

PENERAPAN ALGORITMA *CLASSIFICATION AND REGRESSION TREES* PADA KLASIFIKASI KETEPATAN WAKTU LULUS MAHASISWA S1 FIP UNIVERSITAS NEGERI PADANG TAHUN 2022

TUGAS AKHIR

Diajukan sebagai salah satu persyaratan untuk memperoleh gelar Ahli Madya



Oleh

**SEPTIA YULIANA
NIM 19037072**

**PRODI DIPLOMA III STATISTIKA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI PADANG
2023**

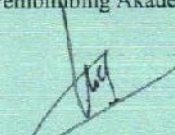
PERSETUJUAN UJIAN TUGAS AKHIR

**PENERAPAN ALGORITMA *CLASSIFICATION AND REGRESSION TREES*
PADA KLASIFIKASI KETEPATAN WAKTU LULUS MAHASISWA S1 FIP
UNIVERSITAS NEGERI PADANG TAHUN 2022**

Nama : Septia Yuliana
NIM/TahunMasuk : 19037072/2019
Program Studi : DIII Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Padang, 6 Februari 2023

Distetujui oleh:
Pembimbing Akademik


Admi Salma, S.Pd., M.Si

NIDN. 00251290003

HALAMAN PENGESAHAN LULUS UJIAN TUGAS AKHIR


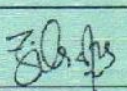

Nama : Septia Yuliana
NIM/TahunMasuk : 19037072/2019
Program Studi : DIII Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

**PENERAPAN ALGORITMA *CLASSIFICATION AND REGRESSION TREES*
PADA KLASIFIKASI KETEPATAN WAKTU LULUS MAHASISWA S1 FIP
UNIVERSITAS NEGERI PADANG TAHUN 2022**

**Dinyatakan lulus setelah dipertahankan di depan Tim Penguji Tugas Akhir
Program Studi DIII Statistika Departemen Statistika Fakultas Matematika dan
Ilmu Pengetahuan Alam Universitas Negeri Padang**

Padang, 6 Februari 2023

Tim Penguji

Nama	Tanda Tangan
1. Ketua : Admi Salma, S.Pd., M.Si	1. 
2. Anggota : Dodi Vionanda, M.Si., Ph.D	2. 
3. Anggota : Zilrahmi, S.Pd., M.Si	3. 

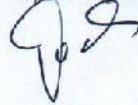
SURAT PERNYATAAN TIDAK PLAGIAT

Nama : Septia Yuliana
NIM/Tahun Masuk : 19037072/2019
Program Studi : DIII Statistika
Jurusan : Statistika
Fakultas : Matematika dan Ilmu Pengetahuan Alam

Dengan ini menyatakan bahwa, Tugas Akhir saya dengan judul **“Penerapan Algoritma Classification And Regression Trees Pada Klasifikasi Ketepatan Waktu Lulus Mahasiswa S1 FIP Universitas Negeri Padang Tahun 2022”** adalah benar merupakan hasil karya saya dan bukan merupakan plagiat dari karya orang lain atau pengutipan dengan cara-cara yang tidak sesuai dengan etika yang berlaku dalam tradisi keilmuan. Apabila suatu saat terbukti saya melakukan plagiat, maka saya bersedia diproses dan menerima sanksi akademis maupun hukum sesuai dengan hukum dan ketentuan yang berlaku, baik di instansi UNP maupun di masyarakat dan negara.

Demikian pernyataan ini saya buat dengan penuh kesadaran rasa tanggung jawab sebagai anggota masyarakat ilmiah.

