



TUGAS AKHIR - KS184822

**PENENTUAN BATAS MINIMAL NILAI MATA UJI
PADA TES TULIS MASUK PERGURUAN TINGGI
NEGERI TAHUN 2018 DENGAN PENDEKATAN
*CLASSIFICATION TREE***

**RHAVIDA ANNIZA ANDYANI
NRP 062115 4000 0006**

**Dosen Pembimbing
Dr. Dra. Ismaini Zain, M.Si.
Dr. Dra. Agnes Tuti Rumiati, M.Sc.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**



TUGAS AKHIR - KS184822

**PENENTUAN BATAS MINIMAL NILAI MATA UJI
PADA TES TULIS MASUK PERGURUAN TINGGI
NEGERI TAHUN 2018 DENGAN PENDEKATAN
*CLASSIFICATION TREE***

**RHAVIDA ANNIZA ANDYANI
NRP 062115 4000 0006**

**Dosen Pembimbing
Dr. Dra. Ismaini Zain, M.Si.
Dr. Dra. Agnes Tuti Rumiati, M.Sc.**

**PROGRAM STUDI SARJANA
DEPARTEMEN STATISTIKA
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**



FINAL PROJECT - KS184822

**DETERMINATION OF WRITING TESTS
MINIMUM LIMITS FOR ENTERING STATE
UNIVERSITIES 2018 WITH
CLASSIFICATION TREE APPROACH**

**RHAVIDA ANNIZA ANDYANI
NRP 062115 4000 0006**

**Supervisor
Dr. Dra. Ismaini Zain, M.Si.
Dr. Dra. Agnes Tuti Rumiati, M.Sc.**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2019**

LEMBAR PENGESAHAN

PENENTUAN BATAS MINIMAL NILAI MATA UJI PADA TES TULIS MASUK PERGURUAN TINGGI NEGERI TAHUN 2018 DENGAN PENDEKATAN CLASSIFICATION TREE

TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Statistika
pada
Program Studi Sarjana Departemen Statistika
Fakultas Matematika, Komputasi, dan Sains Data
Institut Teknologi Sepuluh Nopember

Oleh :

Rhavida Anniza Andyani
NRP. 062115 4000 0006

Disetujui oleh Pembimbing:
Dr. Dra. Ismaini Zain, M.Si.
NIP. 19600525 198803 2 001

Dra. Agnes Tuti Rumiati, M.Sc.
NIP. 19570724 198503 2 002



Mengetahui,
Dekan Departemen

Dr. Suhartono

NIP. 19710929 199512 1 001

SURABAYA, JANUARI 2019

(Halaman ini sengaja dikosongkan)

PENENTUAN BATAS MINIMAL NILAI MATA UJI PADA TES TULIS MASUK PERGURUAN TINGGI NEGERI TAHUN 2018 DENGAN PENDEKATAN CLASSIFICATION TREE

Nama Mahasiswa : Rhavida Anniza Andyani
NRP : 062115 4000 0006
Departemen : Statistika
Dosen Pembimbing 1 : Dr. Dra. Ismaini Zain, M.Si.
Dosen Pembimbing 2 : Dr. Dra. Agnes Tuti Rumiati, M.Sc.

Abstrak

Jalur tes tulis masuk perguruan tinggi selalu memiliki tingkat persaingan yang ketat. Penerimaan peserta berdasarkan peringkat skor total tes tulis yang sesuai dengan jumlah daya tampung, namun peserta yang diterima belum tentu memiliki nilai tertinggi pada mata uji yang sesuai dengan program studi pilihannya. Hal ini menarik untuk dianalisis mengenai mata uji apa saja yang diduga berpengaruh terhadap status penerimaan beserta batas nilai minimal. Penelitian menggunakan data peserta tes tulis masuk perguruan tinggi negeri tahun 2018, dengan pilihan program studi Statistika, Matematika, dan Aktuaria di Institut Teknologi Sepuluh Nopember, Institut Pertanian Bogor, dan Universitas Indonesia. Penentuan batas dilakukan dengan pendekatan Classification Tree, dimana status penerimaan sebagai variabel respon, dan variabel prediktor berupa 10 nilai mata uji. Hasil Penelitian menunjukkan bahwa variabel terpenting dalam penentuan penerimaan peminat Statistika adalah Nilai Numerik untuk Statistika ITS dan UI, serta nilai Figural untuk Statistika IPB. Prodi Matematika menggunakan nilai Kimia sebagai pemilah utama dalam penerimaan peserta tes tulis masuk PTN 2018 di ITS dan IPB, sedangkan Matematika UI menggunakan nilai Numerik dalam penentuan awal status penerimaan mahasiswa. Prodi Aktuaria ITS menggunakan nilai Fisika, IPB dengan nilai Figural, dan UI menggunakan nilai Matematika Dasar sebagai penentuan penerimaan.

