

**KLASIFIKASI MASA STUDI MAHASISWA MENGGUNAKAN  
ALGORITMA DECISION TREE ( Studi Kasus : Mahasiswa Program Studi  
Informatika UMM)**

**LAPORAN TUGAS AKHIR**

Diajukan Untuk Memenuhi Persyaratan Guna Meraih Gelar Sarjana Strata 1  
Teknik Informatika Universitas Muhammadiyah Malang



Achmad Fauzi

201710370311183

**Bidang Minat :**

Data Science

**JURUSAN TEKNIK INFORMATIKA**

**FAKULTAS TEKNIK**

**UNIVERSITAS MUHAMMADIYAH MALANG**

**TAHUN 2024**

# LEMBAR PERSETUJUAN

**KLASIFIKASI MASA STUDI MAHASISWA MENGGUNAKAN  
ALGORITMA DECISION TREE ( Studi Kasus : Mahasiswa Program  
Studi Informatika UMM)**



**TUGAS AKHIR**

**Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1**

**Informatika Universitas Muhammadiyah Malang**

Menyetujui,

Malang, 16 Juli 2024

Dosen Pembimbing 1

Dosen Pembimbing 2



**Didih Rizki Chandranegara S.kom.,**

**Vinna Rahmayanti S S.Si., M.Si**

**M.Kom**

**NIP. 180306071990PNS.**

**NIP. 180302101992PNS.**

**LEMBAR PENGESAHAN**  
**KLASIFIKASI MASA STUDI MAHASISWA MENGGUNAKAN**  
**ALGORITMA DECISION TREE ( Studi Kasus : Mahasiswa**  
**Program Studi Informatika UMM)**  
**TUGAS AKHIR**

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1  
Informatika Universitas Muhammadiyah Malang

Disusun Oleh :

**Achmad Fauzi**

**201710370311183**

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis penguji  
pada tanggal 16 Juli 2024

Menyetujui,

Dosen Penguji 1



**Christian Sri Kusuma Aditya S.Kom.,**

**M.Kom**

**NIP. 180327021991PNS.**

Dosen Penguji 2



**Setio Basuki MT., Ph.D.**

**NIP. 10809070477PNS.**

Mengetahui,  
Ketua Jurusan Informatika



**Ir. Galih Wasis Wicaksono S.kom. M.Cs.**

**NIP. 10814100541PNS.**

## LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

**NAMA : Achmad Fauzi**

**NIM : 201710370311183**

**FAK./JUR. : Informatika**

Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul “**KLASIFIKASI MASA STUDI MAHASISWA MENGGUNAKAN ALGORITMA DECISION TREE ( Studi Kasus : Mahasiswa Program Studi Informatika UMM)**” beserta seluruh isinya adalah karya saya sendiri dan bukan merupakan karya tulis orang lain, baik sebagian maupun seluruhnya, kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya ini, atau ada klaim dari pihak lain terhadap keaslian karya saya ini maka saya siap menanggung segala bentuk resiko/sanksi yang berlaku.