Diketahui oleh,
Kepala Departemen Statistika



Dodi Vionanda, M.Si., Ph.D
NIP.19760611 200501 1 002

Saya yang menyatakan,



Septia Yuliana
NIM. 19037072

ABSTRAK

Septia Yuliana : Penerapan Algoritma *Classification and Regression Trees* Pada Klasifikasi Ketepatan Waktu Lulus Mahasiswa S1 FIP Universitas Negeri Padang Tahun 2022

Mahasiswa UNP masih mengalami masalah ketepatan waktu lulus, pada periode wisuda 126-127 Tahun 2022 mahasiswa yang lulus tepat waktu kurang dari 40% dari jumlah mahasiswa yang diwisuda pada periode tersebut. Hal ini dikarenakan beragamnya karakteristik yang dimiliki oleh mahasiswa selama menjalankan studi sehingga mempengaruhi lama waktu studi yang dibutuhkan untuk lulus. Penelitian ini bertujuan untuk mendapatkan model keputusan sebagai pencari ketepatan waktu lulus berdasarkan karakteristik yang dimiliki mahasiswa dengan teknik klasifikasi yaitu metode analisis pohon keputusan menggunakan algoritma *Classification And Regression Trees* (CART).

Penelitian ini berjenis penelitian terapan. Data yang digunakan adalah data sekunder yang diperoleh dari Biro Akademik dan Kemahasiswaan UNP Tahun 2022 yaitu data wisudawan FIP UNP periode wisuda 126-127. Pohon keputusan akan menghasilkan model keputusan dari klasifikasi ketepatan waktu lulus mahasiswa FIP UNP berdasarkan karakteristik dari mahasiswa yaitu IPK, jalur masuk perguruan tinggi dan jenis kelamin sebagai variabel prediktor (X) dan status ketepatan waktu lulus (Y) sebagai variabel respon.

Hasil penelitian menggunakan metode pohon keputusan optimal dengan algoritma CART menghasilkan 5 model keputusan dimana terdapat 3 model dengan keputusan tepat waktu dan 2 model dengan keputusan tidak tepat waktu dengan persentase ketepatan model sebesar 77,7%. Model keputusan mahasiswa yang tepat waktu apabila memiliki IPK di atas 3,64 atau memiliki IPK 3,61-3,64 jalur SBMPTN serta memiliki IPK 3,54 - 3,64 jalur masuk SNMPTN dan seleksi mandiri.

Kata Kunci: Ketepatan Waktu Lulus, Klasifikasi, Algoritma CART.

ABSTRACT

Septia Yuliana : Application Of The Classification And Regression Trees Algorithm to The Timely Classification Of Graduate FIP Students at Universitas Negeri Padang In 2022

During the 126-127 graduation period in 2022 UNP students still having problems graduating on time, students who graduated on time were less than 40%. The various characteristics possessed by students it affects the length of study time needed to graduate. This study using the Classification And Regression Trees (CART) algorithm to obtain a decision model as a marker of timeliness of graduation based on the characteristics possessed by students with a classification technique.

This research is an applied research type using the secondary data obtained from the UNP Academic and Student office in 2022, data on FIP graduates for the 126-127 graduation period. The CART method will produce a decision model from the classification of the punctuality of graduating FIP UNP students based on the characteristics of the students, namely GPA, college entrance and gender as the predictor variable (X) and the punctuality of graduation status (Y) as the response variable.

The results of the study using the CART algorithm produce 3 models with timely decisions and 2 models with untimely decisions with a percentage of model accuracy of 77.7%. The student decision model is timely if you have a GPA above 3.64 or have a GPA of 3.61-3.64 for the SBMPTN route and have a GPA of 3.54 - 3.64 for SNMPTN entry and independent selection.

Keywords: Timely of Graduate, Classification, CART Algorithm.

KATA PENGANTAR

Assalamu 'alaikum Wr. Wb

Puji dan syukur peneliti ucapkan kepada Allah SWT, yang telah melimpahkan rahmat dan karunia-Nya kepada peneliti, sehingga peneliti dapat menyelesaikan Tugas Akhir yang berjudul **“Penerapan Algoritma *Classification and Regression Trees* Pada Klasifikasi Ketepatan Waktu Lulus Mahasiswa S1 FIP Universitas Negeri Padang Tahun 2022”** dapat terselesaikan dengan baik.