Kata kunci: *Batas Minimal, Classification Tree, Mata Uji, Tes Tulis*

(Halaman ini sengaja dikosongkan)

DETERMINATION OF WRITING TESTS MINIMUM LIMITS FOR ENTERING STATE UNIVERSITIES 2018 WITH CLASSIFICATION TREE APPROACH

Name : Rhavida Anniza Andyani
Student Number : 062115 4000 0006
Department : Statistics
Supervisor : Dr. Dra. Ismaini Zain, M.Si.
Co-Supervisor : Dr. Dra. Agnes Tuti Rumiati, M.Sc.

Abstract

The college entrance test path always has a level of intense competition. Acceptance of written test participants is based on the ranking of the total written test scores according to the amount of capacity. However, participants who are accepted do not necessarily have the highest score in the subject that is in accordance with their chosen study program. Therefore, it is interesting to do an analysis of the test subjects that influence the acceptance status along with the passing score. The study used the data of college entrance written test participants in 2018, with a choice of Statistics, Mathematics, and Actuarial study programs at the Institut Teknologi Sepuluh Nopember (ITS), Institut Pertanian Bogor (IPB), and the University of Indonesia (UI). The analysis used the Classification Tree approach, with acceptance status as the response variable, and 10 test subjects as predictor variables. The result showed that the most important variable in determining the acceptance of interest in Department of Statistics was the Numeric for Department of Statistics in ITS and UI, and the Figural for Statistics IPB. Whereas in Department of Mathematics, it used the score from Chemistry as the main factor in the acceptance of college entrance written test participants in 2018 in Department of Mathematics ITS and IPB, while in UI uses Numeric score as the initial determination of student acceptance status. For Department of Actuarial ITS use Physic score, IPB use Figural score and UI with Basic Mathematic for acceptance status.

Keywords: Classification Tree, Minimum Limits, Test Subject, Written test

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur penulis panjatkan atas rahmat dan hidayah yang diberikan Allah SWT sehingga penulis dapat menyelesaikan laporan Tugas Akhir yang berjudul **“Penentuan Batas Minimal Nilai Mata Uji Pada Tes Tulis Masuk Perguruan Negeri Tahun 2018 dengan Pendekatan *Classification Tree*”** dengan lancar.

Penulis menyadari bahwa Tugas Akhir ini dapat terselesaikan tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada :

1. Ibu, ayah, kakak, dan keluarga atas segala do'a, semangat, dukungan, dan nasehat yang selalu diberikan kepada penulis demi kesuksesan dan kebahagiaan penulis.
2. Dr. Dra. Ismaini Zain, M.Si. dan Dr. Dra. Agnes Tuti Rumiati, M.Sc. selaku dosen pembimbing yang telah meluangkan waktu dan dengan sabar memberikan arahan, bimbingan, saran, dukungan serta motivasi selama penyusunan Tugas Akhir.
3. Dr. Vita Ratnasari, S.Si., M.Si. dan Erma Oktania Permatasari, S.Si., M.Si. selaku dosen penguji yang telah banyak memberikan kritik dan saran demi kesempurnaan Tugas Akhir ini.
4. Dr. Suhartono selaku Ketua Departemen Statistika ITS, yang telah banyak memberikan saran dan arahan dalam proses belajar di Departemen Statistika.
5. Dr. Sutikno, M.Si. selaku Ketua Program Studi Sarjana Statistika ITS yang telah memberikan fasilitas, sarana, dan prasarana dalam proses belajar di Departemen Statistika.
6. Muhammad Sjahid Akbar, S.Si., M.Si., selaku dosen wali, seluruh dosen pengajar, dan karyawan Statistika ITS atas ilmu dan pengalaman yang telah diberikan kepada penulis.
7. Teman-teman seperjuangan Tugas Akhir, khususnya Natasya, Anggi, Shahnaz, Nanda, dan Naufal yang selama ini telah

