

2. Program Studi Manajemen Unismuh Makassar

Program Studi Manajemen Unismuh Makassar didirikan dengan Surat Keputusan Pendirian Program Studi Nomor 028/DIKTI/1983 Tanggal 11 September 1983 yang ditanda tangani oleh Menteri Pendidikan dan Kebudayaan Prof. Dr. Nugroho Notosusanto. Penyelenggaraan Program Studi Manajemen dimulai pada Bulan Juli Tahun 1991 dengan SK Izin Operasional Nomor : 1712/D/T/K-IX/2009 Tanggal 31 Maret 2009.

Adapun visi keilmuan Program Studi Manajemen Unismuh Makassar Mewujudkan program studi yang memiliki daya saing tinggi dan dipercaya untuk mengembangkan ilmu pendidikan dalam bidang Manajemen yang berkarakter Islami di kawasan timur Indonesia melalui penguatan kegiatan pendidikan dan pengajaran, penelitian, pengabdian pada masyarakat serta kemuhammadiyah.

Perumusan Visi, Misi, Tujuan, dan Sasaran Program Studi Manajemen dilakukan berdasarkan pertimbangan kebutuhan dunia kerja sehingga Program Studi Manajemen sebagai institusi diharapkan dapat berkontribusi dalam mengisi kebutuhan tenaga kerja dan bahkan dapat menciptakan lapangan kerja. Adapun rumusan Visi, Misi, Tujuan dan Sasaran Program Studi Manajemen Fakultas

Ekonomi Dan Bisnis Universitas Muhammadiyah Makassar. Pada Program studi Manajemen dalam penyelenggaraan pendidikan menawarkan 4 program peminatan bagi mahasiswa, yaitu: 1). Manajemen SDM, 2). Manajemen Keuangan, 3) Manajemen Pemasaran.

3. Mahasiswa

Mahasiswa adalah seseorang yang terdaftar secara resmi untuk mengambil program studi pada suatu perguruan tinggi dengan batas usia 18 sampai 30 tahun dan dapat dianggap sebagai kelompok masyarakat yang memperoleh statusnya karena keterkaitan (Alfian, 2014)

Mahasiswa adalah orang yang menempuh pendidikan di perguruan tinggi, baik itu universitas, maupun akademi. Mereka yang terdaftar sebagai murid di perguruan tinggi dapat disebut mahasiswa. Pada dasarnya makna dari mahasiswa tidak sesempit itu, Mendaftar sebagai mahasiswa di perguruan tinggi hanyalah persyaratan administratif. Secara etimologis, mahasiswa terdiri dari dua kata, yaitu “maha” dan “siswa”. Maha artinya sangat, amat dan besar, sedangkan siswa artinya murid atau pelajar (Kurniawan, 2021).

Mahasiswa merupakan generasi penerus bangsa yang diharapkan mampu untuk bersaing dan mengharumkan nama bangsa. Mahasiswa juga diharapkan mampu untuk menyatukan dan mampu untuk menyampaikan pikirannya untuk memajukan bangsa. Mahasiswa dianggap sebagai kaum intelektual oleh masyarakat. Perpaduan antara kesadaran akan kepercayaan masyarakat terhadap Indonesia yang lebih baik dan kesempatan untuk menjadi intelektual yang dapat

menjadi kekuatan besar dalam menjadikan Indonesia negara yang hebat. Selain itu, mahasiswa merupakan asset yang berharga dan juga harapan tinggi suatu bangsa karena mahasiswa merupakan penerus yang diharapkan memiliki loyalitas tinggi terhadap kemajuan bangsa terutama dalam dunia Pendidikan.

Berdasarkan beberapa pendapat di atas mengenai pengertian mahasiswa, penulis menyimpulkan bahwa mahasiswa adalah individu sekaligus sivitas akademika yang terdaftar secara resmi untuk mengambil mata kuliah di perguruan tinggi dan secara aktif mengetahui perkembangan dirinya. potensi belajar, , pencarian kebenaran ilmiah, dan/atau penguasaan, pengembangan dan pengamalan suatu cabang ilmu untuk menjadi calon intelektual yang terdidik.

4. Data Mining

Data mining merupakan tahapan pengelolaan data yang menggunakan teknik statistika, kecerdasan buatan, machine learning untuk mengekstrak serta mengidentifikasi informasi yang bermanfaat beserta indikator penting dari berbagai database seperti Kaggle. Tujuan dilakukan data mining adalah untuk memahami lebih jauh mengenai data yang diamati (Pangaribuan & Barus, n.d.)

Data mining adalah analisa terhadap data untuk menemukan hubungan yang jelas serta menyimpulkannya yang belum diketahui sebelumnya dengan cara terkini dipahami dan berguna bagi pemilik data tersebut (Benri et al., 2015). Secara garis besar, data mining dapat dikelompokkan menjadi 2 kategori utama, yaitu :

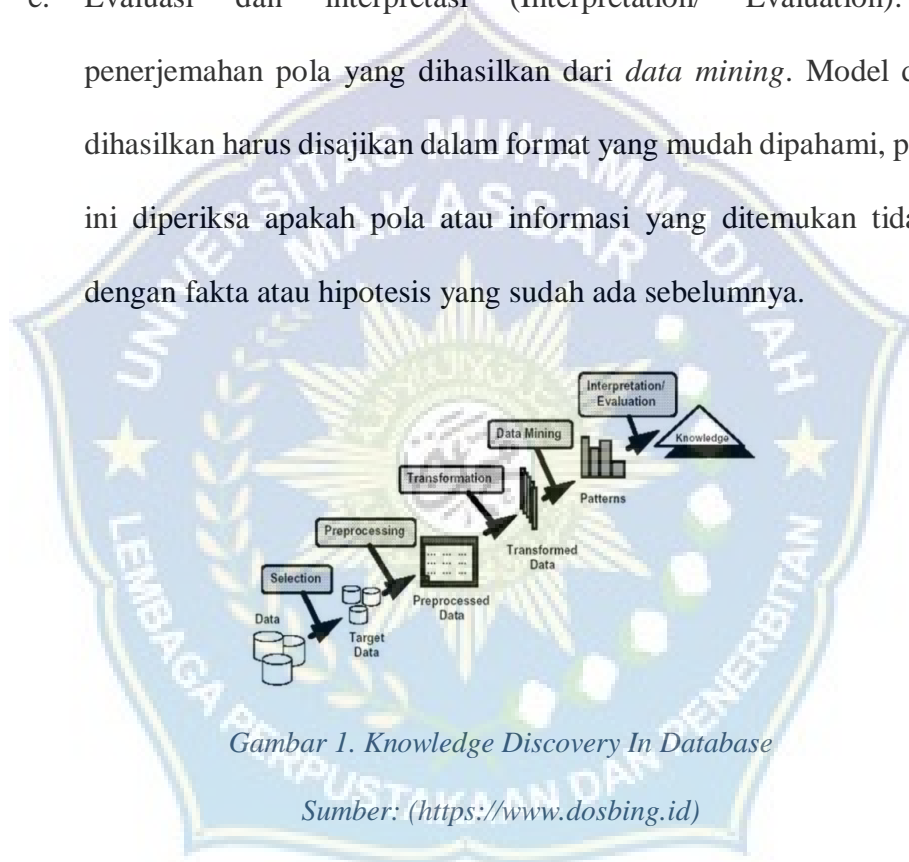
- a. Descriptive mining yaitu proses untuk menemukan karakteristik penting dari data dalam satu basis data. Teknik data mining yang termasuk descriptive mining adalah clustering, asosiasi, dan sequential mining.
- b. Predictive yaitu proses untuk menemukan pola dari data dengan menggunakan beberapa variabel di masa depan. Salah satu teknik yang terdapat dalam predictive mining adalah klasifikasi.

Menurut (Benri Melpa Metisen, 2015) *data mining* terdiri dari beberapa langkah, antara lain :

- a. Pemilihan data (*Data Selection*): Buat set data target, pilih set data atau fokuskan pada subset variabel atau sampel data untuk membuat penemuan. Hasil pemilihan disimpan dalam file terpisah dari database operasional.
- b. *Pre-processing / cleaning* data: *Pre-processing* dan pembersihan data adalah operasi dasar, seperti pengurangan kebisingan. Proses pembersihan melibatkan penghapusan data duplikat, memeriksa data yang bertentangan dan memperbaiki kesalahan dalam data seperti kesalahan ketik. Data dapat diperkaya dengan data atau informasi eksternal terkait.
- c. *Transformation*: Merupakan proses integrasi data yang dipilih untuk membuat data cocok untuk proses penambangan data. Ini adalah proses yang sangat tergantung pada jenis atau pola data yang dicari dalam database.
- d. Pemodelan Data (*Data mining*): Pemilihan tugas *data mining* adalah pemilihan tujuan dari proses *knowledge discovery in database*, seperti karakterisasi, klasifikasi, regresi, pengelompokan, asosiasi, dll. Pemilihan

tugas penambangan data adalah pemilihan target dari proses *knowledge discovery in database*. Karakterisasi, klasifikasi, regresi, pengelompokan, asosiasi, dll. Pemilihan teknik, metode atau algoritma yang tepat sangat bergantung pada tujuan dan keseluruhan proses *knowledge discovery in database*.

- e. Evaluasi dan interpretasi (Interpretation/ Evaluation): Yakni, penerjemahan pola yang dihasilkan dari *data mining*. Model data yang dihasilkan harus disajikan dalam format yang mudah dipahami, pada tahap ini diperiksa apakah pola atau informasi yang ditemukan tidak sesuai dengan fakta atau hipotesis yang sudah ada sebelumnya.



Gambar 1. Knowledge Discovery In Database

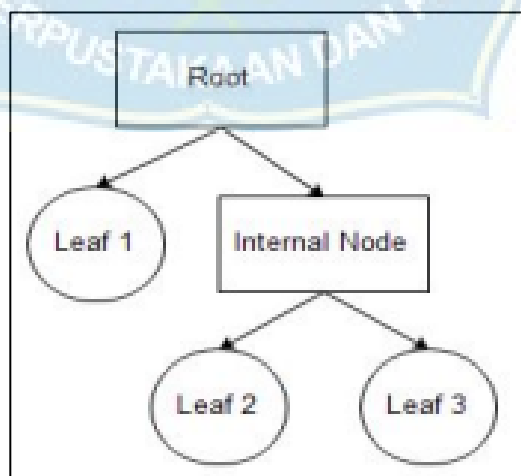
Sumber: (<https://www.dosbing.id>)

Data mining dapat membantu mengidentifikasi model dan tren bisnis yang dapat digunakan untuk membuat keputusan yang lebih baik dan lebih efisien. Misalnya, penambangan data dapat membuat operasi lebih efisien, meningkatkan penjualan, meningkatkan layanan pelanggan, dan meningkatkan keamanan informasi. Namun, penggunaan penambangan data juga dapat menimbulkan

masalah privasi, keamanan, dan etika. Oleh karena itu, saat melakukan analisis data menggunakan teknik penambangan data, perhatian khusus harus diberikan pada perlindungan data pribadi dan data sensitif.

5. Decision Tree

Pembuatan model untuk memprediksikan kelas suatu objek berdasarkan atribut yang dimilikinya merupakan hal yang harus dilakukan dalam data mining dengan predictive method. Decision tree merupakan salah satu metode untuk melakukan pembuatan model tersebut dan cukup terkenal karena mudah untuk diinterpretasikan, tingkat akurasi yang baik, dan efisien dalam menangani atribut diskret ataupun numerik/kontinyu (Bening, 2014). Decision tree mempunyai konsep mengubah data menjadi pohon keputusan lalu pohon keputusan diubah menjadi aturan-aturan keputusan. Data yang terdapat pada pohon keputusan biasanya berbentuk tabel yang memiliki atribut dan record. Atribut menyatakan parameter yang digunakan sebagai kriteria dalam pembentukan tree.



Gambar 2. Decision Tree

Decision tree terdiri dari Simpul akar, yang tidak memiliki cabang masukan dan berpengaruh paling besar pada suatu kelas tertentu. Simpul internal yang merepresentasikan test atau subset dari sebuah Atribute dan simpul daun yang merupakan sebuah sambungan node dari tree berupa class label.

6. Classification and Regression Tree (CART)

CART merupakan salah satu metode atau algoritma dari salah satu teknik pohon keputusan. CART terbilang sederhana namun merupakan metode yang kuat. CART bertujuan untuk mendapatkan suatu kelompok data yang akurat sebagai tanda dari suatu pengklasifikasian, selain itu CART juga dapat digunakan untuk menggambarkan hubungan antara variabel terikat dengan satu atau lebih variabel bebas. Model pohon yang dihasilkan bergantung pada skala variabel terikat, jika variabel terikat data berbentuk kontinu maka model pohon yang dihasilkan adalah regression tree (pohon regresi) sedangkan bila variabel terikat mempunyai skala data kategorik maka pohon yang dihasilkan adalah classification tree (pohon klasifikasi) (Breiman dkk dalam Pratiwi dan Zain (2014).

CART mempunyai beberapa kelebihan dibandingkan metode lainnya, yaitu hasilnya lebih mudah diinterpretasikan, lebih akurat dan lebih cepat perhitungannya, selain itu CART bisa diterapkan untuk himpunan data yang mempunyai jumlah besar, variabel yang sangat banyak dengan skala variabel campuran melalui prosedur pemilahan biner. Data training digunakan untuk pembentukan pohon klasifikasi optimal sedangkan data testing digunakan untuk

validasi model yaitu seberapa besar kemampuan model dalam memprediksi data baru.

Proses pembentukan pohon klasifikasi pada algoritma CART melalui tiga tahapan, yaitu:

a. Pemilihan Pemilah

Data yang digunakan merupakan sampel data learning. Himpunan bagian yang dihasilkan dari proses pemilahan harus lebih homogen dibandingkan pemilahan sebelumnya. Menurut Breiman dalam Akbar dkk (2010), Rumus pemilah disajikan seperti berikut:

- Variabel bebas kontinu = $b-1$ pemilahan
- Variabel bebas kategori nominal = $1 \ 2 \ 1 \ L- -$ pemilahan
- Variabel bebas kategori ordinal = $L - 1$ pemilahan

$$j(t) = \sum_{j \neq k}^m P(j|t)P(k|t)$$

Keterangan :

b : Banyaknya data pada suatu variabel

L : Banyaknya kategori pada suatu variabel

Fungsi keheterogenan yang digunakan adalah Indeks Gini karena akan selalu memisahkan kelas dengan anggota paling besar/kelas terpenting dalam simpul terlebih dahulu. Fungsi Indeks Gini ditunjukkan pada persamaan berikut:

Keterangan :

j, k : Kelas

$P(j/t)$: Probabilitas bersyarat kelas j yang berada dalam node t

$P(k/t)$: Probabilitas bersyarat kelas k yang berada dalam node t

Dengan $j(t)$ adalah fungsi keheterogenan indeks gini, $P(j/t)$ adalah peluang j pada node t , dan $P(k/t)$ adalah peluang k pada node t .