Tugas akhir ini disusun untuk memenuhi salah satu persyaratan memperoleh gelar Ahli Madya pada Program Studi Diploma III Statistika Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang. Dalam penulisan Tugas Akhir ini, peneliti banyak mengalami kendala, namun berkat bantuan, bimbingan kerjasama dari berbagai pihak yang bersifat membangun, dan berkah Allah SWT sehingga kendala-kendala yang dihadapi dapat diatasi. Dalam kesempatan ini, peneliti mengucapkan terimakasih kepada:

1. Ibu Admi Salma, M.Si., Dosen Penasehat dan Pembimbing Tugas Akhir yang telah memberikan dorongan, arahan dan bimbingan dalam menyelesaikan Tugas Akhir ini.
2. Bapak Dodi Vionanda, M.Si., Ph.D., dosen penguji Tugas Akhir sekaligus Kepala Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang yang telah banyak memberikan masukan dan arahan demi kesempurnaan Tugas Akhir ini.

3. Ibu Zilrahmi, M.Si., dosen penguji Tugas Akhir yang telah banyak memberikan masukan dan arahan demi kesempurnaan Tugas Akhir ini.
4. Kedua Orang tua beserta keluarga yang selalu mendoakan, memberi semangat, nasihat dan materi dalam menyelesaikan Tugas Akhir ini.
5. Bapak dan Ibu Dosen Departemen Statistika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Padang.
6. Sahabat dan rekan-rekan Mahasiswa kelas A dan B Program Studi DIII Statistika angkatan 2019 khususnya Ayu, Anisa, Nadia dan Miftah.
7. Serta semua pihak yang telah membantu pembuatan tugas akhir ini yang tidak dapat peneliti sebutkan satu persatu.

Peneliti telah berusaha semaksimal mungkin untuk memberikan yang terbaik dalam penulisan Tugas akhir ini, namun peneliti mengharapkan saran dan kritik yang membangun demi kesempurnaan tugas akhir ini. Akhir kata, peneliti berharap agar tugas akhir ini dapat bermanfaat bagi semua pihak.

Padang, Februari 2023

Septia Yuliana

DAFTAR ISI

ABSTRAK.....	i
KATA PENGANTAR.....	iii
DAFTAR ISI	v
DAFTAR TABEL	vii
DAFTAR GAMBAR.....	viii
DAFTAR LAMPIRAN	ix
BAB I PENDAHULUAN.....	1
A. Latar Belakang	1
B. Batasan Masalah	4
C. Rumusan Masalah	5
D. Tujuan Masalah.....	5
E. Manfaat Penulisan	5
BAB II KERANGKA TEORITIS	6
A. Jalur Masuk Perguruan Tinggi.....	6
B. Indek Prestasi Kumulatif.....	8
C. Jenis Kelamin.....	8
D. Masa Studi	9
E. Data Mining	9
F. Klasifikasi	11
G. Decision Tree	13
H. CART (Classification and Regression Tree).....	15
I. Confusion Matrix	23
BAB III METODE PENELITIAN	25
A. Jenis penelitian.....	25
B. Jenis dan Sumber data	25
C. Variabel penelitian	25
D. Struktur Data.....	26
E. Teknik Analisis Data	26

BAB IV HASIL DAN PEMBAHASAN.....	28
A. Hasil Penelitian	28
B. Pembahasan	48
BAB V PENUTUP	51
A. Kesimpulan	51
B. Saran.....	51
DAFTAR PUSTAKA.....	53
LAMPIRAN	56

DAFTAR TABEL

Tabel	Halaman
1. Syarat Beban SKS dan IPK Minimal Kelulusan	8
2. Confusion Matrix	23
3. Variabel Dan Jenis Data	25
4. Struktur Data.....	26
5. Calon Simpul Variabel Jalur Masuk Perguruan Tinggi	36
6. Calon Simpul Terbaik Variabel IPK Mahasiswa.....	37
7. Pemilah Terbaik Pada Pembelahan $IPK \geq 3,54$	40
8. Pemilah Terbaik Pada Pembelahan $IPK < 3,54$	42
9. Complexity Parameter Simpul internal Pohon Maksimal.....	44
10. Nilai Complexity Parameter Pohon Keputusan CART Maksimal	45
11. Hasil Klasifikasi Pohon Optimal	47
12. Uji Ketepatan Klasifikasi	47