membantu, berjuang bersama dan saling memberikan semangat

8. Teman-teman Statistika ITS khususnya Σ 26, para sahabat, Sri Mulyani, Juwitasari, Fraschiska, Regita, Fibia, Shindy, Lianna, Hendri yang selama ini telah membantu, mendukung, dan mendengarkan keluh kesah penulis selama masa perkuliahan berlangsung.
9. Semua pihak yang turut membantu dalam pelaksanaan Tugas Akhir yang tidak bisa penulis sebutkan satu persatu.

Semoga kebaikan dan bantuan yang telah diberikan kepada penulis dibalas dengan kebaikan yang lebih oleh Allah SWT. Aamiin.

Penulis menyadari bahwa Tugas Akhir ini masih terdapat kekurangan, oleh karena itu sangat diharapkan kritik dan saran yang bersifat membangun. Semoga Tugas Akhir ini dapat memberikan manfaat baik bagi penulis, pembaca, dan semua pihak.

Surabaya, Januari 2019
Penulis

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
TITLE PAGE	iii
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xix
DAFTAR LAMPIRAN	xxi
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan.....	5
1.4 Manfaat.....	5
1.5 Batasan Masalah.....	6
BAB II TINJAUAN PUSTAKA	7
2.1 Ujian Tulis Masuk Perguruan Tinggi Negeri.....	7
2.2 <i>Classification Tree</i>	8
2.2.1 Proses Pembentukan Pohon Klasifikasi.....	11
2.2.2 Pemangkasan Pohon Klasifikasi (<i>Prunning</i>).....	13
2.2.3 Penentuan Ukuran Pohon Klasifikasi Optimal ..	14
2.3 Evaluasi Ketepatan Klasifikasi.....	15
BAB III METODOLOGI PENELITIAN	19
3.1 Sumber Data.....	19
3.2 Variabel Penelitian.....	19
3.3 Langkah Analisis.....	20
3.4 Diagram Alir.....	22
BAB IV ANALISIS DAN PEMBAHASAN	25
4.1 Deskripsi Peminat Prodi Statistika, Matematika, dan Aktuaria di ITS, IPB dan UI.....	25

4.1.1 Deskripsi Peminat Berdasarkan Status Penerimaan, Jenis Kelamin, dan Asal Sekolah	25
4.1.2 Karakteristik Nilai Peminat Statistika, Matematika, dan Aktuaria ITS, IPB, dan UI	29
4.2 Penentuan Mata Uji yang Signifikan dan Batas Nilai Minimal	45
4.2.1 Penentuan Mata Uji yang Signifikan dan Batas Nilai Minimal Program Studi Statistika	45
4.2.2 Penentuan Mata Uji yang Signifikan dan Batas Nilai Minimal Program Studi Matematika	58
4.2.3 Penentuan Mata Uji yang Signifikan dan Batas Nilai Minimal Program Studi Aktuaria	69
4.3 Perbandingan Batasan Nilai Peminat yang Diterima pada Prodi Statistika, Matematika, dan Aktuaria	78
BAB V KESIMPULAN DAN SARAN	83
5.1 Kesimpulan	83
5.2 Saran	84
DAFTAR PUSTAKA	85
LAMPIRAN	87
BIODATA PENULIS	133