Rumus indeks gini dapat dituliskan:

$$j(t) = 1 - \sum_{j=1}^m P^2(j|t)$$

$Node t$ dibelah menjadi 2 subset $D1$ dan $D2$ dengan ukuran masing-masing $b1$ dan $b2$, indeks gini dari pembelahan tersebut didefinisikan sebagai berikut:

$$Gini_{pembelahan}(t) = \frac{b_1}{b} gini(D_1) + \frac{b_2}{b} gini(D_2)$$

Keterangan :

$Gini_{pembelahan}$: Nilai indeks $gini$ setiap variabel

$gini(D1)$: Nilai indeks $gini$ subset $D1$ pada setiap variabel

$gini(D2)$: Nilai indeks $gini$ subset $D2$ pada setiap variabel

b : Banyaknya data pada suatu variabel

$b1$: Banyaknya data pada subset $D1$

$b2$: Banyaknya data pada subset $D2$

b. Penentuan $Node$ Terminal

Suatu *node t* akan menjadi *node* terminal atau tidak, akan dipilah kembali apabila terdapat batasan minimum n seperti hanya terdapat satu pengamatan pada tiap *node* anak. Umumnya jumlah kasus minimum dalam suatu terminal akhir adalah 5, dan apabila hal itu terpenuhi maka pengembangan pohon akan dihentikan

c. Penandaan Label Kelas

Penandaan label kelas pada simpul terminal berdasarkan aturan jumlah terbanyak dengan persamaan:

$$P(j_0 | t) = \max_j P(j | t) = \max_j \frac{m_j(t)}{m(t)}$$

Keterangan :

$P(j | t)$: Probabilitas bersyarat kelas j yang berada pada *node t*

$m_j(t)$: Jumlah pengamatan pada kelas j pada *node t*

$m(t)$: Jumlah pengamatan pada *node t*

Label kelas *node* terminal t adalah j_0 yang memberi nilai dugaan kesalahan pengklasifikasian *node t* terbesar

7. Pengujian *White Box*

Metode pengujian aplikasi yang menggunakan struktur kontrol yang dianotasi dalam kerangka umum tingkat komponen untuk membuat kasus pengujian disebut Pengujian *White box*. *White box testing* menggunakan beberapa cara, antara lain data flow testing, control flow testing, path/basic path testing, dan loop testing. Dalam pengujian ini, penguji harus memiliki pengetahuan yang mendalam tentang sumber hukum yang akan diuji. Berkat Pengujian Kotak Putih,

dimungkinkan untuk mendeteksi kesalahan dalam penerapan aplikasi. Metode ini dapat diterapkan pada tingkat integrasi, unit dan sistem (Londjo, 2021).

Ada beberapa kelebihan dan kekurangan pengujian dengan menggunakan metode White Box, diantaranya (Londjo, 2021)

- a. Metode white box dapat mengungkapkan kesalahan pada kode dengan menghilangkan baris-baris yang tidak diperlukan.
- b. Manfaat maksimal dari cakupan pengujian aplikasi dapat dicapai saat menguji skenario.
- c. Kelemahannya, Pengujian dengan metode White Box membutuhkan biaya yang tinggi karena membutuhkan tester yang berpengalaman dalam down ini.

Dalam pengujian sistem, terdapat beberapa teknik pengujian White Box yang dapat digunakan seperti yang dijelaskan oleh Londjo (2021). Salah satunya adalah :

- a. Teknik Basis Path yang pertama kali dikemukakan oleh Tom McCabe. Teknik ini memungkinkan penguji untuk mengevaluasi kompleksitas logika dari perancangan prosedural. Saat menggunakan Teknik Basis Path, kasus uji coba yang dibuat akan mengeksekusi setiap pernyataan aplikasi yang diujikan dan dijalankan setidaknya satu kali selama tahap pengujian (Londjo, 2021).
- b. Teknik flowchart, yaitu gambaran tentang langkah dan prosedur dari sebuah aplikasi. Flowchart membantu analis dan pemrogram

memvisualisasikan aplikasi dan memecahnya menjadi segmen yang lebih kecil untuk analisis yang lebih mudah. Flowchart digunakan dalam desain aplikasi untuk menggambarkan logika aplikasi (Londjo, 2021).

- c. Teknik Flowgraph yaitu grafik program yang dihasilkan dari pemetaan flowchart program yang ada. Flowchart digunakan untuk mewakili logika program yang terdapat dalam flowchart. Pada tahap pengujian, flowchart digunakan untuk fokus mewakili aliran program. (Londjo, 2021)

8. Daftar Simbol

Bagan alir atau *flowchart* adalah representasi visual yang menunjukkan aliran logika dalam program sistem atau prosedur. Flowchart digunakan untuk memodelkan input, urusan, proses, dan transaksi menggunakan simbol khusus. Tujuan utama pembuatan flowchart adalah untuk memberikan kemudahan bagi pengguna untuk memahami alur atau transaksi dari sistem (Jogiyanto, 2001).

A. Daftar Simbol *flowchart*

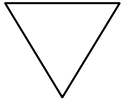


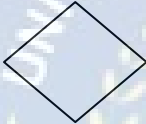
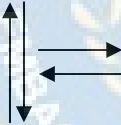

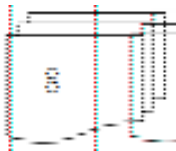
Tabel 1. Simbol Flow Map

Sumber : Jogiyanto Hartono, MBA, Ph.D.(2004:754)

NO	SIMBOL	KETERANGAN	DESKRIPSI
1		Kesatuan luar (Eksternal Entity)	Merupakan kesatuan diluar lingkungan sistem yang dapat berupa orang, organisasi, atau sistem lain adalah entitas eksternal.
2		Arus Data	Merupakan arus dari data yang dapat berupa masukan untuk sistem atau hasil dari proses sistem disebut aliran data. Aliran data .
3		Proses	Kegiatan atau kerja yang dilakukan oleh orang, mesin, atau komputer dari hasil suatu aliran data yang masuk ke dalam proses untuk menghasilkan aliran data yang akan keluar dari proses disebut aktivitas.
4		Simpanan Data	Simpanan data dapat berupa suatu file, arsip catatan, buku, atau simbol catatan adalah penyimpanan data.

Tabel 2. Simbol Flow Map

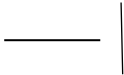
Sumber : Jogiyanto Hartono, MBA, Ph.D.(2004:754)

NO	SIMBOL	KETERANGAN	DESKRIPSI
1		Simbol Arsip	Simbol digunakan untuk menunjukkan data yang diarsipkan dapat berupa simbol arsip.
2		Simbol Dokumen	Simbol digunakan untuk mewakili data/dokumen dapat berupa simbol kotak atau simbol dokumen.
3		Simbol proses	Simbol digunakan untuk mewakili suatu proses komputerisasi dapat berupa simbol persegi atau simbol elips.
4		Simbol keputusan	Simbol digunakan untuk suatu penyeleksian kondisi di dalam program dapat berupa simbol berbentuk diamond.
5		Simbol garis alir	Simbol digunakan untuk menunjukkan arus dari proses dapat berupa panah yang menghubungkan simbol-simbol proses.
6		Simbol proses	Simbol digunakan untuk mewakili suatu proses manual dapat berupa simbol berbentuk panah yang menghubungkan simbol-simbol proses manual.
7		Simbol Rangkap Dokumen	Simbol digunakan untuk mewakili data/dokumen yang dirangkap tiga dapat berupa simbol berbentuk tiga kotak yang berdekatan.

Tabel 3. Simbol Entity Relationship Diagram (ERD)

Sumber : Fathansyah, (2002 : 65)

NO	SIMBOL	KETERANGAN	DESKRIPSI
1		Simbol Entity /objek data	Simbol ini adalah kumpulan objek atau suatu yang dapat dibedakan atau dapat diidentifikasi secara unik, kumpulan entitas yang sejenis dapat diwakili oleh simbol berbentuk persegi panjang yang mewakili tabel atau entitas dalam basis data.
2		Simbol Relationship	Simbol ini adalah hubungan yang terjadi antara satu entitas atau lebih, kumpulan relationship yang sejenis dapat diwakili oleh simbol berbentuk panah yang menghubungkan entitas atau tabel dalam basis data..
3		Simbol garis	Simbol digunakan untuk menghubungkan antar tabel dalam basis data dapat berupa garis atau panah yang menghubungkan kolom yang berkaitan antara tabel-tabel tersebut.
4		Simbol Entity /objek data	Simbol ini adalah kumpulan objek atau suatu yang dapat dibedakan atau dapat diidentifikasi secara unik, kumpulan entitas yang sejenis dapat diwakili oleh simbol berbentuk persegi panjang yang mewakili tabel atau entitas dalam basis data.
5		Simbol Relationship	Simbol ini adalah hubungan yang terjadi antara satu entitas atau lebih, kumpulan relationship yang sejenis dapat diwakili oleh simbol berbentuk panah yang menghubungkan entitas atau tabel dalam basis data..

6		Simbol garis	Simbol digunakan untuk menghubungkan antar tabel dalam basis data dapat berupa garis atau panah yang menghubungkan kolom yang berkaitan antara tabel-tabel tersebut.
---	---	--------------	--

9. Visual Studio Code

Visual Studio Code merupakan editor kode sumber yang sangat fleksibel dan dapat digunakan untuk beberapa bahasa pemrograman serta memiliki fitur yang sangat berguna untuk pengembangan perangkat lunak, seperti debugging dan integrasi Git, dan Webpack dalam bukunya Mastering React (2017),

Visual Basic .NET adalah versi Visual Basic yang didesain ulang yang dirancang untuk digunakan pada platform .NET. Menggunakan Visual Basic .NET, aplikasi yang dihasilkan dapat berjalan di sistem komputer apa pun dan dapat memulihkan data dari beragam jenis apa pun, asalkan diinstal .NET Framework, Fridayanthie Wulansari (2015)

Visual Studio Code, di sisi lain, adalah editor teks tujuan umum yang dikembangkan oleh Microsoft. Editor ini tersedia untuk Windows, Linux dan Mac. Visual Studio Code secara native mendukung berbagai bahasa pemrograman seperti JavaScript, TypeScript, dan Node.js. Selanjutnya melalui plugin yang bisa dipasang di Visual Studio Code, kita bisa mengembangkan dengan bahasa pemrograman lain seperti C, C#, Python, Go, Java dan masih banyak lainnya..

B. Penelitian Terkait

Peneliti memberikan segudang inspirasi dan informasi untuk penyempurnaan proposal ini dari penelitian sebelumnya. Penelitian yang terkait sebelumnya meliputi:

1. (Analisis Pola Pemilihan Konsentrasi Ilmu Jurusan Sistem Informasi Di STMIK TIME, Herman., 2017).

Penelitian pertama dilakukan oleh Herman (2017) dengan judul “Analisis Pola Pemilihan Konsentrasi Ilmu Jurusan Sistem Informasi Di STMIK TIME. Pada penelitian yang menggunakan decision tree, dengan klasifikasi yang dilakukan mengacu pada 34 mata kuliah, dan setelah dilakukan pengolahan mata kuliah yang berpengaruh pada penentuan pola karakteristik pemilihan konsentrasi ilmu jurusan sistem informasi, diperoleh pola karakteristik berupa pohon keputusan atas mata kuliah yang mempengaruhi pemilihan program studi, Berdasarkan perhitungan data mining dengan metode decision tree atas data mahasiswa STMIK TIME diperoleh kesimpulan bahwa mata kuliah yang paling berpengaruh terhadap pemilihan konsentrasi ilmu jurusan sistem informasi adalah Pengantar Akuntansi.

2. (Perancangan Aplikasi Data Mining Menggunakan Metode Decision Tree Untuk Analisa Pemberian Kredit Pada Bri Unit Lalabata Rilau, M. Afdal Tahir, 2019).

Penelitian kedua dilakukan oleh M. Afdal Tahir (2019) dengan judul “ Perancangan Aplikasi Data Mining Menggunakan Metode Decision Tree Untuk Analisa Pemberian Kredit Pada Bri Unit Lalabata Rilau ". Tujuan dari penelitian ini untuk membuat aplikasi analisa pemberian kredit berbasis komputer yang

memiliki kemampuan dalam mengolah data secara cepat dan memiliki data yang terintegrasi. Metode yang digunakan yaitu decision tree untuk melakukan proses perhitungan analisis keputusan. Hasil penerapan aplikasi data mining dengan metode decision tree sangat membantu pihak BRI dalam proses analisis dan memudahkan untuk mengambil keputusan.

3. (Sistem Pendukung Keputusan Penentuan Skincare Untuk Kulit Wajah Menggunakan Metode Decision Tree (Studi Kasus Pada Kosmetik Wardah), Lenny Estika Sari , Wiwien Hadikurniawati 2020).

Penelitian ketiga dilakukan oleh Lenny Estika Sari , Wiwien Hadikurniawati (2020), berjudul “ Sistem Pendukung Keputusan Penentuan Skincare Untuk Kulit Wajah Menggunakan Metode Decision Tree (Studi Kasus Pada Kosmetik Wardah) “. Penelitian ini menggunakan Metode Decision Tree dan dapat disimpulkan bahwa Dengan adanya sistem pengambilan keputusan pemilihan skincare wardah yang sesuai dengan jenis kulit wajah yang bermanfaat untuk mempermudah konsumen dalam memilih skincare yang tepat untuk kulit wajah. Penerapan metode decision tree dalam penentuan skincare wardah yang tepat untuk permasalahan kulit wajah telah berhasil diterapkan dan memiliki tingkat keakuratan pengambilan keputusan yang tinggi.

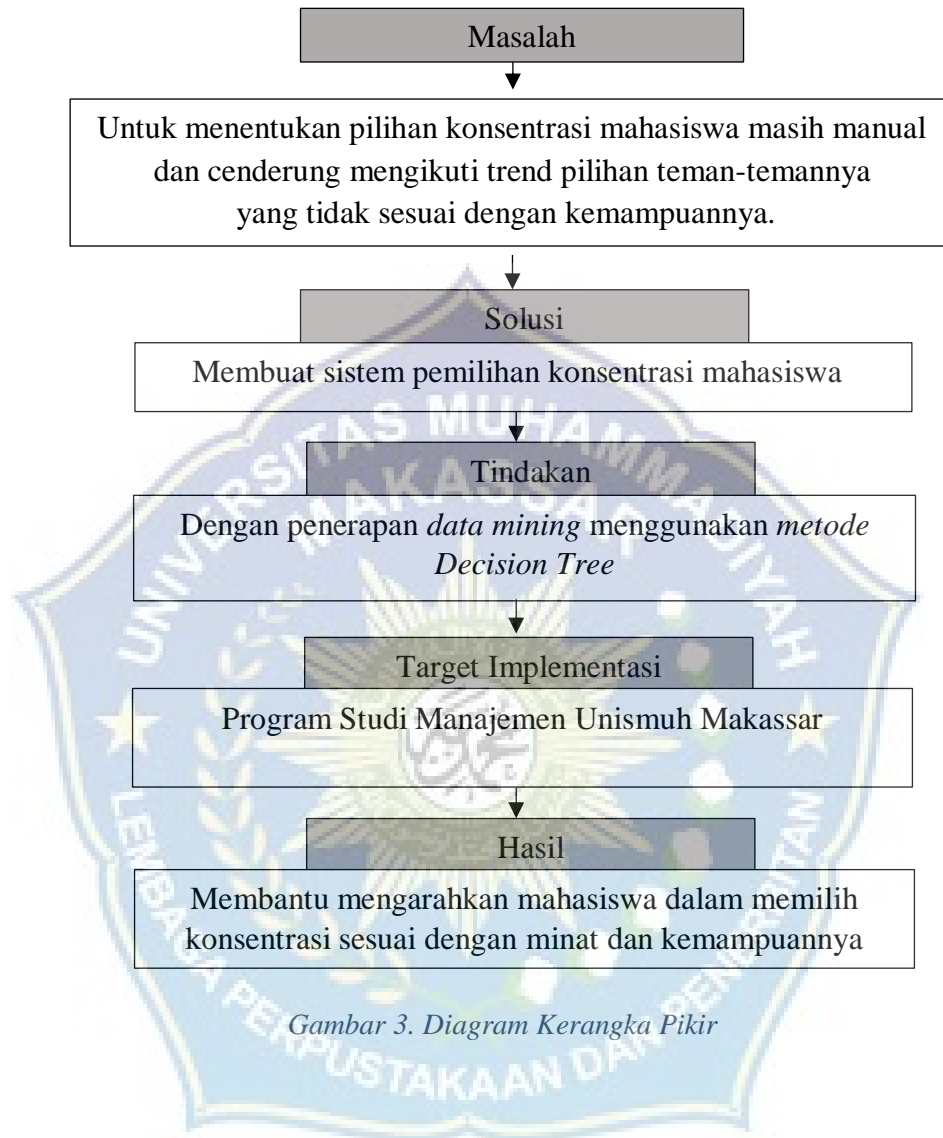
4. (Sistem Pendukung Keputusan Penentuan Skincare Untuk Kulit Wajah Menggunakan Metode Decision Tree (Studi Kasus Pada Kosmetik Wardah, Sutrisman Sulaeman, Verry Ronny Palilingan, Olivia Eunike Selvie Liando 2021).