Mengetahui,  
Dosen Pembimbing



Didih Rizki Chandranegara S.kom.,  
M.Kom

Malang, 16 Juli 2024  
Yang Membuat Pernyataan



Achmad Fauzi

## ABSTRAK

Agar terwujud efisiensi perguruan tinggi maka pengelola perguruan tinggi dituntut untuk mengelola secara profesional. Salah satu faktor yang memengaruhi efisiensi adalah lama studi mahasiswa. Berdasarkan data akademik dari program studi Informatika Universitas Muhammadiyah Malang, banyak mahasiswa yang lama studinya mencapai lebih dari empat tahun. Lama studi menjadi salah satu faktor masalah pengelola program studi sehingga dapat memengaruhi kinerja akademik. Pada lingkungan akademik masih ditemukan permasalahan berdasarkan data - data mahasiswa lulus yang tidak dapat menuntaskan masa studinya dalam jangka waktu yang ditentukan. Hal tersebut dapat dimanfaatkan sebagai sumber informasi strategis bagi program studi untuk melakukan klasifikasi masa studi lulusan mahasiswa dengan menggunakan teknik data mining. Berdasarkan latar belakang tersebut, maka penelitian ini perlu dilakukan untuk menerapkan teknik data mining guna mengklasifikasi masa studi mahasiswa yaitu dengan menggunakan salah satu algoritma Decision Tree C4.5. Metode yang dilakukan berupa analisis data, preprocessing data, dan proses pengklasifikasian dengan Algoritma C4.5. Pada penelitian ini dilakukan 2 pengujian. Yang pertama yaitu pengujian klasifikasi algoritma C4.5 dengan hasil yang didapat bahwa algoritma tersebut menghasilkan akurasi klasifikasi sebesar 82,05 % dengan root node atau simpul paling atas yaitu nilai IPK. Untuk pengujian kedua yaitu dilakukan perbandingan algoritma C4.5 dengan algoritma sebelumnya yaitu algoritma ID3 dimana algoritma ID3 menghasilkan akurasi sebesar 81 % dengan root node yang sama yaitu nilai IPK. Penelitian ini menunjukkan bahwa algoritma C4.5 merupakan algoritma paling efektif dan paling tepat digunakan dalam pengklasifikasian suatu data, dengan nilai akurasi ketepatan prediksi yang cukup tinggi di bandingkan dengan algoritma sebelumnya yaitu Algoritma ID3.

**Kata Kunci:** Klasifikasi, Masa studi, Algoritma Decision Tree, Algoritma C4.5, Algoritma ID3.

## ABSTRACT

In order to realize higher education efficiency, university managers are required to manage it professionally. One factor that influences efficiency is the student's length of study. Based on academic data from the Informatics study program at the University of Muhammadiyah Malang, many students have studied for more than four years. Length of study is one of the problematic factors for study program managers so that it can influence academic performance. In the academic environment, problems are still found based on data on graduating students who are unable to complete their studies within the specified time period. This can be used as a source of strategic information for study programs to classify the study period of graduate students using data mining techniques. Based on this background, this research needs to be carried out to apply data mining techniques to classify student study periods, namely by using one of the C4.5 Decision Tree algorithms. The method used is data analysis, data preprocessing, and classification process using the C4.5 Algorithm. In this research, 2 tests were carried out. The first is testing the C4.5 algorithm classification with the results obtained that the algorithm produces a classification accuracy of 82.05% with the root node or top node, namely the GPA value. For the second test, a comparison of the C4.5 algorithm was carried out with the previous algorithm, namely the ID3 algorithm, where the ID3 algorithm produced an accuracy of 81% with the same root node, namely the GPA value. This research shows that the C4.5 algorithm is the most effective and most appropriate algorithm to use in classifying data, with a fairly high prediction accuracy value compared to the previous algorithm, namely the ID3 Algorithm.

**Keywords:** Classification, Study period, Decision Tree Algorithm, C4.5 Algorithm, ID3 Algorithm.



## LEMBAR PERSEMBAHAN

Puji Syukur penulis panjatkan kepada Allah SWT yang telah memberikan kesehatan, rahmat dan hidayah-Nya sehingga penulis masih diberikan kesempatan untuk menyelesaikan tugas akhir ini sebagai salah satu syarat untuk mendapatkan gelar kesarjanaan. Walaupun jauh dari kata sempurna, namun suatu kebanggaan tersendiri bisa menyelesaikan tugas akhir dengan judul “KLASIFIKASI MASA STUDI MAHASISWA MENGGUNAKAN ALGORITMA DECISION TREE ( Studi Kasus : Mahasiswa Program Studi Informatika UMM)” secara optimal tepat waktu. Akan tetapi terselesaikannya tugas akhir ini tidak lepas dari bantuan beberapa pihak yang telah mendukung dan menemani penulis selama pengerjaan skripsi berlangsung. Sehingga penulis menyampaikan banyak terimakasih kepada :