DAFTAR GAMBAR

	Halaman
Gambar 2.1	Ilustrasi <i>Classification Tree</i> 10
Gambar 3.1	Diagram Alir Penelitian 22
Gambar 3.2	Diagram Alir Pembentukan Pohon Klasifikasi Maksimal 24
Gambar 4.1	Perbandingan Jumlah Peserta Diterima dan Tidak Diterima 26
Gambar 4.2	Perbandingan Jumlah Peminat Prodi Statistika, Matematika, dan Aktuaria Berdasarkan Jenis Kelamin..... 27
Gambar 4.3	Persentase Jumlah Peminat Berdasarkan Asal Sekolah 28
Gambar 4.4	<i>Density Plot</i> Nilai TKPA Peminat Statistika 30
Gambar 4.5	<i>Density Plot</i> Nilai TKD Saintek Peminat Statistika..... 31
Gambar 4.6	<i>Box Plot</i> Nilai Mata Uji Tes Tulis Peminat Statistika ITS..... 32
Gambar 4.7	<i>Box Plot</i> Nilai Mata Uji Tes Tulis Peminat Statistika IPB 33
Gambar 4.8	<i>Box Plot</i> Nilai Mata Uji Tes Tulis Peminat Statistika UI 34
Gambar 4.9	<i>Density Plot</i> Nilai TKPA Peminat Matematika .. 35
Gambar 4.10	<i>Density Plot</i> Nilai TKD Saintek Peminat Matematika 36
Gambar 4.11	<i>Box Plot</i> Nilai Mata Uji Tes Tulis Peminat Matematika ITS 37
Gambar 4.12	<i>Box Plot</i> Nilai Mata Uji Tes Tulis Peminat Matematika IPB 38
Gambar 4.13	<i>Box Plot</i> Nilai Mata Uji Tes Tulis Peminat Matematika UI 39
Gambar 4.14	<i>Density Plot</i> Nilai TKPA Peminat Aktuaria 40
Gambar 4.15	<i>Density Plot</i> Nilai TKD Saintek Peminat Aktuaria 41

Gambar 4.16	<i>Box Plot</i> Nilai Mata Uji Tes Tulis Peminat Aktuaria ITS.....	42
Gambar 4.17	<i>Box Plot</i> Nilai Mata Uji Tes Tulis Peminat Aktuaria IPB	43
Gambar 4.18	<i>Box Plot</i> Nilai Mata Uji Tes Tulis Peminat Aktuaria UI	44
Gambar 4.19	Pohon Klasifikasi Maksimal Prodi Statistika ITS.....	46
Gambar 4.20	<i>Relative Cost Plot</i> Data Peminat Statistika ITS ..	47
Gambar 4.21	Hasil Pohon Klasifikasi Optimal Peminat Prodi Statistika ITS	49
Gambar 4.22	Pohon Klasifikasi Maksimal Prodi Statistika IPB	52
Gambar 4.23	<i>Relative Cost Plot</i> Data Peminat Statistika IPB ..	52
Gambar 4.24	Pohon Klasifikasi Optimal Data Peminat Statistika IPB	53
Gambar 4.25	Pohon Klasifikasi Maksimal Data Peminat Statistika UI	55
Gambar 4.26	<i>Relative Cost Plot</i> Data Peminat Statistika UI....	55
Gambar 4.27	Pohon Klasifikasi Optimal Data Peminat Statistika UI	57
Gambar 4.28	Pohon Klasifikasi Maksimal Data Peminat Prodi Matematika ITS.....	59
Gambar 4.29	<i>Relative Cost Plot</i> Data Peminat Matematika ITS.....	59
Gambar 4.30	Pohon Klasifikasi Optimal Data Peminat Matematika ITS.....	60
Gambar 4.31	Pohon Klasifikasi Maksimal Data Peminat Matematika IPB	62
Gambar 4.32	<i>Relative Ccost Plot</i> Data Peminat Matematika IPB	62
Gambar 4.33	Pohon Klasifikasi Optimal Data Peminat Matematika IPB	63
Gambar 4.34	Pohon Klasifikasi Maksimal Data Peminat Matematika UI	65