DAFTAR GAMBAR

Gambar	Halaman
1. Persentase Wisudawan UNP Periode 126-127.....	2
2. Kerangka Kerja Klasifikasi	12
3. Contoh Pemecahan Fitur Bertipe Nominal	14
4. Contoh Pemecahan Fitur Bertipe Ordinal	14
5. Contoh Pemecahan Fitur Bertipe Numerik	15
6. Contoh Decision Tree Menggunakan CART	22
7. Persentase Mahasiswa Berdasarkan Ketepatan Waktu Lulus	28
8. Persentase Ketepatan Waktu Lulus Wisudawan Berdasarkan Jenis Kelamin...	29
9. Persentase Ketepatan Waktu Lulus Wisudawan FIP Berdasarkan Jalur Masuk	30
10. Pohon CART Pembelahan Pertama	39
11. Pohon CART Pembelahan Kedua.....	41
12. Pohon CART Pembelahan Ketiga	42
13. Pohon Keputusan CART Maksimal.....	43
14. Plot Complexity parameter dan Cross Validation Error	46
15. Pohon Optimal	46

DAFTAR LAMPIRAN

Lampiran	Halaman
1. Data Penelitian.....	56
2. Proses import dan pre-proccesing data menggunakan R studio	57
3. Proses Splitting data latih dan data uji menggunakan R studio	58
4. Proses pembangunan Model pohon maksimal menggunakan R	59
5. Visualisasi Pohon Keputusan <i>CART</i> Maksimal.....	61
6. Tabel simpul terminal dan label kelas.....	62
7. Proses Pruning pohon maksimal.....	63
8. Visualisasi pohon pruning	64
9. Uji Ketepatan Model Pohon Optimal.....	65

BAB I

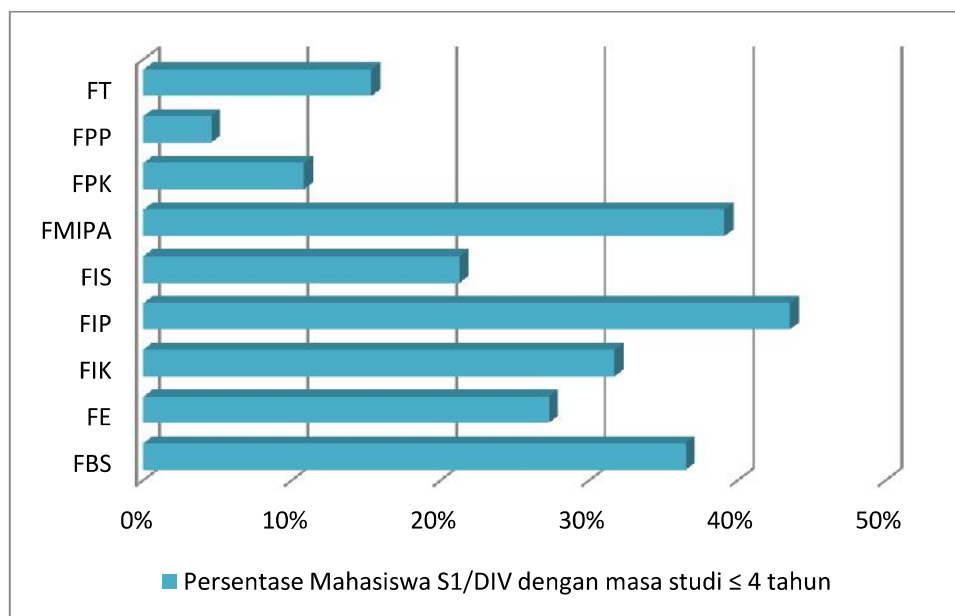
PENDAHULUAN

A. Latar Belakang

Perguruan tinggi merupakan lembaga yang menyelenggarakan pendidikan lanjutan yang ditempuh setelah pendidikan menengah. Perguruan Tinggi adalah satuan pendidikan yang menyelenggarakan Pendidikan Tinggi (UU No.12 tahun 2012 pasal 1 ayat 6). Tujuan pendidikan tinggi salah satunya adalah untuk mengembangkan potensi mahasiswa sebagai peserta didik dari perguruan tinggi sehingga dihasilkan lulusan yang menguasai cabang Ilmu Pengetahuan dan/atau Teknologi untuk memenuhi kepentingan nasional dan peningkatan daya saing bangsa.