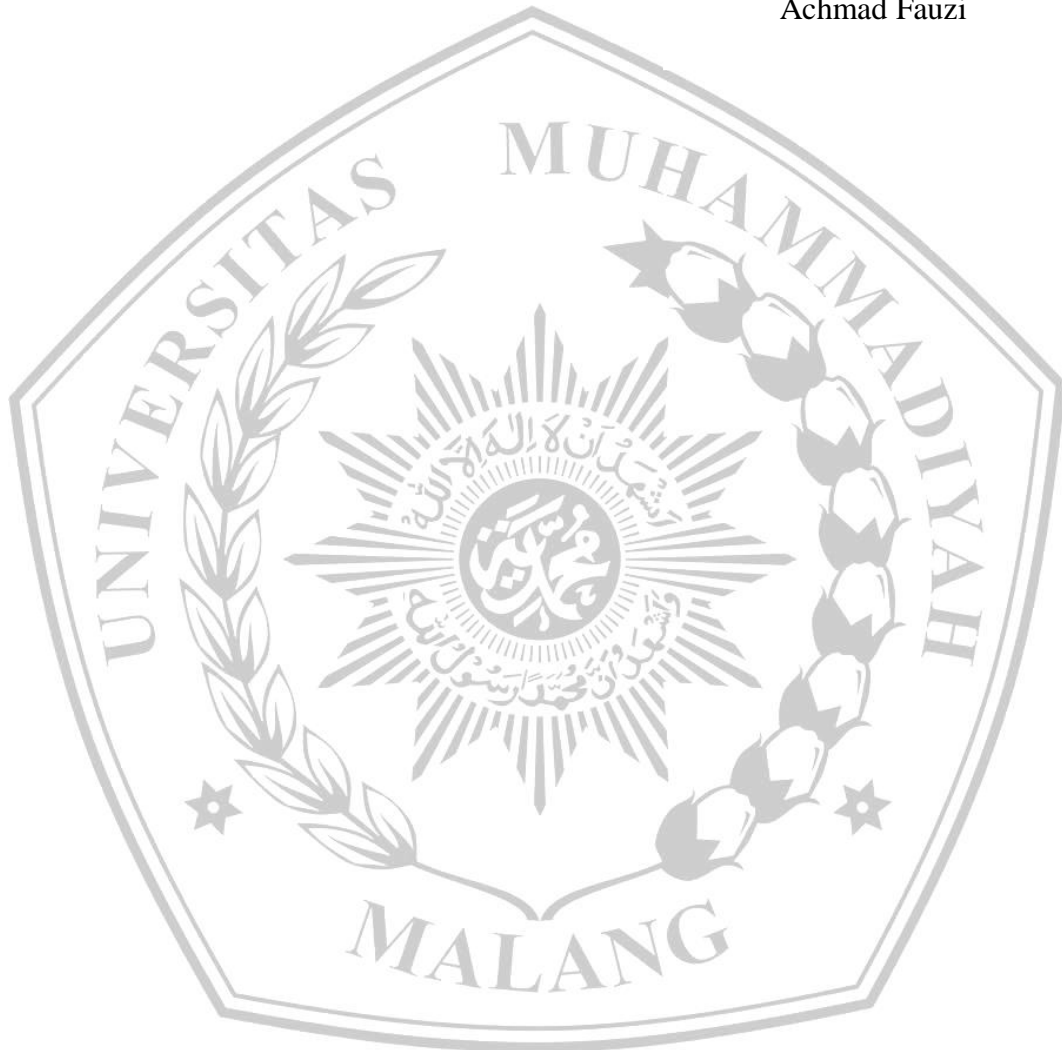
1. Allah SWT yang telah memberikan kesehatan, rahmat dan hidayah-Nya sehingga diberikan kemudahan dalam menyelesaikan tugas akhir ini.
2. Orang Tua saya, Bapak Jamili dan Ibu Andi Suriati serta saudara yang telah memotivasi dan memberikan dukungan secara moral dan materiil selama penulis mengerjakan tugas akhir ini.
3. Dosen pembimbing saya, Bapak Didih Rizki Chandranegara, S.Kom, M.Kom dan juga Ibu Vinna Rahmayanti S., S.Si, M.Si yang telah memberikan dukungan dan telah meluangkan waktu, tenaga, dan pikiran serta kesabarannya dalam membimbing penulis dalam mengerjakan tugas akhir.
4. Sahabat saya Dani, Rian, Lutfi, Arya, Enrico, Ucan, Koko, Wahyu, Ain, Yudi, Aldi, Walid, Naim, dan Tesar yang telah menemani saya dan memberikan dukungan dalam mengerjakan tugas akhir ini.
5. Teman – teman seperkontrakan saya Rian dan Lutfi yang telah memberikan dukungan dalam mengerjakan tugas akhir ini.
6. Teman – teman dari Informatika E 2017, yang telah memberikan dukungan serta motivasi dalam proses perkuliahan.
7. Dan pihak-pihak lain yang tidak dapat di sebutkan satu persatu.