Gambar 4.35	<i>Relative Cost Plot</i> Data Peminat Matematika UI	66
Gambar 4.36	Pohon Klasifikasi Optimal Peminat Matematika UI	67
Gambar 4.37	Pohon Klasifikasi Maksimal Data Peminat Aktuarial ITS	69
Gambar 4.38	<i>Relative cost Plot</i> Data Peminat Aktuarial ITS....	70
Gambar 4.39	Pohon Klasifikasi Optimal Data Peminat Aktuarial ITS	70
Gambar 4.40	Pohon Klasifikasi Maksimal Prodi Statistika IPB	72
Gambar 4.41	<i>Relative Cost Plot</i> Data Peminat Statistika IPB	72
Gambar 4.42	Pohon Klasifikasi Optimal Data Peminat Aktuarial IPB	74
Gambar 4.43	Pohon Klasifikasi Maksimal Data Peminat Aktuarial UI	76
Gambar 4.44	<i>Relative Cost Plot</i> Data Peminat Aktuarial UI....	76
Gambar 4.45	Pohon Klasifikasi Optimal Data Peminat Aktuarial UI	77

(Halaman ini sengaja dikosongkan)

DAFTAR TABEL

	Halaman
Tabel 2.1	Jenis Materi Ujian Tes Tulis 7
Tabel 2.2	<i>Crosstab</i> Ketepatan Klasifikasi..... 16
Tabel 3.1	Jumlah Data Penelitian 19
Tabel 3.2	Struktur Data Penelitian..... 19
Tabel 3.3	Variabel Penelitian untuk Analisis <i>Classification Tree</i> 20
Tabel 3.4	Variabel Penelitian untuk Analisis Statistika Deskriptif 20
Tabel 4.1	Jumlah Peserta Diterima Berdasarkan Asal Sekolah..... 29
Tabel 4.2	Skor Kontribusi Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Maksimal 46
Tabel 4.3	Skor Kontribusi Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Optimal 48
Tabel 4.4	Evaluasi Ketepatan Klasifikasi Peminat Statistika ITS..... 51
Tabel 4.5	Evaluasi Ketepatan Klasifikasi Peminat Statistika IPB 54
Tabel 4.6	Evaluasi Ketepatan Klasifikasi Peminat Statistika UI 58
Tabel 4.7	Evaluasi Ketepatan Klasifikasi Peminat Matematika ITS 61
Tabel 4.8	Evaluasi Ketepatan Klasifikasi Peminat Matematika IPB 64
Tabel 4.9	Evaluasi Ketepatan Klasifikasi Peminat Matematika UI 68
Tabel 4.10	Evaluasi Ketepatan Klasifikasi Peminat Aktuaria ITS 71
Tabel 4.11	Evaluasi Ketepatan Klasifikasi Peminat Aktuaria IPB 75
Tabel 4.12	Evaluasi Ketepatan Klasifikasi Peminat Aktuaria UI 77

Tabel 4.13	Perbandingan Batasan Nilai Penerimaan Peserta Tes Tulis Masuk PTN 2018 Prodi Statistika.....	78
Tabel 4.14	Perbandingan Batasan Nilai Penerimaan Peserta Tes Tulis Masuk PTN 2018 Prodi Matematika...	79
Tabel 4.15	Perbandingan Batasan Nilai Penerimaan Peserta Tes Tulis Masuk PTN 2018 Prodi Aktuaria.....	80