Bagi mahasiswa kelulusan merupakan hal yang mesti dilalui untuk bisa mewujudkan tujuan pendidikan tinggi seperti yang tertera pada UU No. 12 tahun 2012 pasal 5. Setiap perguruan tinggi mengharapkan mahasiswanya mampu melaksanakan kelulusan sebelum batas waktu yang telah ditentukan untuk setiap jenjang studi dengan predikat yang bagus. Namun sayangnya masih banyak mahasiswa yang lulus di luar batas waktu jenjang studinya. Waktu lulus mahasiswa sering dikategorikan menjadi tepat waktu dan tidak tepat waktu berdasarkan lama waktu yang dibutuhkan untuk menyelesaikan studi, berdasarkan penelitian Al amin *et al.*,(2013) mahasiswa jenjang S1/DIV dikategorikan lulus tepat waktu apabila menyelesaikan studi kurang atau sama dengan empat tahun. Berdasarkan data dari Biro Akademik dan kemahasiswaan Universitas Negeri Padang (UNP) hanya 30,7% mahasiswa yang lulus tepat waktu dari keseluruhan

mahasiswa jenjang S1/DIV yang diwisuda pada periode Maret (126) dan Juni (127) Tahun 2022. Berikut data persentase wisudawan yang lulus tepat waktu jenjang S1/DIV dari masing-masing fakultas yang ada di UNP pada Gambar 1.



Gambar 1. Persentase Wisudawan UNP Periode 126-127

Sumber: Biro Akademik dan Kemahasiswaan UNP

Berdasarkan informasi dari Gambar 1 dapat dilihat bahwa fakultas dengan mahasiswa yang lulus tepat waktu paling banyak dari keseluruhan mahasiswa yang diwisuda pada periode tersebut dalam satu fakultas adalah Fakultas Ilmu Pendidikan (FIP) sebesar 44%, Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) 39%, Fakultas Bahasa dan Seni (FBS) 36%, dan diikuti oleh Fakultas Ilmu Keolahragaan (FIK), Fakultas Ekonomi (FE), Fakultas Ilmu Sosial, Fakultas Teknik (FT), Fakultas Psikologi dan Keperawatan (FPK) serta Fakultas Pariwisata dan Perhotelan (FPP).

Berdasarkan besarnya persentase yang diperoleh oleh Fakultas Ilmu Pendidikan hal ini menarik untuk diteliti melihat karakteristik mahasiswa yang lulus tepat waktu. Dengan tujuan untuk mendapatkan suatu informasi berupa

karakteristik yang menyerupai model atau pola dari kondisi mahasiswa FIP yang dapat menyelesaikan studi jenjang S1/ DIV kurang atau sama dengan empat tahun. Informasi tersebut nantinya diharapkan dapat dijadikan bahan pertimbangan dalam meminimalisir masalah ketepatan waktu lulus mahasiswa sehingga masing-masing fakultas dapat meningkatkan persentase lulusan tepat waktu.

Menurut Agwil *et al.*,(2020) kelulusan tepat waktu dipengaruhi oleh banyak faktor baik internal maupun eksternal dari mahasiswa, selain itu karakteristik yang berbeda menjadikan lama masa studi yang dibutuhkan untuk lulus juga beragam. Dari beragamnya faktor yang mempengaruhi masalah ketepatan waktu lulus mahasiswa, klasifikasi dirasa perlu dilakukan untuk dapat melihat karakteristik dari masalah tersebut sehingga dapat dilakukan prediksi di masa yang akan datang beserta solusi untuk mengatasinya.