Penelitian keempat dilakukan oleh Sutrisman Sulaeman, Verry Ronny Palilingan, Olivia Eunike Selvie Liando 2021 dengan judul “ c ini dapat membantu mahasiswa Prodi Pendidikan Teknologi Informasi dan Komunikasi untuk memperoleh rekomendasi konsentrasi jurusan apa yang cocok untuk mereka ambil pada semester empat berdasarkan hasil dari 13 mata kuliah yang dijadikan atribut. Walaupun demikian, hasil dari pengambilan keputusan ini bukan suatu hal yang mutlak karena sistem pengambilan keputusan (SPK) merupakan pemberi alternatif solusi keputusan, sedangkan keputusan mutlak tetap diputuskan oleh pihak pengambil keputusan. Pendukung Keputusan Pemilihan Konsentrasi Jurusan Menggunakan Teknik Naive Bayes ”.

5. (Penerapan Decision Tree Dengan Algoritma C4.5 Untuk Menentukan Rekomendasi Kenaikan Jabatan Karyawan, Elmi Devia.2023)

Penelitian kelima dilakukan oleh Elmi Devia. (2023) bernama " Penerapan Decision Tree Dengan Algoritma C4.5 Untuk Menentukan Rekomendasi Kenaikan Jabatan Karyawan ". Penelitian ini menggunakan metode Decision Tree Dengan Algoritma C4.5 Dari hasil percobaan dengan menggunakan perangkat lunak RapidMiner dalam penerapan terhadap masalah penentuan rekomendasi kenaikan jabatan karyawan, decision tree dengan algoritma C4.5 dapat memberikan hasil yang lebih baik dan lebih mudah dibandingkan perhitungan secara manual.

C. Kerangka Pikir



Gambar 3. Diagram Kerangka Pikir

BAB III

METODE PENELITIAN

A. Jenis Penelitian

Jenis penelitian yang dilakukan adalah penelitian kuantitatif. Penelitian ilmiah yang dilakukan secara sistematis dengan menggunakan model-model matematis dan teori-teori yang berkaitan. Penelitian yang dilakukan menggunakan salah satu teknik pada data mining yaitu algoritma Decision Tree.

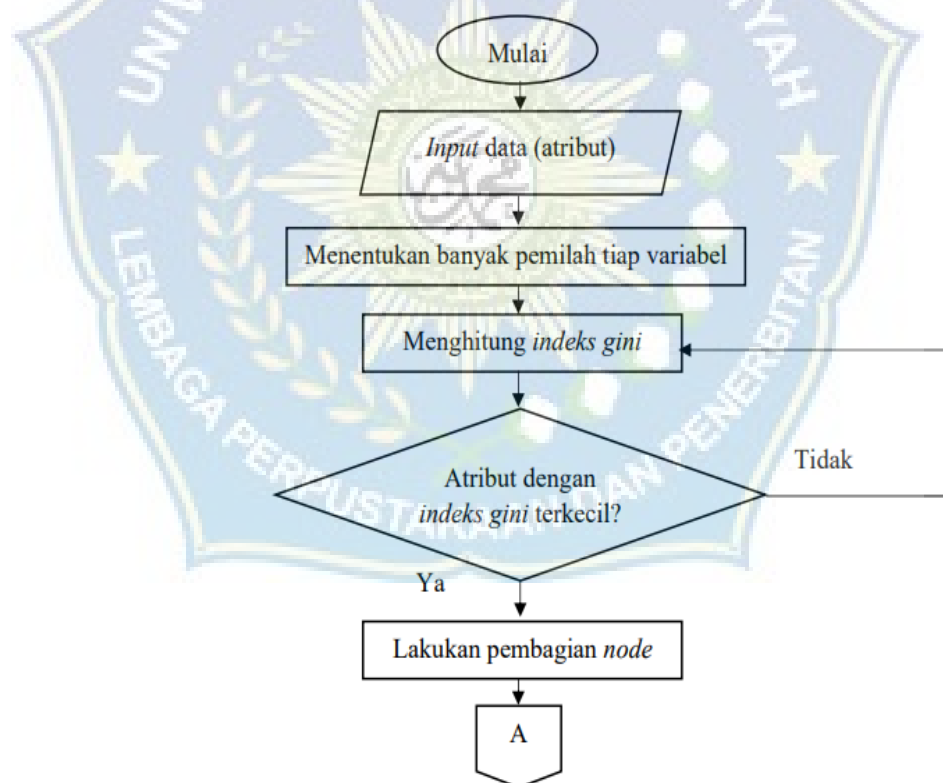
B. Data dan Sumber Data

Dataset yang digunakan pada penelitian ini adalah data nilai mahasiswa Program Studi Manajemen Angkatan 2018, 2019, 2020 berupa data sekunder yang bersumber dari Simak Universitas Unismuh Makassar. Data diambil dalam bentuk tabel dengan format file *XLSX. Pengambilan data dilakukan pada tanggal 10 Agustus 2023.

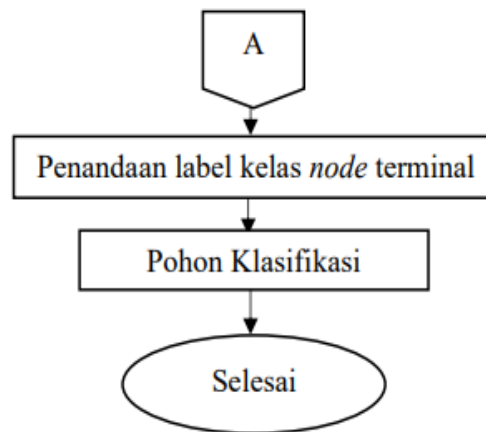
Data selanjutnya yang digunakan pada penelitian ini adalah Daftar Mata Kuliah Program Studi Manajemen semester satu sampai lima dan konsentrasi apa saja yang terdapat pada Program Studi Manajemen yang bersumber dari Pihak Pimpinan Program Studi Manajemen

C. Perancangan Sistem

Rancangan sistem adalah proses memodelkan dan merancang sebuah sistem dengan tujuan memenuhi kebutuhan dan menyelesaikan masalah tertentu, terutama dalam konteks penelitian ini. Rancangan sistem melibatkan pemilihan teknologi yang sesuai, struktur sistem, desain antarmuka pengguna, pemilihan metode dan algoritma pemrograman, serta pengujian menyeluruh terhadap sistem. Peneliti merancang sebuah Flowchart sebagai alat bantu untuk memudahkan pembuatan dan pengembangan sistem, sehingga alur sistem yang dibangun dapat dipahami dengan mudah dan didesain secara terstruktur.



Gambar 4 . Diagram Alir Algoritma CART



Gambar 5. Diagram Alir Algoritma CART (Lanjutan)

Gambar 4. Dalam analisis menggunakan Algoritma CART ada beberapa langkah pertama melakukan penentuan variabel yang akan diteliti kemudian melakukan penentuan banyaknya pemilah per variabel sesuai dengan jenis variabel bebasnya menggunakan persamaan selanjutnya menghitung nilai indeks gini untuk setiap pemilah sesuai dengan persamaan kemudian pemilah yang memiliki nilai indeks gini terkecil akan dipilih menjadi pemilah terbaik selanjutnya ulangi langkah hingga tidak memungkinkan lagi untuk melakukan pemilahan selanjutnya penandaan label kelas node terminal berdasarkan aturan jumlah anggota terbanyak menggunakan persamaan

D. Teknik Pengujian Sistem

Teknik pengujian yang akan digunakan dalam pengujian ini adalah white-box testing. White-box testing merupakan sebuah teknik pengujian perangkat lunak yang melibatkan pemeriksaan dan pengujian struktur internal perangkat lunak. Teknik ini juga dikenal sebagai pengujian berbasis kode atau pengujian struktural.

Dalam white-box testing, pengujian dilakukan dengan memeriksa secara terperinci aliran logika, kontrol aliran, dan struktur kode program. Tujuan dari

pengujian ini adalah untuk memastikan bahwa semua kondisi logis telah diuji dan keputusan yang tepat diambil, bahwa setiap cabang kode diuji, dan setiap pernyataan dalam kode dieksekusi setidaknya satu kali.

Pengujian white-box dilakukan pada tahap akhir pembuatan perangkat lunak atau sistem untuk memastikan bahwa perangkat lunak atau sistem dapat bekerja secara optimal. Tujuan utamanya adalah memastikan bahwa perangkat lunak yang dikembangkan memenuhi persyaratan fungsional, kinerja, dan keamanan yang diinginkan. Selain itu, teknik ini juga bertujuan untuk mendeteksi dan memperbaiki kesalahan dan cacat pada kode program.

Dengan melakukan pengujian white-box, diharapkan dapat menjamin bahwa perangkat lunak atau sistem yang dikembangkan bekerja sesuai dengan harapan dan dapat mengatasi masalah yang mungkin terjadi.

E. Teknik Analisis Data

Teknik analisis data melibatkan serangkaian langkah, termasuk pengumpulan, pembersihan, transformasi, dan pengolahan data, dengan tujuan untuk memperoleh informasi yang berarti dan bernilai. Fokus utama dari teknik analisis data adalah mengidentifikasi pola, tren, dan hubungan dalam data agar dapat digunakan dalam pengambilan keputusan dan memberikan wawasan yang lebih mendalam.

Teknik analisis data mencakup berbagai metode dan alat yang menggunakan pendekatan statistik, matematika, atau komputasi untuk memproses dan memanipulasi data. Beberapa metode yang sering digunakan dalam analisis

data meliputi penambangan data, pembelajaran mesin, statistik deskriptif, dan visualisasi data.

Penambangan data (data mining) digunakan untuk mengeksplorasi data dan menemukan pola atau informasi berharga yang tersembunyi di dalamnya. Pembelajaran mesin (machine learning) melibatkan penggunaan algoritma dan model statistik untuk melatih komputer agar dapat belajar dari data dan membuat prediksi atau keputusan tanpa perlu diatur secara eksplisit. Statistik deskriptif digunakan untuk menggambarkan, merangkum, dan menginterpretasikan data, sementara visualisasi data memberikan representasi visual yang membantu pemahaman dan analisis data secara intuitif.

Proses analisis data pada penelitian ini memiliki 3 langkah yaitu:

1. Reduksi Data

Reduksi data mengacu pada proses mengurangi jumlah data yang diperoleh dari lapangan agar menjadi lebih terkelola dengan baik. Menurut Sugishirono (2015), data yang telah direduksi memberikan gambaran yang lebih jelas dan memudahkan peneliti untuk mengumpulkan dan mengambil data selanjutnya sesuai kebutuhan. Dalam penelitian kualitatif, peneliti harus berhati-hati dalam mereduksi data, terutama jika terdapat outlier, unknown's, atau kualifikasi yang tidak berpola, karena fokus utama adalah hasil penelitian.

2. Display Data

Penyajian data merupakan tahapan kedua dalam analisis data. Penyajian data adalah proses menyusun informasi secara terstruktur dan dijelaskan dalam bentuk narasi, yang memungkinkan peneliti untuk menarik kesimpulan dan mengambil tindakan berdasarkan hasil penelitian. Penyajian data membantu peneliti memahami apa yang sedang terjadi dan memungkinkan dilakukannya analisis lebih lanjut. Dalam penelitian kualitatif, data dapat disajikan dalam bentuk deskripsi singkat, diagram, interaksi antar kategori, flowchart, dan lain sebagainya.

3. Penarikan Kesimpulan

Penarikan kesimpulan merupakan tahapan akhir dalam teknik analisis data kualitatif. Pada tahap ini, peneliti mencari makna dari data yang dikumpulkan dengan mencari hubungan, persamaan, atau perbedaan guna menarik kesimpulan sebagai jawaban dari permasalahan penelitian. Kesimpulan yang dihasilkan masih bersifat sementara dan dapat berubah jika tidak ditemukan bukti pendukung pada tahap pengumpulan data selanjutnya. Validasi data dilakukan untuk memastikan keabsahan kesimpulan yang dihasilkan. Kesimpulan yang didukung oleh bukti-bukti yang valid dianggap kredibel, namun tetap perlu dipertimbangkan bahwa kesimpulan pada penelitian kualitatif dapat berkembang seiring dengan penelitian yang dilakukan di lapangan.

BAB IV HASIL DAN PEMBAHASAN

A. Deskripsi Dataset

Pada bab ini, akan disajikan hasil dan pembahasan mengenai penentuan konsentrasi mahasiswa menggunakan Algoritma Decision Tree. Hasil eksperimen meliputi pengelompokan data mahasiswa berdasarkan atribut nama, nim, nama matakuliah, sks, nilai huruf, nilai angka, bobot semester satu hingga semester lima menggunakan algoritma Decision Tree.

B. Analisis Data Mentah

Data awal yang di olah dari penelitian ini merupakan data mahasiswa program studi manajemen universitas muhammadiyah Makassar angkatan 2018 sampai dengan 2020. Dataset yang didapatkan terdiri dari atribut data yaitu kode prodi, nama prodi, krs, nim, nama, kode matakuliah, mata kuliah dan sks, nilai huruf, nilai angka, bobot.

Tabel 4. Dataset mahasiswa manajemen 2018

NO	KODE PR ODI	NAMA PRODI	PERIODE KRS	NIM	NAMA	KODE MATAKULIAH	NAMA_MATA KULIAH	SKS	NILAI HURUF	NILAI ANGKA	BOBOT
1	61201	Manajemen	20181	105721100118	CITRA ANANDA	KK70201	PENGANTAR AKUNTANSI I	2	C	2	4
2	61201	Manajemen	20181	105721100118	CITRA ANANDA	PK204	BAHASA INDONESIA	2	A	4	8
3	61201	Manajemen	20181	105721100118	CITRA ANANDA	BB203	ILMU SOSIAL BUDAYA DASAR	2	E	0	0
.....
23137	61201	Manajemen	20181	105721153318	SHINTA WULAN DARY	PB72302	AKUNTANSI MANAJEMEN	3	A	4	12

Tabel 4. Dataset mahasiswa manajemen 2018, Dengan menggunakan dataset angkatan 2018 sebanyak 23137 Data sebagai dataset training (dataset latih) untuk menentukan akurasi dari system ini menggunakan metode Decision Tree. Nantinya pengaturan seperti nilai kluster akan di gunakan pada penerapan penentuan konsentrasi untuk dataset angkatan 2018 sebagai data uji.

Tabel 5. Dataset mahasiswa manajemen 2019

NO	KODE PRODI	NAMA PRODI	PERIODE		NAMA	KODE MATAKULIAH	NAMA MATAKULIAH	SKS	NILAI HURUF	NILAI ANGKA	BOBOT
			KRS	NIM							
1	61201	Manajemen	20191	105721100119	ARDINA CHAERUN NISA	KK70302	PENGANTAR BISNIS	3	A	4	12
2	61201	Manajemen	20191	105721100119	ARDINA CHAERUN NISA	PK205	BAHASA INGGRIS	2	A	4	8
3	61201	Manajemen	20191	105721100119	ARDINA CHAERUN NISA	KK70201	PENGANTAR AKUNTANSI	2	B	3	6
.....
20554	61201	Manajemen	20211	105721145719	OGI ANGARA PUTRA	IK104	AIK V	1	A	4	4

Tabel 5. Dataset mahasiswa manajemen 2019, Dengan menggunakan dataset angkatan 2019 sebanyak 20554. Data sebagai dataset training (dataset latih) untuk menentukan akurasi dari system ini menggunakan metode Decision Tree. Nantinya pengaturan seperti nilai kluster akan di gunakan pada penerapan penentuan konsentrasi untuk dataset angkatan 2019 sebagai data uji.