DAFTAR LAMPIRAN

	Halaman
Lampiran 1A. Data Peserta Peminat Prodi Statistika ITS	87
Lampiran 1B. Data Peserta Peminat Prodi Statistika IPB	88
Lampiran 1C. Data Peserta Peminat Prodi Statistika UI.....	89
Lampiran 2A. Data Peserta Peminat Prodi Matematika ITS	90
Lampiran 2B. Data Peserta Peminat Prodi Matematika IPB	91
Lampiran 2C. Data Peserta Peminat Prodi Matematika UI.....	92
Lampiran 3A. Data Peserta Peminat Prodi Aktuaria ITS	93
Lampiran 3B. Data Peserta Peminat Prodi Aktuaria IPB	94
Lampiran 3C. Data Peserta Peminat Prodi Aktuaria UI.....	95
Lampiran 4. <i>Output</i> CART Hasil Analisis Data Prodi Statistika ITS.....	96
Lampiran 5. <i>Output</i> CART Hasil Analisis Data Prodi Statistika IPB	103
Lampiran 6. <i>Output</i> CART Hasil Analisis Data Prodi Statistika UI	106
Lampiran 7. <i>Output</i> CART Hasil Analisis Data Prodi Matematika ITS	111
Lampiran 8. <i>Output</i> CART Hasil Analisis Data Prodi Matematika IPB	114
Lampiran 9. <i>Output</i> CART Hasil Analisis Data Prodi Matematika IPB	117
Lampiran 10. <i>Output</i> CART Hasil Analisis Data Prodi Aktuaria ITS	120
Lampiran 11. <i>Output</i> CART Hasil Analisis Data Prodi Aktuaria IPB	122
Lampiran 12. <i>Output</i> CART Hasil Analisis Data Prodi Aktuaria ITS	127
Lampiran 13. Surat Keterangan Pengambilan Data	131

(Halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Penerimaan mahasiswa baru program sarjana di perguruan tinggi negeri Indonesia dapat dilakukan melalui 3 jalur seleksi, yakni Seleksi Nasional Masuk Perguruan Tinggi Negeri (SNMPTN) atau yang disebut dengan jalur seleksi rapor, Seleksi Bersama Masuk Perguruan Tinggi (SBMPTN) yang selanjutnya disebut dengan jalur tes tulis, dan seleksi Mandiri. Dalam pelaksanaannya, jalur tes tulis menjadi jalur dengan jumlah penerimaan paling banyak dibandingkan jalur seleksi rapor dan seleksi mandiri. Tes tulis dilaksanakan dengan calon mahasiswa dapat memilih 3 program studi, dengan salah satu prodi berasal dari PTN sesuai dengan asal daerahnya.

Tahun 2018, sebanyak 165.831 peserta telah dinyatakan lulus dan diterima di 85 PTN, dari 860.001 pendaftar tes tulis. Rincian peserta diterima berdasarkan kelompok ujian menunjukkan bahwa kelompok saintek memiliki jumlah diterima paling banyak, yakni sebesar 68.333 peserta (Uly, 2018). Banyaknya peminat suatu program studi, menyebabkan tingkat persaingan yang terjadi semakin ketat, karena daya tampung yang dimiliki di setiap perguruan tinggi yang ada relatif sedikit.

Institusi Teknologi Sepuluh Nopember merupakan PTN dengan rumpun program studi Saintek. ITS menyediakan berbagai prodi Saintek, diantaranya adalah Statistika, Matematika dan Aktuaria. Ketiga prodi tersebut tergabung dalam satu fakultas yang baru saja terbentuk, yakni Fakultas Matematika, Komputasi, dan Sains Data. Aktuaria merupakan prodi baru yang dibuka oleh ITS dengan jumlah peminat yang melebihi peminat statistika dan Matematika Tes tulis masuk PTN tahun 2018, peminat Aktuaria berjumlah 1066 peserta dengan daya tampung pada jalur tes sebesar 24. Tingkat ketetapan yang ada adalah 1 : 45, dimana 1 peserta yang diterima harus dapat mengalahkan 45 peserta lainnya.

Peserta tes tulis dinyatakan lulus di salah satu perguruan tinggi apabila peserta berada pada peringkat sesuai dengan jumlah penerimaan di setiap prodi berdasarkan total skor tes tulis. Namun, peserta yang diterima belum tentu memiliki nilai tertinggi pada mata uji yang sesuai dengan program studi pilihannya. Sebagai contoh, peserta diterima prodi matematika belum tentu memiliki nilai matematika dasar yang bagus pula. Selain dengan menggunakan hasil tes tes tulis, strategi dalam memilih program studi dan Perguruan Tinggi Negeri (PTN) yang tepat sangat mempengaruhi diterima atau tidaknya peserta. Urutan pilihan prodi menyatakan prioritas pilihan. Dalam pelaksanaan tes tulis, peserta dapat memilih program studi sebanyak-banyaknya tiga program studi, dengan ketentuan: memilih satu program studi di perguruan tinggi negeri manapun, dan jika lebih dari satu prodi, salah satu pilihan harus di PTN satu wilayah pendaftaran dengan tempat mengikuti ujian, pilihan prodi lain dapat di PTN di luar wilayah pendaftaran (SBMPTN, 2018). Adanya aturan pemilihan prodi berdasarkan wilayah, dapat memberikan karakteristik tersendiri mengenai peserta yang diterima di suatu prodi. Karakteristik tersebut dapat berupa asal daerah, asal sekolah, dan jenis kelamin.