Berdasarkan hal di atas dapat dilakukan *prediction modelling* (model prediksi). Menurut Prasetyo (2014: 5-6) pada model prediksi terdapat dua jenis model yaitu klasifikasi dan regresi. Klasifikasi digunakan untuk permasalahan dengan variabel respon berhingga atau berkategori, sedangkan regresi digunakan untuk permasalahan dengan variabel respon numerik (Russell dan Norvig, 2010: 696). Dalam hal ini menggunakan model klasifikasi untuk mendapatkan klasifikasi pola karakteristik waktu lulus mahasiswa dengan kategori tepat dan tidak tepat berdasarkan karakteristik yang ada pada mahasiswa. Model klasifikasi memerlukan algoritma pelatihan dalam proses pembangunannya diantaranya diperoleh dari beberapa metode klasifikasi yaitu *Decision Tree*, *Artificial Neural Network* (ANN), *Support Vector Machine* (SVM), *Bayesian Classifier*, *K-Nearest Neighbor* (KNN) dan sebagainya.

Sebagai salah satu metode yang populer digunakan *Decision Tree* memiliki beberapa kelebihan yaitu tidak membutuhkan asumsi terhadap distribusinya, tidak membutuhkan waktu yang lama dalam proses *data cleaning* karena dapat menangani *missing value* dan data berjenis campuran sehingga dapat menghemat waktu dan tenaga dalam proses pengolahan. Selain itu *decision tree* memiliki tampilan hasil yang atraktif sehingga menarik dan mudah dipahami.

CART (Classification and Regression Tree) merupakan salah satu metode *Decision Tree* atau pohon keputusan yang digunakan untuk variabel berupa numerik dan kategorik (Agwil *et al*, 2022). Berdasarkan penelitian oleh Rahmayuni (2014) untuk klasifikasi data nilai mahasiswa menggunakan algoritma C4.5 dan CART, diperoleh bahwa algoritma C4.5 lebih cocok digunakan pada data yang bersifat non-numerik seperti data nilai mahasiswa yang dikelompokkan kedalam empat kelompok (*Best, Good, Pass, dan Fail*) sedangkan algoritma *CART* dengan konsep pohon biner lebih cocok digunakan untuk data yang bersifat numerik. Pada kasus ini akan digunakan variabel prediktor berjenis campuran baik numerik ataupun bersifat kategorik tipe nominal dan ordinal maka penelitian ini akan menggunakan Algoritma *CART*. Berdasarkan latar belakang diatas, maka penelitian ini diberi judul “**Penerapan *Classification and Regression Trees* Pada Klasifikasi Ketepatan Waktu Lulus Mahasiswa S1 FIP Universitas Negeri Padang Tahun 2022**”.

B. Batasan Masalah

Berdasarkan latar belakang di atas batasan masalah pada penelitian ini adalah:

1. Data Wisudawan jenjang S1/DIV pada periode wisuda 126-127 Tahun 2022 yaitu Data Wisudawan Fakultas Ilmu Pendidikan UNP

2. Penelitian ini berfokus pada penemuan model dan ketepatan dari model yang dihasilkan.

C. Rumusan Masalah

Berdasarkan latar belakang dan batasan masalah di atas maka permasalahan pada penelitian ini adalah:

1. Bagaimana karakteristik klasifikasi ketepatan waktu kelulusan Wisudawan S1 FIP UNP periode 126-127
2. Bagaimana ketepatan klasifikasi yang dihasilkan dari ketepatan waktu kelulusan Wisudawan S1 FIP UNP periode 126-127

D. Tujuan Masalah

Berdasarkan rumusan masalah diatas, maka diperoleh tujuan sebagai berikut:

1. Mengetahui karakteristik klasifikasi ketepatan waktu kelulusan Wisudawan S1 FIP UNP periode 126-127
2. Mengetahui ketepatan klasifikasi yang dihasilkan dari ketepatan waktu kelulusan Wisudawan S1 FIP UNP periode 126-127

E. Manfaat Penulisan

Manfaat penelitian ini adalah sebagai berikut:

1. Bagi Penulis, dapat menambah ilmu pengetahuan, pemahaman dan wawasan tentang kasus yang diteliti serta menambah pengalaman dalam melakukan penelitian
2. Bagi Pembaca, sebagai informasi tentang penerapan metode *Decision Tree* algoritma *CART*.
3. Bagi Peneliti selanjutnya, diharapkan menjadi bahan masukan untuk mengembangkan dan memperluas cakupan hasil penelitian.