Terimakasih atas segala dukungan dan motivasi dari berbagai pihak.  
Semoga tuhan selalu memberikan perlindungan.

Malang, 20 Februari 2024



Achmad Fauzi





## KATA PENGANTAR

Dengan memanjatkan puji syukur kepada Allah SWT atas limpahan rahmat hidayah dan karunia-Nya sehingga penulis dapat menyelesaikan tugas akhir sebagai persyaratan meraih gelar Sarjana Strata 1 dengan judul :

**“KLASIFIKASI MASA STUDI MAHASISWA MENGGUNAKAN ALGORITMA DECISION TREE ( Studi Kasus : Mahasiswa Program Studi Informatika UMM)”**

Di dalam tulisan ini disajikan pokok-pokok bahasan yang meliputi latar belakang, metode penelitian, serta hasil dan pembahasan yang telah di dapat pada proses penelitian ini. Diberikan kesimpulan berdasarkan hasil yang telah didapatkan pada proses penelitian ini.

Penulis sepenuhnya menyadari bahwa dalam penulisan tugas akhir ini masih banyak kekurangan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun untuk menyempurnakan penulisan tugas akhir ini. Sehingga, dapat bermanfaat bagi berbagai pihak dan juga bagi perkembangan ilmu pengetahuan terutama di bidang Informatika.

Malang, 20 Februari 2024



Achmad Fauzi

## DAFTAR ISI

LEMBAR PERSETUJUAN .....	i
LEMBAR PENGESAHAN .....	ii
LEMBAR PERNYATAAN .....	iii
ABSTRAK .....	iv
ABSTRACT .....	v
LEMBAR PERSEMBAHAN .....	vi
KATA PENGANTAR.....	viii
DAFTAR ISI .....	ix
DAFTAR GAMBAR.....	xi
DAFTAR TABEL.....	xii
BAB 1 .....	1
PENDAHULUAN .....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian.....	3
1.4 Batasan Masalah.....	4
BAB II.....	5
LANDASAN TEORI.....	5
2.1 Pengertian Mahasiswa .....	5
2.2 Data Mining .....	5
2.3 Klasifikasi .....	6
2.4 Decision Tree .....	6
2.4.1 Algoritma C4.5.....	7
2.4.2 Algoritma ID3 .....	7
BAB III.....	8
METODOLOGI PENELITIAN.....	8
3.1 Flowchart sistem .....	8
3.2 Analisis Data .....	9
3.3 Data Mining atau Preprocessing Data .....	10
3.4 Metode klasifikasi Algoritma Decision Tree .....	11

BAB IV .....	13
HASIL DAN PEMBAHASAN .....	13
4.1 Data <i>Preprocessing</i> .....	13
<b>4.2 Hasil pengujian klasifikasi Algoritma Decision Tree .....</b>	<b>18</b>
<b>4.2.1 Perhitungan manual dengan algoritma Decision Tree C4.5 .....</b>	<b>18</b>
4.2.2 Pengujian dengan algoritma Decision Tree C4.5 .....	21
4.2.3 Pengujian dan perbandingan algoritma ID3 dan C4.5 .....	25
<b>BAB V KESIMPULAN DAN SARAN .....</b>	<b>27</b>
<b>5.1 Kesimpulan .....</b>	<b>27</b>
<b>5.2 Saran .....</b>	<b>27</b>
DAFTAR PUSTAKA .....	29
LAMPIRAN .....	32



## DAFTAR GAMBAR

Gambar 3. 1 Flowchart .....	8
Gambar 4. 1 Data Cleaning .....	13
Gambar 4. 2 Penentuan Atribut .....	14
Gambar 4. 3 Penentuan Label .....	16
Gambar 4. 4 Data Training.....	17
Gambar 4. 5 Data Testing.....	17
Gambar 4. 6 Contoh Hasil Pohon Keputusan C4.5.....	21
Gambar 4. 7 Hasil Pohon Keputusan Algoritma C4.5 .....	22
Gambar 4. 8 Desain Proses Modelling .....	23
Gambar 4. 9 Hasil Proses Modelling .....	23
Gambar 4. 10 Desain Proses Evaluasi Performa Algoritma C4.5 .....	24
Gambar 4. 11 Hasil Evaluasi Performa Algoritma C4.5.....	25