Tabel 6. Dataset mahasiswa manajemen 2020

NO	KODE PR ODI	NAMA PRODI	PERIODE	NIM	NAMA	KODE MATAKULIAH	NAMA_MAT AKULIAH	SK S	NIL AI HURUF	NIL AI ANGKA	BOBOT
1	61201	Manajemen	20201	105721100120	RISALDI	PK201	PENDIDIKAN AGAMA ISLAM	2	B	3	6
2	61201	Manajemen	20201	105721100120	RISALDI	BB202	ILMU KEALAMAN DASAR	2	A	4	8
3	61201	Manajemen	20201	105721100120	RISALDI	KK70307	MATEMATIKA EKONOMI	3	B	3	9
.....
15824	61201	Manajemen	20221	105721136420	MUHAMMAD FATHURRACHMAN	KK72311	PENGANGGARAN PERUSAHAAN	3	A	4	12

Tabel 6. Dataset mahasiswa manajemen 2020, Dengan menggunakan dataset angkatan 2020 sebanyak 15824. Data sebagai dataset training (dataset latihan) untuk menentukan akurasi dari system ini menggunakan metode Decision Tree. Nantinya pengaturan seperti nilai kluster akan di gunakan pada penerapan penentuan konsentrasi untuk dataset angkatan 2020 sebagai data uji.

C. Data Preprocessing

Data preprocessing adalah tahap yang krusial dalam penelitian ini. Langkah-langkah preprocessing data ini bertujuan untuk memastikan bahwa data yang digunakan dalam analisis dengan algoritma Decision Tree adalah bersih, konsisten, dan siap digunakan. Proses ini dimulai dengan pengumpulan data dari data yang sudah didapatkan. Diman terlebih dahulu data yang diambil dan dilanjutkan dengan proses perhitungan yaitu data nilai angka dimana :

- Nilai huruf A = 4
- Nilai huruf B = 3

- Nilai huruf C = 2
- Nilai huruf D = 1
- Nilai huruf A = 0

Tabel 8. Dataset Nilai mahasiswa manajemen 2018

NO	NAMA	AKUNTANSI MANAJEMEN	EKONOMI INTERNASIONAL	EKONOMI MANAJERIAL	KOMUNIKASI BISNIS	LEMBAGA KEUANGAN SYARIAH	MANAJEMEN KEUANGAN I	MANAJEMEN KEUANGAN II	MANAJEMEN OPERASIONAL I	MANAJEMEN OPERASIONAL II
1	A. ALDI MAPPIONANG	4	4	4	4	4	4	4	4	4
2	A. HAERUL EKAPUTRA	3	4	3	4	4	4	4	4	4
3	A. HASMA TENRILALA	4	4	4	3	4	4	3	3	4
...
435	ZULFITRA KISHAN	2	2	3	4	1	3	3	3	3

Tabel 8. Dataset nilai mahasiswa manajemen 2018, Dengan menggunakan dataset nilai mahasiswa Angkatan 2018 yang sudah dilakukan preprocessing sebanyak 435 data mahasiswa sebagai data testing untuk menguji sejauh mana metode Decision Tree dapat melakukan penentuan yang akurat pada data yang tidak pernah dilihat sebelumnya. Dengan anggota atribut nama, nama mata kuliah, nilai angka mata kuliah semester 1 sampai 5.

Tabel 9. Dataset Nilai mahasiswa manajemen 2019

NO	NAMA	AKUNT	EKONOM	EKONO	KOMU	LEMB	MANAJ	MANAJ	MANAJ	MANAJ
		ANSI	I	MI		AGA	EMEN	EMEN	EMEN	EMEN
		MANAJ	INTERNA	MANAJ	NIKASI	KEUA	KEUAN	KEUAN	OPERAS	OPERAS
		EMEN	SIONAL	ERIAL	BISNIS	NGAN	GAN I	GAN II	IONAL I	IONAL II
1	A. EKA WAHYU NI	4	3,75	4	3	4	4	4	4	4
2	A. FARA REZKY RAHMA DANI RASYID	4	3,5	4	4	4	4	3	4	4
3	A. MEGA WATI	2,75	4	3,75	4	3	3	4	4	3
...
404	ZURAC HMAD	3	3	4	4	4	4	4	4	4

Tabel 9. Dataset nilai mahasiswa manajemen 2019, Dengan menggunakan dataset nilai mahasiswa Angkatan 2019 yang sudah dilakukan preprocessing sebanyak 404 data mahasiswa sebagai data testing untuk menguji sejauh mana metode Decision Tree dapat melakukan penentuan yang akurat pada data yang tidak pernah dilihat sebelumnya, Dengan anggota atribut nama, nama mata kuliah, nilai angka mata kuliah semester 1 sampai 5.

Tabel 10. Dataset Nilai mahasiswa manajemen 2020

NO	NAMA	AKUNT	EKONOM	EKONO	KOMU	LEMB	MANAJ	MANAJ	MANAJ	MANAJ
		ANSI	I	MI		AGA	EMEN	EMEN	EMEN	EMEN
		MANAJ	INTERNA	MANAJ	NIKASI	KEUA	KEUAN	KEUAN	OPERAS	OPERAS
		EMEN	SIONAL	ERIAL	BISNIS	NGAN	GAN I	GAN II	IONAL I	IONAL II
1	A. ADI SUDAR MAN	3,5	4	4	3	4	3	4	4	3,75
2	A. AFIFAH RAMA DANI	4	3,5	4	4	3,75	3	4	4	4
...
311	ZULFA DLI	0	2	3,5	4	3,5	4	4	4	4

Tabel 10. Dataset nilai mahasiswa manajemen 2020, Dengan menggunakan dataset nilai mahasiswa Angkatan 2020 yang sudah dilakukan preprocessing sebanyak 311 data mahasiswa sebagai data testing untuk menguji sejauh mana metode Decision Tree dapat melakukan penentuan yang akurat pada data.

D. Pelabelan Data Mahasiswa

Hasil pengamatan dalam tahap pelabelan konsentrasi merupakan langkah penting dalam penelitian ini. Fokus pengamatan terutama ditujukan untuk mengidentifikasi dan memahami karakteristik unik yang membedakan tiga konsentrasi yang ada, yaitu Keuangan, Pemasaran, dan Keuangan. Pengamatan ini bertujuan untuk memberikan dasar yang kuat dalam proses pengambilan keputusan dan penentuan konsentrasi mahasiswa

1. Identifikasi Mata Kuliah Konsentrasi

Identifikasi terlebih dahulu mata kuliah inti atau mata kuliah yang paling relevan dengan masing-masing konsentrasi.

Tabel 11. Dataset Matakuliah Untuk Konsentrasi SDM

NO	MATA KULIAH	SEMESETER
1	MANAJEMEN PEMASARAN I	3
2	MANAJEMEN PEMASARAN II	4
3	EKONOMI MANAJERIAL	5
4	KOMUNIKASI BISNIS	2
5	MANAJEMEN OPERASIONAL I	3
6	MANAJEMEN OPERASIONAL II	4
7	SISTEM OPERASI MANAJEMEN	3
8	RISET OPERASIONAL	4
9	PEREKONOMIAN INTERNASIONAL	5
10	EKONOMI INTERNASIONAL	5
11	SITUS KELAYAKAN BISNIS	5

Tabel 11. Daftar mata kuliah dan semester untuk pemilihan konsentrasi SDM

Tabel 12. Dataset Matakuliah Untuk Konsentrasi Pemasaran

NO	MATA KULIAH	SEMESTER
1	MANAJEMEN KEUANGAN I	3
2	MANAJEMEN KEUANGAN II	4
3	LEMBAGA KEUANGAN SYARIAH	4
4	PENGANGGARAN PERUSAHAAN	5
5	AKUTANSI MANAJEMEN	5
6	KOMUNIKASI BISNIS	2
7	MANAJEMEN OPERASIONAL I	3
8	MANAJEMEN OPERASIONAL II	4
9	SISTEM OPERASI MANAJEMEN	3
10	RISET OPERASIONAL	4
11	PEREKONOMIAN INTERNASIONAL	5
12	EKONOMI INTERNASIONAL	5
13	SITUS KELAYAKAN BISNIS	5

Tabel 12. Daftar mata kuliah dan semseter untuk pemilihan konsentrasi Pemasaran.

Tabel 13. Dataset Matakuliah Untuk Konsentrasi Keuangan

NO	MATA KULIAH	SEMESTER
1	MANAJEMEN SDM I	3
2	MANAJEMEN SDM II	4
3	TEORI PENGAMBILAN KEPUTUSAN	5
4	KOMUNIKASI BISNIS	2
5	MANAJEMEN OPERASIONAL I	3
6	MANAJEMEN OPERASIONAL II	4
7	SISTEM OPERASI MANAJEMEN	3
8	RISET OPERASIONAL	4
9	PEREKONOMIAN INTERNASIONAL	5
10	EKONOMI INTERNASIONAL	5
11	SITUS KELAYAKAN BISNIS	5

Tabel 13. Daftar mata kuliah dan semseter untuk pemilihan konsentrasi Keuangan

2. Identifikasi Pelabelan Konsentrasi

Dari data yang sudah didapatkan yaitu 20 dataset, maka selanjutnya dilakukan pelabelan data dengan cara melakukan mengcleaning data yang nantinya akan diolah dengan melakukan perhitungan terhadap setiap nilai mata kuliah dari data yang telah didapatkan menggunakan Microsoft excel. Yang kemudian akan diolah dengan algoritma Decision Tree.

Tabel 14. Dataset pelabelan konsentrasi mahasiswa

NO	NAMA MAHASISWA	HASIL NILAI PEMASARAN N	HASIL NILAI SDM	HASIL NILAI KEUANGA N	KONSENTRASI MAHASISWA
1	A SUBHANATULLA HAS	3,82	3,64	3,62	PEMASARAN
2	A SUSANTI	3,91	3,91	4	KEUANGAN
3	A M FITRAH RAMADHANA	2,73	2,64	2,46	PEMASARAN
4	ABDUL RAZAK M UMAR	2,82	2,45	2,38	PEMASARAN
5	ALAMIAH ALAM	3,18	3,18	2,85	PEMASARAN-SDM
.....
20	ALIF	3	3	3	PEMASARAN-SDM-KEUANGAN

Tabel 14. Dataset pelabelan konsentrasi mahasiswa manajemen, Dengan menggunakan dataset 20 mahasiswa yang sudah dilakukan preprocessing sebagai data testing untuk melakukan penentuan konsentrasi.

Adapun langkah yang dilakukan dalam pelabelan konsentrasi dimana dalam perhitungan manual pada metode Decision Tree sebagai berikut :

1. Melakukan pemisahan mata kuliah terhadap ketiga konsentrasi yaitu keuangan, pemasaran dan sdm.
2. Melakukan Penjumlahan terhadap setiap nilai mata kuliah dari ketiga konsentrasi yang telah dilakukan pemisahan kemudian melakukan pembagian terhadap jumlah mata kuliah di setiap konsentrasi, Adapun mekanisme perhitungan sebagai berikut :

- Konsentrasi SDM

Total Mata Kuliah SDM + Nilai Mahasiswa

Jumlah Mata Kuliah SDM

- Konsentrasi Pemasaran

Total Mata Kuliah Pemasaran + Nilai Mahasiswa

Jumlah Mata Kuliah Pemasaran

- Konsentrasi Keuangan

Total Mata Kuliah Keuangan + Nilai Mahasiswa

Jumlah Mata Kuliah Keuangan

3. Setelah dilakukan penjumlahan dan pemabagian maka dilakukan penentuan secara manual terhadap konsentrasi yang nilainya tertinggi. Seperti contoh yang ada ditabel dibawah

Tabel 15. Dataset Penjumlahan Untuk Konsentrasi SDM

1	A.SUSANTI	MANAJEMEN SDM I	4
2		MANAJEMEN SDM II	4
3		TEORI PENGAMBILAN KEPUTUSAN	4
4		KOMUNIKASI BISNIS	4
5		MANAJEMEN OPERASIONAL I	4
6		MANAJEMEN OPERASIONAL II	4
7		SISTEM OPERASI MANAJEMEN	4
8		RISET OPERASIONAL	4
9		PEREKONOMIAN INTERNASIONAL	4
10		EKONOMI INTERNASIONAL	4
11		SITUS KELAYAKAN BISNIS	4
JUMLAH			4

Tabel 16. Dataset Penjumlahan Untuk Konsentrasi Pemasaran

NO	NAMA	MATA KULIAH	NILAI ANGKA
1	A.SUSANTI	MANAJEMEN PEMASARAN I	4
2		MANAJEMEN PEMASARAN II	4
3		EKONOMI MANAJERIAL	4
4		KOMUNIKASI BISNIS	4
5		MANAJEMEN OPERASIONAL I	4
6		MANAJEMEN OPERASIONAL II	3
7		SISTEM OPERASI MANAJEMEN	4
8		RISET OPERASIONAL	4
9		PEREKONOMIAN INTERNASIONAL	4
10		EKONOMI INTERNASIONAL	4
11		SITUS KELAYAKAN BISNIS	4
JUMLAH			4

Tabel 17. Dataset Penjumlahan Untuk Konsentrasi Keuangan

NO	NAMA	MATA KULIAH	NILAI ANGKA
1	A.SUSANTI	MANAJEMEN KEUANGAN I	4
2		MANAJEMEN KEUANGAN II	4
3		LEMBAGA KEUANGAN SYARIAH	4
4		PENGANGGARAN PERUSAHAAN	4
5		AKUTANSI MANAJEMEN	4
6		KOMUNIKASI BISNIS	4
7		MANAJEMEN OPERASIONAL I	4
8		MANAJEMEN OPERASIONAL II	4
9		SISTEM OPERASI MANAJEMEN	4
10		RISET OPERASIONAL	4
11		PEREKONOMIAN INTERNASIONAL	4
12		EKONOMI INTERNASIONAL	4
13		SITUS KELAYAKAN BISNIS	4
		JUMLAH	4

4. Seperti yang terlihat pada sampel data pada tabel 15, 16,17 diatas yang telah dilakukan pemisahan mata kuliah dan penjumlahan nilai mata kuliah secara manual dapat disimpulkan bahwa :

b. Nilai Konsentrasi Pemasaran : 4

c. Nilai Konsentrasi SDM : 4

d. Nilai Konsentrasi Keuangan : 4

5. Dari perhitungan yang telah dilakukan dapat disimpulkan bahwa hasil proses pelabelan secara manual dapat disimpulkan bahwa untuk mahasiswa A.Susanti dinyatakan rekomendasi konsentrasi yaitu Pemasaran, Sdm, dan Keuangan.

E. Perhitungan Algoritma CART

Terdapat beberapa tahapan dalam proses perhitungan manual dalam Algoritma CART sebagai berikut:

1. Tahap pertama adalah menentukan calon cabang yang dilakukan dari beberapa atribut prediktor yang telah di tentukan.

Tabel 18. Calon Cabang Mutakhir

Cabang	Cabang Kanan	Cabang Kiri
1	Studi kelayakan bisnis ≤ 3	Studi kelayakan bisnis ≤ 3
2	Manajemen keuangan I ≤ 4	Manajemen keuangan I ≤ 3
3	Manajemen keuangan I ≤ 3	Manajemen keuangan I ≤ 4
4	Komunikasi Bisnis ≤ 2	Komunikasi Bisnis ≤ 1
5	Akuntansi Manajemen ≤ 2	Akuntansi Manajemen ≤ 1

Tabel 18. Adapun penjelasan terkait tabel diatas adalah memperlihatkan perbandingan antara Cabang Kanan dan Cabang Kiri, dengan setiap baris mewakili suatu atribut predictor yang relevan untuk menentukan pilihan cabang dimana didapatkan :

a. Studi Kelayakan Bisnis:

- Cabang Kanan: Memiliki syarat bahwa studi kelayakan bisnis harus diselesaikan dengan nilai tidak lebih dari 3.
- Cabang Kiri: Menetapkan batasan nilai yang sama, yaitu studi kelayakan bisnis harus diselesaikan dengan nilai tidak lebih dari 3.

b. Manajemen Keuangan I:

- Cabang Kanan: Menetapkan bahwa mahasiswa harus mendapatkan nilai tidak lebih dari 4 dalam mata kuliah Manajemen Keuangan I.
- Cabang Kiri: Mengharuskan mahasiswa untuk mendapatkan nilai tidak lebih dari 3 dalam mata kuliah Manajemen Keuangan I.

c. Komunikasi Bisnis:

- Cabang Kanan: Menetapkan batasan nilai maksimal 2 untuk mata kuliah Komunikasi Bisnis.
- Cabang Kiri: Mengharuskan mahasiswa mendapatkan nilai tidak lebih dari 1 dalam mata kuliah Komunikasi Bisnis.

d. Akuntansi Manajemen:

- Cabang Kanan: Memiliki syarat bahwa nilai mata kuliah Akuntansi Manajemen tidak boleh lebih dari 2.
- Cabang Kiri: Menetapkan batasan nilai maksimal 1 untuk mata kuliah Akuntansi Manajemen.