Berdasarkan uraian sebelumnya, pada penelitian ini dilakukan analisis untuk mengetahui mata uji tes tulis apa saja yang dapat mempengaruhi status diterimanya peserta pada program studi Statistika, Matematika, dan Aktuaria Institut Teknologi Sepuluh Nopember beserta batasan nilai minimal pada mata uji tersebut. Selain ITS, akan dilakukan pula penentuan batasan pada prodi Statistika dan Matematika Universitas Indonesia (UI) dan Institut Pertanian Bogor (IPB). Pemilihan kedua universitas tersebut berdasarkan hasil klasterisasi perguruan tinggi non vokasi tahun 2018 yang dikeluarkan oleh RISTEKDIKTI, dimana kedua universitas tersebut memiliki peringkat diatas peringkat ITS (Ristekdikti, 2018). Selain itu, pemilihan kedua Universitas tersebut mempertimbangkan adanya prodi Statistika, Matematika, dan Aktuaria. Dalam hal ini variabel

yang digunakan adalah status penerimaan peserta sebagai variabel respon, dengan variabel prediktor berupa nilai mata uji TKPA, TKD Saintek. Selain itu akan digunakan pula variabel jenis kelamin, dan status sekolah untuk mengetahui karakteristik peserta.

Metode yang umum digunakan pada data dengan variabel berupa kategorik adalah regresi logistik. Metode ini merupakan metode parametrik, yakni metode yang harus memenuhi beberapa asumsi agar dapat memberikan hasil yang akurat. Menurut Webb dan Johannes (1999), regresi logistik memiliki nilai probabilitas yang cenderung sulit dilakukan interpretasi. Adanya kelemahan pada penerapan metode regresi logistik, maka pada penelitian ini menggunakan metode nonparametrik untuk menyelesaikan permasalahan diatas, yakni dengan menggunakan metode *classification tree*. Metode *Classification tree* merupakan bagian dari metode klasifikasi dari *Classification and Regression Tree* (CART). Menurut Breiman, *et al.* (1993), CART merupakan metodologi statistik nonparametrik yang dikembangkan untuk topik analisis klasifikasi, baik untuk variabel respon kategorik maupun kontinyu.

CART menghasilkan suatu pohon klasifikasi jika variabel responnya kategorik, dan menghasilkan pohon regresi jika variabel responnya kontinyu. *Classification tree* merupakan metode nonparametrik yang dapat digunakan untuk melihat hubungan antara peubah respon kategorik dengan peubah variabel prediktor yang berukuran besar dan kompleks. Menurut Lewis (2000), kelebihan menggunakan metode *classification tree* antara lain dapat bekerja pada data dengan dimensi dan struktur yang lebih kompleks, interpretasi hasilnya lebih mudah daripada proses regresi biasa, karena identifikasi adanya pengaruh dari variabel prediktor dalam pohon regresi dilakukan dalam masing-masing subgroup data bukan dalam keseluruhan data seperti halnya regresi biasa.