## DAFTAR TABEL

Tabel 3. 1 .....	10
Tabel 4. 1 .....	14
Tabel 4. 2 .....	18
Tabel 4. 3 .....	19



## DAFTAR PUSTAKA

- [1] A. Fadli, M. I. Zulfa, and Y. Ramadhani, "Performance Comparison of Data Mining Classification Algorithms for Early Warning System of Students Graduation Timeliness," *J. Teknol. dan Sist. Komput.*, vol. 6, no. 4, pp. 158–163, 2018, doi: 10.14710/jtsiskom.6.4.2018.158-163.
- [2] K. Suhada, A. Elanda, and A. Aziz, "Klasifikasi Predikat Tingkat Kelulusan Mahasiswa Program Studi Teknik Informatika dengan Menggunakan Algoritma C4.5 (Studi Kasus: STMIK Rosma Karawang)," *Dirgamaya J. Manaj. dan Sist. Inf.*, vol. 1, no. 2, pp. 14–27, 2021, doi: 10.35969/dirgamaya.v1i2.182.
- [3] E. L. Hadisaputro, J. N. Utamajaya, P. Studi, and S. Informasi, "Performa Algoritma Data Mining Untuk Klasifikasi Data Perceraian," vol. 21, no. 04, pp. 14–17, 2022.
- [4] N. Indah Prabawati, Widodo, and H. Ajie, "Kinerja Algoritma Classification And Regression Tree (Cart) dalam Mengklasifikasikan Lama Masa Studi Mahasiswa yang Mengikuti Organisasi di Universitas Negeri Jakarta," *PINTER J. Pendidik. Tek. Inform. dan Komput.*, vol. 3, no. 2, pp. 139–145, 2019, doi: 10.21009/pinter.3.2.9.
- [5] S. Yunianita, S., Setiani, N., dan Mulyati, "Prediksi Ketepatan Masa Studi Mahasiswa dengan Algoritma Pohon Keputusan C45," *SNATi*, pp. 23–29, 2018.
- [6] P. Seminar and N. Ilmu, "Prosiding Seminar Nasional Ilmu Komputer Vol. 1, No 1, 2021," vol. 1, no. 1, pp. 220–225, 2021.
- [7] L. R. Haidar, E. Sedyono, and A. Iriani, "ANALISA PREDIKSI MAHASISWA DROP OUT MENGGUNAKAN METODE DECISION TREE DENGAN ALGORITMA ID3 dan C4.5," *J. Transform.*, vol. 17, no. 2, p. 97, 2020, doi: 10.26623/transformatika.v17i2.1609.
- [8] J. Kurniawati and S. Baroroh, "Literasi Media Digital Mahasiswa



Universitas Muhammadiyah Bengkulu,” *J. Komun.*, vol. 8, no. 2, pp. 51–66, 2016, [Online]. Available:

[https://r.search.yahoo.com/\\_ylt=AwrjbwLmv\\_NiIYcJ9B1XNyoA;\\_ylu=Y29sbwNncTEEcG9zAzEEdnRpZANEMTEyNV8xBHNIYwNzcg--/RV=2/RE=1660170343/RO=10/RU=https%3A%2F%2Fjournal.umy.ac.id%2Findex.php%2Fjkm%2Farticle%2Fview%2F2069/RK=2/RS=81QU2oK5sxo7ghZTIsrFj4EtGCI-](https://r.search.yahoo.com/_ylt=AwrjbwLmv_NiIYcJ9B1XNyoA;_ylu=Y29sbwNncTEEcG9zAzEEdnRpZANEMTEyNV8xBHNIYwNzcg--/RV=2/RE=1660170343/RO=10/RU=https%3A%2F%2Fjournal.umy.ac.id%2Findex.php%2Fjkm%2Farticle%2Fview%2F2069/RK=2/RS=81QU2oK5sxo7ghZTIsrFj4EtGCI-)