2. Setelah menentukan calon cabang selanjutnya menentukan dan menghitung calon cabang mutakhir yaitu menghitung jumlah calon cabang kanan dan kiri berdasarkan jumlah dari data latih yang sudah ditentukan dari noktah keputusan.

Tabel 19. Jumlah Calon Cabang

Cabang	Cabang Kanan	Cabang Kiri
1	5	1
2	6	7
3	6	1
4	1	5

Tabel 19. Adapun penjelasan terkait tabel diatas adalah proses untuk memahami distrubusi data latih di setiap cabang dan bagaimana keputusan pada noktah keputusan memperngaruhi pembagian data :

3. Setelah mendapatkan jumlah dari calon cabang mutakhir selanjutnya menghitung candidate split purity left dan candidate split purity right dengan penilaian calon cabang kanan atau pun kiri di bagi dengan jumlah dari data latih.

Tabel 20. Nilai Candidate Split

Cabang	Cabang Kanan	Cabang Kiri
1	0,278	0,627
2	0,0	0,0
3	0,449	0,278
4	0,278	0,0

Tabel 20. Adapun penjelasan terkait tabel diatas merupakan proses penjumlahan candidate split purity left dan candidate split purity right dengan penilaian calon cabang kanan atau pun kiri di bagi dengan jumlah dari data latih.

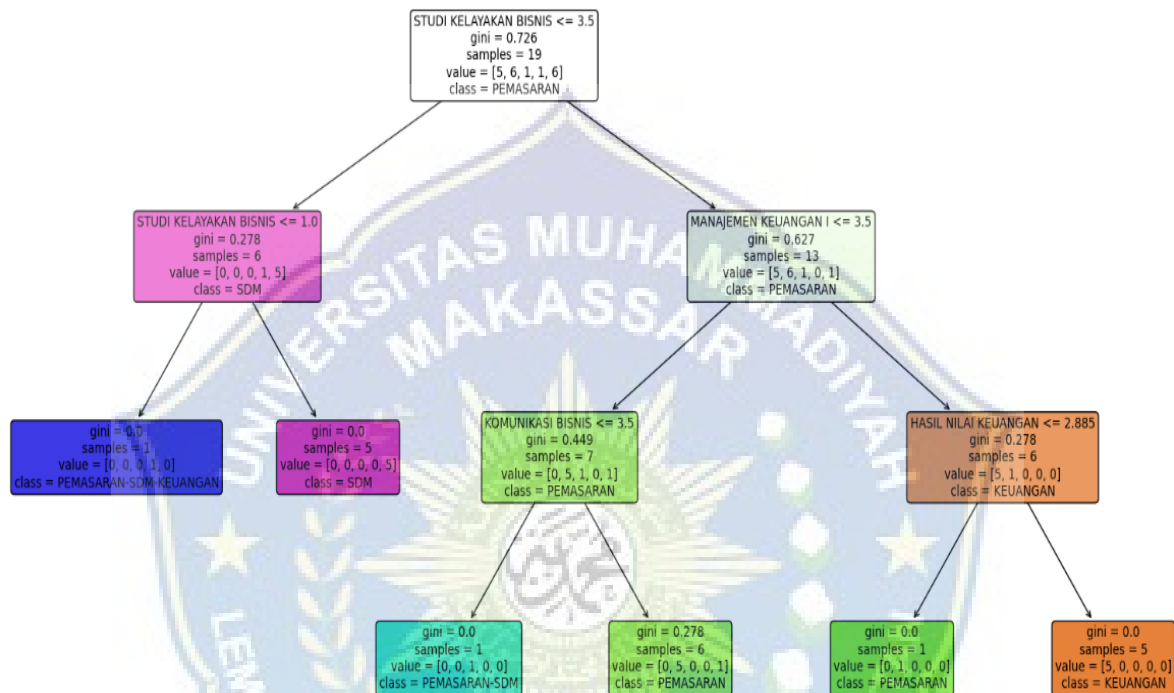
4. Hasil penentuan tersebut selanjutnya menghitung nilai $Q(S|t)$ yaitu dengan menghitung selisih nilai absolut dari $P J|tl$ dan $P J|tr$ dan kemudian dijumlahkan

Tabel 21. Perhitungan Nilai $Q(S|t)$

Gini Impurity for NAMA: 0.9473684210526316
Gini Impurity for AKUNTANSI MANAJEMEN: 0.6260387811634349
Gini Impurity for EKONOMI INTERNASIONAL: 0.3545706371191135
Gini Impurity for EKONOMI MANAJERIAL: 0.4986149584487536
Gini Impurity for KOMUNIKASI BISNIS: 0.48199445983379496
Gini Impurity for LEMBAGA KEUANGAN SYARIAH: 0.6315789473684211
Gini Impurity for MANAJEMEN KEUANGAN I: 0.7091412742382271
Gini Impurity for MANAJEMEN KEUANGAN II: 0.6094182825484765
Gini Impurity for MANAJEMEN OPERASIONAL I: 0.6260387811634349
Gini Impurity for MANAJEMEN OPERASIONAL II: 0.45983379501385035
Gini Impurity for MANAJEMEN PEMASARAN I: 0.5595567867036011
Gini Impurity for MANAJEMEN PEMASARAN II: 0.6094182825484765
Gini Impurity for MANAJEMEN SDM I: 0.5429362880886427
Gini Impurity for MANAJEMEN SDM II: 0.47645429362880887
Gini Impurity for PENGANGGARAN PERUSAHAAN: 0.5761772853185596
Gini Impurity for PEREKONOMIAN INDONESIA: 0.48199445983379496
Gini Impurity for RISET OPERASIONAL: 0.6094182825484765
Gini Impurity for SISTEM INFORMASI MANAJEMEN: 0.5429362880886427
Gini Impurity for STUDI KELAYAKAN BISNIS: 0.49307479224376727
Gini Impurity for TEORI PENGAMBILAN KEPUTUSAN: 0.5761772853185596
Gini Impurity for HASIL NILAI PEMASARAN: 0.8919667590027701
Gini Impurity for HASIL NILAI SDM: 0.8698060941828255
Gini Impurity for HASIL NILAI KEUANGAN: 0.9196675900277008

Tabel 21. Adapun penjelasan terkait tabel diatas adalah proses perhitungan Gini Impurity untuk setiap kategori atau mata kuliah yang disebutkan. Gini Impurity merupakan suatu metrik yang mengukur ketidakmurnian (impurity) dari suatu himpunan data. Semakin tinggi nilai Gini Impurity, semakin besar ketidakmurnian atau keberagaman dalam kategori tersebut.

5. Berdasarkan dari root node tersebut didapatkan pohon keputusannya dengan root node studi kelayakan bisnis ≤ 3.5 dengan gini awal yaitu 0,726 diikuti dengan cabang kanan manajemen keuangan ≤ 3.5 dan cabang kirinya studi kelayakan bisnis ≤ 1.0



Gambar 6. Pohon Keputusan CART

Tabel 20. Adapun penjelasan terkait tabel diatas adalah visualisasi pohon keputusan yang telah dilakukan perhitungan.

a. Root Node (Node Awal):

- Root node adalah titik awal pohon keputusan.
- Pada contoh Anda, kondisi awal adalah "Studi Kelayakan Bisnis ≤ 3.5 " dengan tingkat kecampuran (Gini Impurity) sebesar 0.726.

b. Branches (Cabang):

- Setiap cabang menggambarkan kemungkinan hasil berdasarkan kondisi tertentu.
- Pada cabang kiri, jika "Studi Kelayakan Bisnis" kurang dari atau sama dengan 3.5, kita masuk ke cabang ini.
- Pada cabang kanan, jika "Studi Kelayakan Bisnis" lebih dari 3.5, kita masuk ke cabang ini.

c. Leaf Nodes (Node Akhir):

- Leaf nodes adalah node akhir yang memberikan hasil atau keputusan.
- Pada cabang kiri, kita mungkin memiliki keputusan tertentu (hasil klasifikasi).
- Pada cabang kanan, kita mungkin memiliki keputusan lain.

d. Cabang-Cabang Tambahan:

- Pohon keputusan dapat memiliki cabang-cabang tambahan untuk membuat keputusan yang lebih rinci.
- Contohnya, pada cabang kiri, kita mungkin membuat keputusan berdasarkan kondisi "Studi Kelayakan Bisnis ≤ 1.0 ".

F. Proses Machine Learning Perhitungan Algoritma CART

Proses Machine Learning dalam perhitungan Dengan algoritma CART, atribut terkuat dipilih sebagai akar pohon keputusan. Melalui perhitungan ajaib, data dibagi menjadi cabang berdasarkan ambang batas yang optimal. Proses ini berulang hingga terbentuk pohon keputusan yang elegan setelah pemangkasan cabang yang kurang penting.

- Langkah pertama adalah Mengidentifikasi kelas unik dan menghitung frekuensinya dalam himpunan data dan Menghitung probabilitas masing-masing kelas selanjutnya Menghitung tingkat ketidakmurnian Gini menggunakan rumus $1 - \sum(p^2)$.

```
def GiniImpurity(y):  
    classes, counts = np.unique(y, return_counts=True)  
    probabilitas = counts / len(y)  
    gini = 1 - np.sum(probabilitas**2)  
    return gini
```

- Selanjutnya membagi data menjadi dua berdasarkan suatu fitur dan nilai ambang tertentu. Ini dilakukan dengan membuat dua masker: satu untuk data yang kurang dari atau sama dengan ambang, dan satu lagi untuk data yang lebih besar dari ambang. Fungsi mengembalikan empat himpunan data baru yang masing-masing sesuai dengan kondisi tersebut.

```
def MembagiData(X,y, feature, threshold):  
    left_mask = X[feature] <= threshold  
    right_mask = -left_mask  
    return X[left_mask], y[left_mask], X[right_mask], y[right_mask]
```

- Selanjutnya Fungsi Temukan Pembagian Terbaik mencari pembagian optimal untuk membangun pohon keputusan dengan iterasi melalui setiap

fitur dan nilai ambang, serta menghitung Gini pondered untuk memilih kombinasi yang menghasilkan Gini terendah.

```
def TemukanPembagianTerbaik(X, y):
    m, n = X.shape
    gini_terbaik = float('inf')
    fitur_terbaik = None
    threshold_terbaik = None

    for feature in X.columns:
        unique_values = np.unique(X[feature])
        for threshold in unique_values:
            X_left, y_left, X_right, y_right = MembagiData(X, y, feature, threshold)
            gini_left = GiniImpurity(y_left)
            gini_right = GiniImpurity(y_right)
            bobot_gini = (len(y_left) / m) * gini_left + (len(y_right) / m) * gini_right

            if bobot_gini < gini_terbaik:
                gini_terbaik = bobot_gini
                fitur_terbaik = feature
                threshold_terbaik = threshold

    return gini_terbaik, fitur_terbaik, threshold_terbaik
```

- Selanjutnya Fungsi Bangun Pohon Keputusan membangun pohon keputusan dengan rekursif, memeriksa batasan kedalaman maksimum dan homogenitas data, serta memilih fitur dan ambang terbaik untuk membagi data, kemudian membuat dua subpohon untuk bagian kiri dan kanan dari pembagian tersebut.

```
def BangunPohonKeputusan(X, y, depth=0, max_depth=None):
    if depth == max_depth or len(np.unique(y)) == 1:
        return {'class': np.argmax(np.bincount(y))}

    gini_terbaik, fitur_terbaik, threshold_terbaik = TemukanPembagianTerbaik(X, y)

    if fitur_terbaik is None:
        return {'class': np.argmax(np.bincount(y))}

    X_left, y_left, X_right, y_right = MembagiData(X, y, fitur_terbaik, threshold_terbaik)

    subpohon_kiri = BangunPohonKeputusan(X_left, y_left, depth + 1, max_depth)
    subpohon_kanan = BangunPohonKeputusan(X_right, y_right, depth + 1, max_depth)

    return {'feature': fitur_terbaik, 'threshold': threshold_terbaik,
            'left': subpohon_kiri, 'right': subpohon_kanan}
```

- Selanjutnya Dalam kode tersebut, LabelEncoder digunakan untuk mengkodekan variabel target (y) dari dataframe df yang memiliki kolom 'KONSENTRASI MAHASISWA'. Kemudian, hasil transformasi disimpan dalam variabel y_encoded. Selanjutnya, fitur (X) diambil dari df dengan menghapus kolom 'NAMA' dan 'KONSENTRASI MAHASISWA', dan disimpan dalam variabel X.


```

label_encoder = LabelEncoder()
y_encoded = label_encoder.fit_transform(df['KONSENTRASI MAHASISWA'])

# Memisahkan fitur (X) dan target yang telah diencode (y_encoded)
X = df.drop(['NAMA', 'KONSENTRASI MAHASISWA'], axis=1)
y = y_encoded

```

- Selanjutnya Dalam kode tersebut, `BangunPohonKeputusan` digunakan untuk membuat pohon keputusan dari data fitur (X) dan target (y) dengan kedalaman maksimum 3. Selanjutnya, fungsi `print_decision_tree` digunakan untuk mencetak struktur pohon keputusan. Pada akhirnya, pohon keputusan hasilnya dicetak dengan memanggil fungsi tersebut pada pohon keputusan yang telah dibangun (`decision_tree`).

```

def convert_tree_structure(tree, X):
    if 'class' in tree:
        return tree['class']

    feature = tree['feature']
    threshold = tree['threshold']
    left = tree['left']
    right = tree['right']

    if isinstance(threshold, str):
        condition = f"{feature} == {threshold}"
    else:
        condition = f"{feature} <= {threshold}"

    return (condition, convert_tree_structure(left, X), convert_tree_structure(right, X))

converted_tree = convert_tree_structure(decision_tree, X)

clf = DecisionTreeClassifier(max_depth=3)
clf.fit(X, y)

from sklearn.tree import plot_tree
import matplotlib.pyplot as plt

plt.figure(figsize=(20, 10))
plot_tree(clf, feature_names=X.columns, class_names=label_encoder.classes_, filled=True, rounded=True, fontsize=10)
plt.show()

```

- Proses Manual Calculation Algoritma CART
- Dalam kode tersebut, sebuah dataframe (df) dibaca dari file Excel yang berlokasi di `/content/testing.xlsx`. Kolom target y diambil dari kolom 'KONSENTRASI MAHASISWA' pada dataframe. Selanjutnya, nilai ketidakmurnian Gini (`gini_value`) dari target tersebut dihitung menggunakan fungsi `GiniImpurity`, dan hasilnya dicetak.

```

df = pd.read_excel('/content/testing.xlsx')

y_column = 'KONSENTRASI MAHASISWA'
y = df[y_column]

gini_value = GiniImpurity(y)

print(f"Gini Impurity for {y_column}: {gini_value}")

Gini Impurity for KONSENTRASI MAHASISWA: 0.7257617728531855

```

- Selanjutnya kode tersebut, sebuah dataframe (df) dibaca dari file Excel yang berlokasi di '/content/testing.xlsx'. Kemudian, untuk setiap kolom pada dataframe (kecuali kolom target 'KONSENTRASI MAHASISWA'), nilai ketidakmurnian Gini (gini_value) dihitung menggunakan fungsi Gini Impurity. Setiap hasil Gini Impurity dicetak bersama dengan nama kolomnya. Ini memberikan pemahaman tentang tingkat ketidakmurnian Gini untuk setiap fitur dalam dataframe.

```
df = pd.read_excel('/content/testing.xlsx')

target_column = 'KONSENTRASI MAHASISWA'

# Menampilkan Gini Impurity untuk setiap kolom
for column in df.columns:
    if column != target_column:
        y = df[column]
        gini_value = GiniImpurity(y)
        print(f"Gini Impurity for {column}: {gini_value}")
```

- Selanjutnya kode tersebut, sebuah dataframe (df) dibaca dari file Excel yang berlokasi di '/content/testing.xlsx'. Fitur (X) diambil dari dataframe dengan menghapus kolom 'NAMA' dan 'KONSENTRASI MAHASISWA', sementara target (y) diambil dari kolom 'KONSENTRASI MAHASISWA'. Selanjutnya, fungsi TemukanPembagianTerbaik digunakan untuk mencari pembagian terbaik berdasarkan tingkat ketidakmurnian Gini. Hasilnya, nilai Gini terbaik (gini_best), fitur terbaik (feature_best), dan ambang terbaik (threshold_best) dicetak.

```
df = pd.read_excel('/content/testing.xlsx')

X = df.drop(['NAMA', 'KONSENTRASI MAHASISWA'], axis=1)
y = df['KONSENTRASI MAHASISWA']

gini_best, feature_best, threshold_best = TemukanPembagianTerbaik(X, y)

print(f"Best Gini Impurity: {gini_best}")
print(f"Best Feature: {feature_best}")
print(f"Best Threshold: {threshold_best}")

Best Gini Impurity: 0.5168690958164642
Best Feature: STUDI KELAYAKAN BISNIS
Best Threshold: 3
```

- Selanjutnya Dalam kode tersebut, akurasi pada data pelatihan dihitung menggunakan fungsi calculate_accuracy dengan menerapkan pohon keputusan

yang telah dibangun (decision_tree) pada fitur (X) dan target (y). Hasil akurasi kemudian dicetak. Ini memberikan gambaran tentang sejauh mana pohon keputusan cocok dengan data yang digunakan untuk melatihnya.

```
# Menghitung kesesuaian pada data yang sama digunakan untuk membangun pohon
accuracy = calculate_accuracy(decision_tree, X, y)
print(f"Accuracy on the training data: {accuracy}")
```



G. Proses Machine Learning Decision Tree

Proses Machine Learning menggunakan Decision Tree melibatkan serangkaian langkah untuk membuat, melatih, dan mengevaluasi model Decision Tree. Berikut adalah penjelasan umum tentang tahapan proses tersebut:

- Berikut adalah penjelasan Baris kode `df=pd.read_excel("Training-M.xlsx", sheet_name='Sheet1')` digunakan untuk mengimpor dataset dari file Excel ke dalam Pandas DataFrame.

```
#Import manajemen dataset
df=pd.read_excel("Training-M.xlsx", sheet_name = 'Sheet1')
```

- Selanjutnya Baris kode `df.describe()` digunakan untuk menghasilkan statistik deskriptif dari DataFrame `df`. Saat metode `describe()` dipanggil pada DataFrame, itu akan memberikan ringkasan statistik dari data yang terdapat dalam DataFrame, Jika Anda menjalankan `df.describe()`, Anda akan melihat tabel dengan statistik ini untuk setiap kolom numerik dalam DataFrame.

```
df.describe()
```

	AKUNTANSI MANAJEMEN	EKONOMI INTERNASIONAL	EKONOMI MANAJERIAL	KOMUNIKASI BISNIS	LEMBAGA KEUANGAN SYARIAH	MANAJEMEN KEUANGAN I	MANAJEMEN KEUANGAN II	MANAJEMEN OPERASIONAL I
count	1554.000000	1554.000000	1554.000000	1554.000000	1554.000000	1554.000000	1554.000000	1554.000000
mean	3.604891	3.636583	3.681950	3.541828	3.512870	3.603604	3.574807	3.608430
std	0.716898	0.728445	0.653691	0.714031	0.806742	0.757903	0.792853	0.735771
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	3.500000	3.500000	3.750000	3.000000	3.000000	3.000000	3.000000	3.000000
50%	4.000000	4.000000	4.000000	4.000000	4.000000	4.000000	4.000000	4.000000
75%	4.000000	4.000000	4.000000	4.000000	4.000000	4.000000	4.000000	4.000000
max	7.000000	8.000000	6.750000	4.000000	6.000000	7.000000	4.000000	6.000000

- Selanjutnya Baris kode df Presentasi dataset terdapat dalam bentuk DataFrame yang memudahkan penampilan dan manipulasi data.

	NAMA MHS ANG 2018	AKUNTANSI MANAJEMEN	EKONOMI INTERNASIONAL	EKONOMI MANAJERIAL	KOMUNIKASI BISNIS	LEMBAGA KEUANGAN SYARIAH	MANAJEMEN KEUANGAN I	MANAJEMEN KEUANGAN II	MANAJEMEN OPERASIONAL I	CONTOH
0	A. ALDI MAPPIONANG	4.00	4.00	4.0	4	4.00	4.00	4.0	4.0	
1	A. HAERUL EKAPUTRA	3.00	4.00	3.0	4	4.00	4.00	4.0	4.0	
2	A. HASMA TENRILALA	4.00	4.00	4.0	3	4.00	4.00	3.0	3.0	
3	A. KHAERIL	3.00	3.00	3.0	4	2.00	3.00	3.0	3.0	
4	A. MILASARI	4.00	4.00	4.0	4	4.00	4.00	3.0	4.0	

- Selanjutnya Baris kode Pengaturan Rentang Kolom start_col dan end_col menentukan rentang kolom yang akan digunakan. Selanjutnya Pemilihan Subset Data dimana x menyimpan fitur (kolom 1 hingga 19) dari DataFrame y menyimpan target (kolom 20) dari DataFrame. Dengan kata lain, kita mendeklarasikan dan mengalokasikan dataset ke dalam dua variabel, yaitu x dan y. Variabel x berfungsi sebagai dataset inputan, merepresentasikan nilai-nilai dari berbagai kolom mata kuliah. Sementara itu, variabel y berperan sebagai dataset target atau label, mencakup informasi dari kolom Konsentrasi.

```

start_col = 1
end_col = 19

x= df.iloc[:, start_col:end_col+1]
y=df.iloc[:,20]

y

0    KEUANGAN
1    KEUANGAN
2    KEUANGAN
3    KEUANGAN
4    KEUANGAN
...
1052 KEUANGAN
1053 KEUANGAN
1054 KEUANGAN
1055 KEUANGAN
1056 KEUANGAN
Name: KONSENTRASI MAHASISWA, Length: 1057, dtype: object

```

- Selanjutnya Baris kode x adalah dataset x inputan

AKUNTANSI MANAJEMEN	EKONOMI INTERNASIONAL	EKONOMI MANAJERIAL	KOMUNIKASI BISNIS	LEMBAGA KEUANGAN SYARIAH	MANAJEMEN KEUANGAN I	MANAJEMEN KEUANGAN II	MANAJEMEN OPERASIONAL I	MANAJEMEN OPERASIONAL II	MANAJEMEN PEMASARAN I	MANAJEMEN PEMASARAN II
4.00	4.0	4.00	4	4.0	4.0	4.0	4.0	4.00	4.00	3.00
3.00	4.0	3.00	4	4.0	4.0	4.0	4.0	4.00	4.00	3.00
4.00	4.0	4.00	3	4.0	4.0	3.0	3.0	4.00	3.00	4.00
3.00	3.0	3.00	4	2.0	3.0	3.0	3.0	4.00	3.00	3.00
4.00	4.0	4.00	4	4.0	4.0	3.0	4.0	4.00	4.00	4.00
...
3.75	4.0	4.00	3	4.0	3.0	4.0	4.0	4.00	4.00	4.00
4.00	4.0	4.00	4	4.0	3.5	4.0	4.0	3.50	4.00	3.50
3.50	4.0	3.75	4	4.0	4.0	4.0	4.0	3.75	3.75	4.00
4.00	3.5	4.00	4	3.0	4.0	3.0	3.5	3.50	4.00	3.75
4.00	4.0	4.00	4	3.5	3.0	4.0	4.0	3.00	4.00	4.00

- Baris kode ini menggunakan modul `train_test_split` dari pustaka `scikit-learn` (`sklearn`) untuk membagi dataset menjadi dua bagian: data pelatihan (`x_train` dan `y_train`) dan data pengujian (`x_test` dan `y_test`), Dengan kata lain proses `splitting` data secara `random` data `training` 90% dan data `testing` 10%.

```

from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test = train_test_split(x,y,random_state=0,test_size=0.1)
✓ 0.3s

```

- Baris kode tersebut terlibat dalam pembuatan dan pelatihan model `Decision Tree` untuk tugas `klasifikasi` menggunakan pustaka `scikit-learn` atau sebagai proses `training` dan `testing` algoritma `decision tree`. Langkah-langkahnya sebagai berikut:

1. `from sklearn.tree import DecisionTreeClassifier` adalah Mengimpor kelas `Decision Tree Classifier` dari pustaka `scikit-learn`, yang menyediakan implementasi algoritma `Decision Tree` untuk `klasifikasi`.

2. `clf = DecisionTreeClassifier()` adalah Membuat objek model Decision Tree yang disebut `clf` dengan parameter default. Objek ini akan digunakan untuk melatih model pada data pelatihan.
3. `clf.fit(x_train, y_train)` menggunakan metode `.fit()` untuk melatih model dengan dataset pelatihan (`x_train` untuk fitur dan `y_train` untuk target).

```
from sklearn.tree import DecisionTreeClassifier
clf=DecisionTreeClassifier()
clf.fit(x_train,y_train) #model training
```

```
DecisionTreeClassifier
DecisionTreeClassifier()
```

- Selanjutnya adalah tampilan hasil testing dari Decision Tree

```
print (x_test)
print (akurasi)
```

	AKUNTANSI MANAJEMEN	EKONOMI INTERNASIONAL	EKONOMI MANAJERIAL	\
418	4.0	4.0	4.00	
202	4.0	4.0	4.00	
681	4.0	4.0	3.75	
316	3.0	4.0	3.00	
267	4.0	4.0	3.00	
..	
483	3.5	3.0	3.75	
629	4.0	4.0	4.00	
192	3.0	2.0	3.00	
141	3.0	4.0	4.00	
77	3.0	3.0	3.00	
	KOMUNIKASI BISNIS	LEMBAGA KEUANGAN SYARIAH	MANAJEMEN KEUANGAN I	\
418	3	4.0	4.0	
202	4	4.0	4.0	
681	3	4.0	4.0	
316	4	4.0	3.0	
267	3	4.0	4.0	
..	
483	4	3.0	3.0	
629	4	4.0	4.0	
192	0	0.0	0.0	
141	4	2.0	3.0	
77	4	1.0	3.0	

- Selanjutnya, kita mengambil dataset tambahan yang terdiri dari 93 data untuk melakukan pengujian. Dataset ini diambil dari file Excel. Proses pengujian bertujuan untuk mengevaluasi sejauh mana model Decision Tree yang telah dilatih dapat menggeneralisasi dan melakukan prediksi dengan akurat pada data yang belum pernah dilihat sebelumnya. Dengan menggunakan data pengujian ini, kita dapat mengukur performa model secara objektif sebelum menerapkannya pada situasi dunia nyata

```
#Import manajemen dataset untuk testing
#df=pd.read_excel("Training-N.xlsx", sheet_name = 'Sheet2')
df=pd.read_excel("testing.xlsx", sheet_name = 'Sheet1')
```



```
df
```

	NAMA	AKUNTANSI MANAJEMEN	EKONOMI INTERNASIONAL	EKONOMI MANAJERIAL	KOMUNIKASI BISNIS	LEMBAGA KEUANGAN SYARIAH	MANAJEMEN KEUANGAN I	MANAJEMEN KEUANGAN II	MANAJEMEN OPERASIONAL I	MANAJEMEN OPERASIONAL II	...
0	A. SUSANTI	4.0	4.00	4	4	4	4.00	4	4	4.0	...
1	A.M.FITRAH RAMADHANA	2.0	2.00	3	4	2	4.00	0	4	3.0	...
2	ABDUL RAZAK N. UMAR	2.0	2.00	4	4	0	3.00	3	0	4.0	...
3	ABI DZAR AL GHIFARI	3.0	3.00	4	0	2	2.00	4	4	4.0	...
4	AENUL	4.0	4.00	3	4	0	0.00	3	2	4.0	...
5	AGUNG GUNAWAN PRASETIO	4.0	4.00	3	3	3	3.00	4	3	4.0	...

- Selanjutnya, kita menetapkan variabel x sebagai dataset input yang akan digunakan dalam proses pengujian. Dataset ini berperan sebagai input untuk menguji performa model atau algoritma pada data yang belum pernah dilihat sebelumnya. Dengan menggunakan variabel x, kita dapat melakukan evaluasi terhadap kemampuan model untuk melakukan prediksi secara akurat pada situasi yang baru. Proses pengujian ini membantu memastikan generalisasi model yang baik dan memahami sejauh mana model dapat diterapkan pada data dunia nyata dan tampilan datanya.

```
x = df.iloc[:, start_col:end_col+1]
nama = df.iloc[:, 0]
```



```
x
```

	AKUNTANSI MANAJEMEN	EKONOMI INTERNASIONAL	EKONOMI MANAJERIAL	KOMUNIKASI BISNIS	LEMBAGA KEUANGAN SYARIAH	MANAJEMEN KEUANGAN I	MANAJEMEN KEUANGAN II	MANAJEMEN OPERASIONAL I	MANAJEMEN OPERASIONAL II	MANAJEMEN PEMASARAN I
0	4.0	4.00	4	4	4	4.00	4	4	4.0	4
1	2.0	2.00	3	4	2	4.00	0	4	3.0	4
2	2.0	2.00	4	4	0	3.00	3	0	4.0	3
3	3.0	3.00	4	0	2	2.00	4	4	4.0	4
4	4.0	4.00	3	4	0	0.00	3	2	4.0	4
5	4.0	4.00	3	3	3	3.00	4	3	4.0	4
6	3.0	4.00	3	3	3	2.00	3	3	4.0	4
7	1.0	3.00	4	2	4	3.00	4	4	4.0	3
8	3.0	4.00	3	4	4	4.00	3	3	2.0	1
9	3.0	4.00	4	4	3	2.00	0	2	4.0	3
10	3.0	4.00	4	4	2	0.00	0	2	3.0	4
11	4.0	0.00	4	0	0	3.00	0	4	4.0	3
12	4.0	0.00	3	3	3	3.00	3	3	4.0	4

- Langkah berikutnya melibatkan pengujian model dengan menggunakan panggilan model Decision Tree yang telah dilatih, yang disimpan dalam variabel clf. Proses ini merupakan bagian dari evaluasi kinerja model pada data pengujian untuk mengukur seberapa baik model dapat membuat

prediksi pada situasi yang belum pernah dilihat sebelumnya yang kemudian ditampilkan.

```
y_predict = clf.predict(x)

nama=pd.DataFrame(nama)
nama=nama.reset_index(drop=True)
nama['ID']=nama.index+0
ey = pd.DataFrame(nama)

y_predict=pd.DataFrame(y_predict)
y_predict=y_predict.reset_index(drop=True)
y_predict['ID']=y_predict.index+0

ex = pd.DataFrame(y_predict)
```

```
ex.describe

<bound method NDFrame.describe of 0 ID
0 KEUANGAN 0
1 KEUANGAN 1
2 PEMASARAN 2
3 SDM 3
4 SDM 4
5 KEUANGAN 5
6 KEUANGAN 6
7 KEUANGAN 7
8 KEUANGAN 8
9 PEMASARAN 9
10 PEMASARAN 10
11 SDM 11
12 KEUANGAN 12
```

- Tahapan terakhir melibatkan pengambilan data nama dan menggabungkannya dengan hasil prediksi yang kemudian diekspor ke dalam file Excel dengan nama "hasil_konsentrasi.xls". Proses ini memungkinkan kita untuk menyimpan hasil prediksi bersama dengan informasi identifikasi atau nama yang terkait, memberikan keluaran yang lengkap dan dapat dibaca untuk analisis lebih lanjut atau pelaporan.

```
<bound method NDFrame.describe of 0 ID
0 KEUANGAN 0
1 KEUANGAN 1
2 KEUANGAN 2
3 KEUANGAN 3
4 KEUANGAN 4
..
..
88 KEUANGAN 88
89 KEUANGAN 89
90 KEUANGAN 90
91 KEUANGAN 91
92 KEUANGAN 92

[93 rows x 2 columns]>

result = pd.merge(ex,ey, on='ID', how='inner')

result.to_excel("nama_file_hasil_prediksi.xlsx", index=False)
```

1. Teknik Pengujian Sistem

Teknik pengujian sistem memainkan peran krusial untuk memastikan kualitas dan kehandalan sistem. Proses ini melibatkan serangkaian langkah yang dirancang untuk mengevaluasi berbagai aspek kinerja dan fungsionalitas sistem.