- [9] A. Romadhona, Suprapedi, and H. Himawan, “Prediksi Kelulusan Mahasiswa Tepat Waktu Berdasarkan Usia, Jenis Kelamin, Dan Indeks Prestasi Menggunakan Algoritma Decision Tree,” *J. Teknol. Inf.*, vol. 13, no. 1, pp. 69–83, 2017.
- [10] P. Kasih, “Pemodelan Data Mining Decision Tree Dengan Classification Error Untuk Seleksi Calon Anggota Tim Paduan Suara,” *Innov. Res. Informatics*, vol. 1, no. 2, pp. 63–69, 2019, doi: 10.37058/innovatics.v1i2.918.
- [11] Y. Irawan, “Penerapan Algoritma Decision Tree C4.5 Untuk Memprediksi Kelayakan Calon Pendorong Melakukan Donor Darah Dengan Klasifikasi Data Mining,” *JTIM J. Teknol. Inf. dan Multimed.*, vol. 2, no. 4, pp. 181–189, 2021, doi: 10.35746/jtim.v2i4.75.
- [12] A. Susanto, S. R. Riady, S. D. Ranti, and R. Mandala, “Penerapan Perhitungan Metode Decision Tree Menggunakan Algoritma Iterative Dichotomiser 3 (ID3) Berbasis Website Application of Decision Tree Method Calculation Using Website Based Iterative Dichotomiser 3 (ID3) Algorithm,” *Indones. J. Sci.*, vol. 1, no. 2, pp. 59–68, 2020, [Online]. Available: <http://journal.pusatsains.com/index.php/jsi>
- [13] A. Khaerunnisa, “Analisis Tingkat Kelulusan Mahasiswa di Unisba dengan menggunakan Algoritma K-Means Clustering,” *J. Ris. Mat.*, pp. 67–76, 2022, doi: 10.29313/jrm.v2i1.1018.
- [14] I. K. Swarjana, “Konsep Pengetahuan Sikap, Prilaku, Persepsi, Stres,

Kecemasan, Nyeri, Dukungan Sosial, Kepatuhan, Motivasi, Kepuasan, Pandemi Covid-19, Akses Layanan Kesehatan,” *Andi*, vol. 4, pp. 3–12, 2022.

- [15] E. Indrayuni, “Klasifikasi Text Mining Review Produk Kosmetik Untuk Teks Bahasa Indonesia Menggunakan Algoritma Naive Bayes,” *J. Khatulistiwa Inform.*, vol. 7, no. 1, pp. 29–36, 2019, doi: 10.31294/jki.v7i1.1.
- [16] I. C. Wibowo, A. C. Fauzan, M. D. P. Yustiana, and F. A. Qhabib, “Komparasi Algoritma Naive Bayes dan Decision Tree Untuk Memprediksi Lama Studi Mahasiswa,” *Ilk. J. Comput. Sci. Appl. Informatics*, vol. 1, no. 2, pp. 65–74, 2019, doi: 10.28926/ilkomnika.v1i2.21.
- [17] M. A. C, “Prediksi Keberlangsungan Studi Mahasiswa Fakultas Ilmu Pendidikan dan Sosial Universitas Nahdlatul Ulama Blitar Menggunakan Algoritma C4.5,” vol. 1, no. 2, pp. 110–119, 2021.



### FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

**Nama Mahasiswa** : Achmad Fauzi  
**NIM** : 201710370311183  
**Judul TA** : Klasifikasi Masa Studi Mahasiswa Menggunakan  
 Algoritma Decision Tree ( Studi Kasus ; Mahasiswa  
 Program Studi Informatika UMM )

#### Hasil Cek Plagiarisme dengan Turnitin

No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%) *
1.	Bab 1 – Pendahuluan	10 %	2%
2.	Bab 2 – Daftar Pustaka	25 %	8%
3.	Bab 3 – Analisis dan Perancangan	25 %	8%
4.	Bab 4 – Implementasi dan Pengujian	15 %	2%
5.	Bab 5 – Kesimpulan dan Saran	5 %	5%
6.	Makalah Tugas Akhir	20%	12%

*\*) Hasil cek plagiarism diisi oleh pemeriksa (staf TU)*

*\*) Maksimal 5 kali (4 Kali sebelum ujian, 1 kali sesudah ujian)*

Mengetahui,

Pemeriksa (Staff TU)

  
 (.....)