Untuk pengujian sistem, langkah tambahan yang perlu dilakukan adalah membandingkan hasil prediksi yang diperoleh secara manual dengan hasil yang dihasilkan oleh aplikasi. Proses perbandingan ini akan membantu mengidentifikasi sejauh mana kesesuaian antara keputusan yang diambil oleh model Decision Tree dan keputusan yang diambil secara manual. Dengan memperoleh jumlah keseluruhan keputusan yang cocok, kita dapat menghitung persentase akurasi sistem, yang memberikan gambaran tentang seberapa baik model dapat mereplikasi keputusan yang dilakukan oleh manusia. Melalui analisis persentase kesesuaian ini, kita dapat mengevaluasi kinerja sistem dengan lebih komprehensif dan mendapatkan wawasan yang lebih mendalam mengenai keandalan model pada tugas pengambilan keputusan dalam konteks konsentrasi manajemen.

Berikut adalah hasil perbandingan yang telah dilakukan berdasarkan data yang telah diolah dengan model Decision Tree dan keputusan yang diambil secara manual :

Tabel 22. Hasil Pengujian Sistem Secara Manual dan Decision Tree

NAMA	MANUAL	DECISION TREE
A. SUBHANATULLAH. AS	PEMASARAN	PEMASARAN
A. SUSANTI	KEUANGAN	KEUANGAN
A.M.FITRAH RAMADHANA	PEMASARAN	KEUANGAN
ABDUL RAZAK N. UMAR	PEMASARAN	PEMASARAN
ABI DZAR AL GHIFARI	SDM	SDM
AENUL	SDM	SDM
AGUNG GUNAWAN PRASETIO	SDM	SDM
AHMAD IRWANTO	SDM	SDM
AHMAD MUBARAK	SDM	SDM
AIYUB AL ANSHARI SYAIFUL	KEUANGAN	KEUANGAN
AKBAR SAHFAT	PEMASARAN	PEMASARAN
AL MUFAKHIR	PEMASARAN	PEMASARAN
ALAMIA ALAM	PEMASARAN-SDM	PEMASARAN-SDM
ALDA	KEUANGAN	KEUANGAN
ALDA SUKIRMAN	SDM	SDM
ALDI SISWANTO	PEMASARAN	PEMASARAN
ALFIRA HERMAN	KEUANGAN	KEUANGAN
ALFIYYAH FARHANAH	KEUANGAN	KEUANGAN
ALIF	PEMASARAN-SDM-KEUANGAN	PEMASARAN-SDM-KEUANGAN

Pada tabel 22. Ditemukan hasil perbandingan yang dilakukan berdasarkan data yang telah diolah menggunakan metode Decision Tree dan perhitungan yang dilakukan secara manual menyatakan bahwa Melalui analisis perbandingan, hasil dari pemrosesan data menggunakan metode Decision Tree dan perhitungan manual menunjukkan hasil sama. Grafik perbandingan ini didasarkan pada pengujian sistem menggunakan dataset testing yang terdiri dari 20 data mahasiswa. Hasil evaluasi menunjukkan bahwa tingkat akurasi dalam penentuan konsentrasi, baik yang dilakukan oleh metode Decision Tree maupun secara manual, Terdapat satu data yang memiliki perbedaan. Dengan kata lain, model Decision Tree mampu memberikan keputusan dengan tingkat akurasi yang setara dengan keputusan yang dihasilkan secara manual. Temuan ini mencerminkan konsistensi dan keandalan model dalam tugas pengambilan keputusan untuk menentukan konsentrasi mahasiswa dalam konteks manajemen.

BAB V

PENUTUP

A. KESIMPULAN

Adapun beberapa point untuk kesimpulan yang penelitian ini dapat rangkum yaitu sebagai berikut:

1. Perbandingan antara metode Decision Tree dan perhitungan manual menunjukkan kesesuaian yang tinggi dalam penentuan konsentrasi mahasiswa. Konsistensi ini memvalidasi kemampuan model untuk mengambil keputusan dengan tingkat akurasi yang sebanding dengan proses manual.
2. Model Decision Tree mampu menghasilkan tingkat akurasi yang memadai dalam memprediksi konsentrasi mahasiswa. Temuan ini menunjukkan keberhasilan model dalam menggeneralisasi pola dari data pelatihan ke data pengujian. Penggunaan metode
3. Decision Tree memberikan efisiensi dalam proses pengambilan keputusan tanpa mengorbankan tingkat akurasi. Konsistensi antara hasil prediksi model dan perhitungan manual menunjukkan bahwa metode ini dapat diandalkan secara konsisten.

B. SARAN

Berdasarkan hasil penelitian yang telah dilakukan, diperoleh beberapa saran yang dapat diajukan, meliputi:

1. Melakukan penelitian lebih lanjut dengan menguji performa algoritma lain dalam konteks penentuan konsentrasi mahasiswa. Perbandingan dengan algoritma lain dapat memberikan wawasan lebih mendalam tentang mana yang lebih efektif dalam kasus ini
2. Meneliti pengaruh jumlah data terhadap kinerja model Decision Tree. Menggunakan dataset yang lebih besar atau lebih kecil dapat memberikan pemahaman tentang sejauh mana model dapat diandalkan dalam mengatasi variasi jumlah data.

DAFTAR PUSTA

- Alfian, M. (2014). *REGULASI EMOSI PADA MAHASISWA SUKU JAWA, SUKU BANJAR, DAN SUKU BIMA*. 02(02), 263–275.
- Benri, M., Metisen, H., & Latipa, S. (2015). Analisis Clustering Menggunakan Metode K-Means Dalam Pengelompokkan Penjualan Produk Pada Swalayan Fadhila. *Jurnal Media Infotama*, 11(2), 110–118.
<https://core.ac.uk/download/pdf/287160954.pdf>
- Crystallography, X. D. (2016). *pengujian sistem dengan path testing*. 1–23.
- Devia, E. (2023). *Penerapan Decision Tree Dengan Algoritma C4 . 5 Untuk Menentukan Rekomendasi Kenaikan Jabatan Karyawan*. 3, 28–37.
- Guarango, P. M. (2022). No Title *הענינים לנגד שבאמת מה את לראות קשה הכי*, 8.5.2017, 2003–2005.
- Hafizhuddin, M. I. (2019). *Hubungan Antara Self Disclousure Melalui Status Wa Dan Kualitas Hidup Pada Mahasiswa Di Universitas Muhammadiyah Surabaya*. http://repository.um-surabaya.ac.id/id/eprint/3715%0Ahttp://repository.um-surabaya.ac.id/3715/3/BAB_II.pdf
- Haryati, S., Sudarsono, A., & Suryana, E. (2015). Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu). *Jurnal Media Infotama*, 11(2), 130–138.
- hidayat fahrul, D. (2023). *HUBUNGAN MOTIVASI BELAJAR DENGAN TINGKAT KONSENTRASI BELAJAR SISWA KELAS VIII PADA MATA PELAJARAN FIQIH DI MTs NEGERI 2 SURAKARTA TAHUN PELAJARAN 2022/2023*. 31–41.
- Maysaroh, P., Studi, P., Agama, P., Islam, F. A., & Riau, U. I. (2021). *Yayasan Lembaga Pendidikan Islam Universitas Islam Riau Fakultas Agama Islam Beljar Peserta Didik Dalam Pembelajaran Sejarah Kebudayaan Islam Di Smp Taruna Islam Tenayan Raya ,.*

- Muntihana, V., Informatika, J. T., Sains, F., & Teknologi, D. A. N. (2017). *Berbasis Web Dan Android Pada Klinik Gigi Lisda*.
- Ndaumanu, R. I. (2023). Pengujian Sistem Informasi Perpustakaan Berbasis Website dengan Basis Path Testing. *Justek : Jurnal Sains Dan Teknologi*, 6(1), 123. <https://doi.org/10.31764/justek.v6i1.13808>
- Niland, N., Pearce, A. P., Naumann, D. N., O'Reilly, D., Series, P. B., Sataloff, R. T., Johns, M. M., Kost, K. M., Orsini, R. J., Medicine, T., Kalkman, J. P., Sataloff, R. T., Johns, M. M., Kost, K. M., Maiti, Bidinger, Assistance, H., Mitigate, T. O., Eroukhmanoff, C., & Licina, D. (2020). Title. *Global Health*, 167(1), 1–5. <https://www.e-ir.info/2018/01/14/securitisation-theory-an-introduction/>
- Nurjoko, & Kurniawan, H. (2016). Aplikasi Data Mining Untuk Memprediksi Tingkat Kelulusan Mahasiswa Menggunakan Algoritma Apriori di IBI Darmajaya, Bandar Lampung. *Jurnal TIM Darmajaya*, 02(01), 79–93.
- Octaviani, W. (2020). Pengaruh Constructive Play Therapy Dalam Meningkatkan Konsentrasi Belajar Anak Attention Deficit Hyperactivity Disorder (Adhd) Di Yayasan Pembinaan Anak Cacat (YPAC) Palembang. *UIN Raden Fatah Palembang*, 9–29.
- Pangaribuan, J. J., & Barus, O. P. (n.d.). *EXTREME LEARNING MACHINE: PENERAPAN DAN APLIKASI*.
- Putri Ayu, M. (2021). *Perbandingan Metode Extreme Learning Machine dan. September*.
- Ridwan, R., Lubis, H., & Kustanto, P. (2020). Implementasi Algoritma Neural Network dalam Memprediksi Tingkat Kelulusan Mahasiswa. *Jurnal Media Informatika Budidarma*, 4(2), 286. <https://doi.org/10.30865/mib.v4i2.2035>
- Rizmayanti, A. I., Hidayati, N., Nugraha, F. S., & Gata, W. (2021). Penerapan Data Mining Untuk Memprediksi Kompetensi Siswa Menggunakan Metode Decission Tree (Studi Kasus Smk Multicomp Depok). *Swabumi*, 9(1), 9–18. <https://doi.org/10.31294/swabumi.v9i1.8363>
- Sari, L. E., & Hadikurniawati, W. (2020). Sistem Pendukung Keputusan Penentuan Skincare Untuk Kulit Wajah Menggunakan Metode Decision Tree

- (Studi Kasus : Kosmetik Wardah). *Proceeding SENDIU 2020*, 978–979.
- Studi, P., & Informatika, T. (2017). Analisis Pola Pemilihan Konsentrasi Ilmu Jurusan Sistem Informasi Di STMIK TIME. *Jurnal TIMES*, VI(2), 37–43.
- Sulaeman, S., Palilingan, V. R., & Liando, O. E. S. (2021). Pengembangan Sistem Pendukung Keputusan Pemilihan Konsentrasi Jurusan Menggunakan Teknik Naive Bayes. *Edutik : Jurnal Pendidikan Teknologi Informasi Dan Komunikasi*, 1(2), 209–219. <https://doi.org/10.53682/edutik.v1i2.2259>
- Tahir, M. A. (2019). Perancangan Aplikasi Data Mining Menggunakan Metode Decision Tree Untuk Analisa Pemberian Kredit Pada Bri Unit LalabataRilau. *Jurnal Ilmiah Sistem Informasi Dan Teknik Informatika*, 2(1), 1–10.
- Wulansari Fridayanthie, E. (2018). Perancangan Sistem Informasi Penjualan Peralatan Hiking Berbasis Desktop. *Jurnal Khatulistiwa Informatika*, 3(2), 144.



LAMPIRAN 1. Source Code :

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
import sklearn
warnings.simplefilter("ignore")

#Import manajemen dataset
df=pd.read_excel("Training-M.xlsx", sheet_name = 'Sheet1')

df.describe()

df

start_col = 1
end_col = 22

x= df.iloc[:, start_col:end_col+1]

start_col = 28
end_col = 28
y=df.iloc[:, start_col:end_col+1]

y
x

from sklearn.model_selection import train_test_split
x_train, x_test, y_train, y_test =
train_test_split(x,y,random_state=0,test_size=0.1)

from imblearn.over_sampling import SMOTE
smote = SMOTE(random_state=42)
x_resampled, y_resampled = smote.fit_resample(x, y)

from sklearn.tree import DecisionTreeClassifier
clf=DecisionTreeClassifier()
clf.fit(x_train,y_train)
```



```

y_predict = clf.predict(x_test)
akurasi = sklearn.metrics.accuracy_score(y_test, y_predict)

```

```

print (x_test)
print (akurasi)

```

	AKUNTANSI MANAJEMEN	EKONOMI INTERNASIONAL	EKONOMI MANAJERIAL	\
418	4.0	4.0	4.0	4.00
202	4.0	4.0	4.0	4.00
681	4.0	4.0	4.0	3.75
316	3.0	4.0	4.0	3.00
267	4.0	4.0	4.0	3.00
..
483	3.5	3.0	3.0	3.75
629	4.0	4.0	4.0	4.00
192	3.0	2.0	2.0	3.00
141	3.0	4.0	4.0	4.00
77	3.0	3.0	3.0	3.00

I \	KOMUNIKASI BISNIS	LEMBAGA KEUANGAN SYARIAH	MANAJEMEN KEUANGAN
418	3	4.0	
4.0			
202	4	4.0	
4.0			
681	3	4.0	
4.0			
316	4	4.0	
3.0			
267	3	4.0	
4.0			
..	
...			
483	4	3.0	
3.0			
629	4	4.0	
4.0			
192	0	0.0	
0.0			
141	4	2.0	
3.0			
77	4	1.0	
3.0			

```

...
77          4.00

[106 rows x 19 columns]
0.9150943396226415

#Import manajemen dataset untuk testing
#df=pd.read_excel("Training-M.xlsx", sheet_name = 'Sheet2')
df=pd.read_excel("testing.xlsx", sheet_name = 'Sheet1')

df
x = df.iloc[:, start_col:end_col+1]
nama = df.iloc[:, 0]

x
nama
0          A. SUSANTI
1    A.M.FITRAH RAMADHANA
2    ABDUL RAZAK N. UMAR
3    ABI DZAR AL GHIFARI
4          AENUL
5    AGUNG GUNAWAN PRASETIO
6          AHMAD IRWANTO
7          AHMAD MUBARAK
8    AIYUB AL ANSHARI SYAIFUL
9          AKBAR SAHFAT
10         ARI NUGRAHA
11    ASWANDI SADRIL
12    ASWARUDDIN
13         ASWIN
14         NUR HALIMAH
15    REZQI AMALIA
16    FIRMAN KARIM
17    FIRSAL PULDRIANSA
18    MUH. HAMZA DENI
Name: NAMA, dtype: object

y_predict = clf.predict(x)

nama=pd.DataFrame(nama)

nama=nama.reset_index(drop=True)
nama['ID']=nama.index+0

```

```
ey = pd.DataFrame(nama)

y_predict=pd.DataFrame(y_predict)
y_predict=y_predict.reset_index(drop=True)
y_predict['ID']=y_predict.index+0

ex = pd.DataFrame(y_predict)

ex.describe
```

```
<bound method NDFrame.describe of          0 ID
0    KEUANGAN    0
1    KEUANGAN    1
2    PEMASARAN    2
3         SDM    3
4         SDM    4
5    KEUANGAN    5
6    KEUANGAN    6
7    KEUANGAN    7
8    KEUANGAN    8
9    PEMASARAN    9
10   PEMASARAN   10
11         SDM   11
12   KEUANGAN   12
13         SDM   13
14   PEMASARAN   14
15   KEUANGAN   15
16         SDM   16
```

17 SDM 17

18 PEMASARAN 18>

ey.describe

<bound method NDFrame.describe of

NAMA ID

0 A. SUSANTI 0

1 A.M.FITRAH RAMADHANA 1

2 ABDUL RAZAK N. UMAR 2

3 ABI DZAR AL GHIFARI 3

4 AENUL 4

5 AGUNG GUNAWAN PRASETIO 5

6 AHMAD IRWANTO 6

7 AHMAD MUBARAK 7

8 AIYUB AL ANSHARI SYAIFUL 8

9 AKBAR SAHFAT 9

10 ARI NUGRAHA 10

11 ASWANDI SADRIL 11

12 ASWARUDDIN 12

13 ASWIN 13

14 NUR HALIMAH 14

15 REZQI AMALIA 15

16 FIRMAN KARIM 16

17 FIRSAL PULDRIANSA 17

18 MUH. HAMZA DENI 18>



```
result = pd.merge( ex,ey, on='ID', how='inner')
```

```
result.to_excel('hasil_konsentrasi.xlsx', index=False)
```



LAMPIRAN 2. Output Source Code :

	A	B	C
1	0	ID	NAMA
2	PEMASARAN	0	A. SUBHANATULLAH. AS
3	KEUANGAN	1	A. SUSANTI
4	KEUANGAN	2	A.M.FITRAH RAMADHANA
5	PEMASARAN	3	ABDUL RAZAK N. UMAR
6	SDM	4	ABI DZAR AL GHIFARI
7	SDM	5	AENUL
8	SDM	6	AGUNG GUNAWAN PRASETIO
9	SDM	7	AHMAD IRWANTO
10	SDM	8	AHMAD MUBARAK
11	KEUANGAN	9	AIYUB AL ANSHARI SYAIFUL
12	PEMASARAN	10	AKBAR SAHFAT
13	PEMASARAN	11	AL MUFAKHIR
14	PEMASARAN-SDM	12	ALAMIA ALAM
15	KEUANGAN	13	ALDA
16	SDM	14	ALDA SUKIRMAN
17	PEMASARAN	15	ALDI SISWANTO
18	KEUANGAN	16	ALFIRA HERMAN
19	KEUANGAN	17	ALFIYYAH FARHANAH
20	PEMASARAN-SDM-KEUANGAN	18	ALIF

Gambar 7. Output Source Code

LAMPIRAN 3. Surat Pengambilan Data :

UNIVERSITAS MUHAMMADIYAH MAKASSAR
FAKULTAS TEKNIK
PROGRAM STUDI INFORMATIKA
GEDUNG MENARA IQRA LT.3
Jl. Sultan Alauddin No. 259 Telp. (0411) 866 972 Fax (0411) 865 588 Makassar 90221
Website: www.unismuh.ac.id, e-mail: unismuh@gmail.com
Website: <http://teknik.unismuh.makassar.ac.id>

Kampus Merdeka
INDONESIA JAYA

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Nomor 842/05/C.4-VI/VIII/45/2023
Lamp. -
Hal Permintaan Data Penelitian

Makassar, 17 Muharram 1445 H
04 Agustus 2023 M

Kepada yang Terhormat,
Pimpinan Program Studi Manajemen Fakultas Ekonomi dan Bisnis Universitas Muhammadiyah Makassar
Di -
Tempat

Assalamu 'Alaikum Warahmatullahi Wabarakatuh

Dengan Rahmat Allah SWT, Semoga aktivitas kita bernilai ibadah di Sisi - Nya. Dalam rangka penyelesaian Tugas Sarjana / Tugas Akhir Mahasiswa pada Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar dengan judul: **"Penerapan Algoritma K-Means pada Penentuan Konsentrasi Mahasiswa Program Studi Manajemen Unismuh Makassar"** Sehubungan hal tersebut, maka kami meminta kesediaan Bapak/Ibu agar kiranya berkenan membantu perihal surat tersebut. Bersama ini kami sampaikan mahasiswa(i):

No.	Stambuk	Nama
1.	105 84 11031 18	Rifaldi Arifuddin

Demikian surat kami atas perhatian dan kerja samanya kami haturkan banyak terima kasih.
Jazakumullah Khaeran Katsran
Wassalamu 'Alaikum warahmatullah Wabarakatuh

Ketua Program Studi
Informatika
Muhammad A. M. Hayat, S.Kom., MT.

Tembusan: Kepada Yang Terhormat,
1 Dekan Fakultas Teknik
2 Arsip

Gambar 8. Surat Pengambilan Data



UNIVERSITAS MUHAMMADIYAH MAKASSAR
FAKULTAS TEKNIK
PROGRAM STUDI INFORMATIKA



Kampus
Merdeka
INDONESIA - JAYA

GEDUNG MENARA IKIRA LT.3
Jl. Sultan Alauddin No. 259 Telp. (0411) 866 972 Fax (0411) 865 588 Makassar 90221
Website: www.unismuh.ac.id, e-mail: unismuh@gmail.com
Website: <http://teknik.unismuh.makassar.ac.id>

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Nomor : 841/05/C.4-VI/VIII/45/2023
Lamp. :-
Hal : Permintaan Data Penelitian

Makassar, 17 Muharram 1445 H
04 Agustus 2023 M

Kepada yang Terhormat,
Pimpinan Simak Universitas Muhammadiyah Makassar
Di -
Tempat

Assalamu 'Alaikum Warahmatullahi Wabarakatuh

Dengan Rahmat Allah SWT, Semoga aktivitas kita bernilai ibadah di Sisi - Nya. Dalam rangka penyelesaian Tugas Sarjana / Tugas Akhir Mahasiswa pada Program Studi Informatika Fakultas Teknik Universitas Muhammadiyah Makassar dengan judul: "*Penerapan Algoritma K-Means pada Penentuan Konsentrasi Mahasiswa Program Studi Manajemen Unismuh Makassar*" Sehubungan hal tersebut, maka kami meminta kesediaan Bapak/Ibu agar kiranya berkenan membantu perihal surat tersebut. Bersama ini kami sampaikan mahasiswa(i):

No.	Stambuk	Nama
1.	105 84 11031 18	Rifaldi Arifuddin

Demikian surat kami atas perhatian dan kerja samanya kami haturkan banyak terima kasih.

Jazakumullah Khaeran Katsiran
Wassalamu 'Alaikum warahmatullah Wabarakatuh

Ketua Program Studi

Informatika

UNIVERSITAS MUHAMMADIYAH MAKASSAR

FAKULTAS TEKNIK

PRODI INFORMATIKA

A. M. Havat, S.Kom., MT.

Tembusan: Kepada Yang Terhormat,
1 Dekan Fakultas Teknik
2 Arsip

Gambar 9. Surat Pengambilan Data

LAMPIRAN 4. Surat Penelitian :

**MAJELIS PENDIDIKAN TINGGI PIMPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH MAKASSAR**
LEMBAGA PENELITIAN PENGEMBANGAN DAN PENGABDIAN KEPADA MASYARAKAT
Jl. Sultan Alauddin No. 259 Telp.866972 Fax (0411)865588 Makassar 90221 e-mail dp3m@unismuh.ac.id

Nomor : 2204/05/C.4-VIII/VIII/1444/2023 21 Muharram 1445 H
Lamp : 1 (satu) Rangkap Proposal 08 August 2023 M
Hal : Permohonan Izin Penelitian

Kepada Yth,
Ketua Lembaga Perpustakaan dan Penerbitan
Universitas Muhamamdiyah Makassar
di -
Makassar

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

Berdasarkan surat Dekan TEKNIK Universitas Muhammadiyah Makassar, nomor: 841/05/C.4-VI/VIII/45/2023 tanggal 4 Agustus 2023, menerangkan bahwa mahasiswa tersebut di bawah ini :

Nama : RIFALDI ARIFUDDIN
No. Stambuk : 10584 1103118
Fakultas : TEKNIK
Jurusan : INFORMATIKA
Pekerjaan : Mahasiswa
Bermaksud melaksanakan penelitian/pengumpulan data dalam rangka penulisan Skripsi dengan judul :

"PENERAPAN ALGORITMA K-MEANS PADA PENENTUAN KONSENTRASI MAHASISWA PROGRAM STUDI MANAJEMEN UNISMUH MAKASSAR"

Yang akan dilaksanakan dari tanggal 11 Agustus 2023 s/d 11 Oktober 2023.

Sehubungan dengan maksud di atas, kiranya Mahasiswa tersebut diberikan izin untuk melakukan penelitian sesuai ketentuan yang berlaku.
Demikian, atas perhatian dan kerjasamanya diucapkan Jazakumullahu khaeran

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

Ketua LP3M,

Dr. Muhsin Arief Muhsin, M.Pd
NBM 1127761

08-23

Gambar 10. Surat Penelitian

Nm



MAJELIS PENDIDIKAN TINGGI PIMPINAN PUSAT MUHAMMADIYAH
UNIVERSITAS MUHAMMADIYAH MAKASSAR
UPT PERPUSTAKAAN DAN PENERBITAN

Alamat kantor: Jl. Sultan Alauddin No.259 Makassar 90221 Tlp.(0411) 866972,881593, Fax.(0411) 863588

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

SURAT KETERANGAN BEBAS PLAGIAT

UPT Perpustakaan dan Penerbitan Universitas Muhammadiyah Makassar,
Menerangkan bahwa mahasiswa yang tersebut namanya di bawah ini:

Nama : Rifaldi, A

Nim : 105841103118

Program Studi : Teknik Informatika

Dengan nilai:

No	Bab	Nilai	Ambang Batas
1	Bab 1	9 %	10 %
2	Bab 2	23 %	25 %
3	Bab 3	10 %	10 %
4	Bab 4	5 %	10 %
5	Bab 5	5 %	5 %

Dinyatakan telah lulus cek plagiat yang diadakan oleh UPT- Perpustakaan dan Penerbitan Universitas Muhammadiyah Makassar Menggunakan Aplikasi Turnitin.

Demikian surat keterangan ini diberikan kepada yang bersangkutan untuk dipergunakan seperlunya.

Makassar, 09 Januari 2024
Mengetahui,

Kepala UPT: Perpustakaan dan Penerbitan,



Jl. Sultan Alauddin no 259 makassar 90222
Telepon (0411)866972,881 593, fax (0411)865 588
Website: www.library.unismuh.ac.id
E-mail : psrperustakaan@unismuh.ac.id

RIFALDI.A 105841103118 BAB I

by Tahap Tutup



Submission date: 09-Jan-2024 11:36AM (UTC+0700)

Submission ID: 2268233186

File name: TURNITIN_BAB_1_5.docx (240.51K)

Word count: 1537

Character count: 10324

RIFALDI.A 105841103118 BAB I

ORIGINALITY REPORT

27%
SIMILARITY INDEX

27%
INTERNET SOURCES

0%
PUBLICATIONS

14%
STUDENT PAPERS

PRIMARY SOURCES

1	repository.radenfatah.ac.id Internet Source	8%
2	eprints.radenfatah.ac.id Internet Source	7%
3	123dok.com Internet Source	5%
4	repository.uin-suska.ac.id Internet Source	4%
5	digilib.iain-jember.ac.id Internet Source	2%
6	digilibadmin.unismuh.ac.id Internet Source	2%

Exclude quotes On

Exclude matches < 2%

Exclude bibliography On

RIFALDI.A 105841103118 BAB II

by Tahap Tutup



Submission date: 09-Jan-2024 11:36AM (UTC+0700)
Submission ID: 2268233399
File name: TURNITIN_BAB_2_5.docx (1.47M)
Word count: 3769
Character count: 24301

RIFALDI.A 105841103118 BAB II

ORIGINALITY REPORT

17% SIMILARITY INDEX	20% INTERNET SOURCES	8% PUBLICATIONS	8% STUDENT PAPERS
--------------------------------	--------------------------------	---------------------------	-----------------------------

PRIMARY SOURCES

1	123dok.com Internet Source	4%
2	repository.uin-suska.ac.id Internet Source	3%
3	proceeding.unindra.ac.id Internet Source	2%
4	eprints.radenfatah.ac.id Internet Source	2%
5	kc.umn.ac.id Internet Source	2%
6	unismuh.ac.id Internet Source	2%
7	(12-19-13) http://222.124.203.59/files/disk1/474/jbptunikompp-gdl-berrymuham-23693-13-daftars-l.pdf Internet Source	2%

Exclude quotes

On

Exclude matches

< 2%

RIFALDI.A 105841103118 BAB

III

by Tahap Tutup

Submission date: 09-Jan-2024 11:37AM (UTC+0700)

Submission ID: 2268233689

File name: TURNITIN_BAB_3_4.docx (536.87K)

Word count: 1488

Character count: 9630

RIFALDI.A 105841103118 BAB III

ORIGINALITY REPORT

25%
SIMILARITY INDEX

25%
INTERNET SOURCES

5%
PUBLICATIONS

9%
STUDENT PAPERS

PRIMARY SOURCES

1	eprints.radenfatah.ac.id Internet Source	9%
2	etheses.uin-malang.ac.id Internet Source	5%
3	media.neliti.com Internet Source	4%
4	eprints.umm.ac.id Internet Source	3%
5	antarbangsa.ac.id Internet Source	2%
6	eprints.uns.ac.id Internet Source	2%

Exclude quotes On

Exclude matches < 2%

Exclude bibliography On

RIFALDI.A 105841103118 BAB

IV

by Tahap Tutup



Submission date: 09-Jan-2024 11:38AM (UTC+0700)

Submission ID: 2268233875

File name: TURNITIN_BAB_4_4.docx (2.01M)

Word count: 3610

Character count: 21454

RIFALDI.A 105841103118 BAB IV

ORIGINALITY REPORT

8%	8%	0%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	eprints.radenfatah.ac.id Internet Source	2%
2	conference.binadarma.ac.id Internet Source	2%
3	manajemen.unsil.ac.id Internet Source	2%
4	repository.radenfatah.ac.id Internet Source	1%
5	Submitted to Universitas Internasional Batam Student Paper	1%

Exclude quotes On

Exclude matches < 1%

Exclude bibliography On

RIFALDI.A 105841103118 BAB V

by Tahap Tutup



Submission date: 09-Jan-2024 11:39AM (UTC+0700)

Submission ID: 2268234088

File name: TURNITIN_BAB_5_4.docx (240.32K)

Word count: 1283

Character count: 8445

RIFALDI.A 105841103118 BAB V

ORIGINALITY REPORT

20%
SIMILARITY INDEX

20%
INTERNET SOURCES

4%
PUBLICATIONS

10%
STUDENT PAPERS

PRIMARY SOURCES

1	eprints.radenfatah.ac.id Internet Source	10%
2	repository.upnjatim.ac.id Internet Source	2%
3	jurnal.umpwr.ac.id Internet Source	2%
4	repository.umy.ac.id Internet Source	2%
5	etheses.uin-malang.ac.id Internet Source	2%
6	elibrary.bsi.ac.id Internet Source	2%

Exclude quotes On

Exclude matches < 2%

Exclude bibliography On