Penelitian dengan menggunakan metode *classification tree* telah dilakukan oleh Ghasemzadeh, *et al* (2018) dengan

menghasilkan suatu kesimpulan bahwa metode *classification tree* dapat dengan mudah menjelaskan interaksi yang kompleks antara beberapa variabel penjelas, dibandingkan dengan menggunakan regresi logistik. Karpen dan Ellis (2018), menggunakan metode *classification tree* dalam penerimaan sekolah farmasi. Penelitian tersebut memberikan hasil bahwa *classification tree* dapat digunakan untuk memfasilitasi berbagai keputusan dalam pendidikan farmasi dan sangat berguna untuk analisis eksplorasi. Penelitian mengenai tes tulis masuk perguruan tinggi telah dilakukan oleh Pratiwi (2015) dengan menggunakan metode *Classification and Regression Tree* (CART) dalam melakukan klasifikasi pada peserta tes tulis dengan peminat yang mendaftar pada program studi di Institut Teknologi Sepuluh Nopember (ITS) berdasarkan status penerimaan. Menurut Anggreini (2015), faktor yang memberikan pengaruh diterimanya peserta ujian tes tulis adalah Tes Potensial Akademik (TPA) dan Tes Kemampuan Dasar (TKD) Saintek. Selain itu, pemilihan jurusan memiliki pengaruh dalam penerimaan peserta tes tulis masuk perguruan tinggi.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, permasalahan yang dibahas pada penelitian ini adalah sebagai berikut.

1. Bagaimana deskripsi dan gambaran nilai tes tulis peminat program studi Statistika, Matematika, dan Aktuaria pada ITS, IPB, dan UI pada tes tulis masuk perguruan tinggi tahun 2018?
2. Bagaimana hasil penentuan batas minimal nilai mata uji pada tes tulis masuk perguruan tinggi tahun 2018 berdasarkan status penerimaan peserta pada prodi Statistika, Matematika, dan Aktuaria ITS, IPB, dan UI dengan menggunakan pendekatan *Classification Tree*?

3. Bagaimana hasil perbandingan batasan nilai mata uji tes tulis agar dapat diterima pada prodi Statistika, Matematika, dan Aktuaria antara ITS, IPB, dan UI?

1.3 Tujuan

Berdasarkan rumusan masalah yang telah disusun, maka tujuan yang ingin dicapai dalam penelitian ini adalah sebagai berikut.

- 1 Mendeskripsikan dan memberikan gambaran mengenai nilai tes tulis peminat program studi Statistika, Matematika, dan Aktuaria antara ITS, IPB, dan UI pada tes tulis masuk perguruan tinggi negeri tahun 2018.
- 2 Menentukan batas minimal nilai mata uji agar diterima pada program studi Statistika, Matematika, dan Aktuaria ITS, IPB dan UI pada tes tulis masuk perguruan tinggi tahun 2018 dengan menggunakan pendekatan *Classification Tree*.
- 3 Membandingkan hasil batasan nilai minimal peserta untuk diterima di Statistika, Matematika, dan Aktuaria ITS, IPB, dan UI.

1.4 Manfaat

Manfaat yang dapat diberikan dari hasil penelitian ini adalah sebagai berikut.

1. Memberikan informasi mengenai karakteristik peminat program studi Statistika, Matematika, dan Aktuaria di ITS, IPB, dan UI pada tes tulis masuk perguruan tinggi negeri tahun 2018.
2. Mengetahui mata uji yang berpengaruh signifikan terhadap status penerimaan calon mahasiswa baru program studi Statistika, Matematika, dan Aktuaria ITS, IPB, dan UI.
3. Mengetahui batas minimal nilai mata uji tes tulis khusus pada pilihan program studi Statistika dan Matematika, sehingga dapat dijadikan sebagai rekomendasi dalam penerimaan peserta tes tulis masuk perguruan tinggi negeri pada tahun berikutnya.

1.5 Batasan Masalah

Batasan masalah yang digunakan dalam penelitian ini dapat diuraikan sebagai berikut.

1. Data yang digunakan adalah data peserta tes tulis masuk perguruan tinggi tahun 2018 dengan pilihan program studi Statistika, Matematika, dan Aktuaria di ITS, IPB, dan UI, baik pada pilihan 1, 2, ataupun 3.
2. Bagi peserta yang mengikuti ujian campuran, nilai yang digunakan hanya nilai TKPA dan TKD Saintek.
3. Pembagian data *training* dan *testing* adalah sebesar $\frac{2}{3}$ untuk *training* dan $\frac{1}{3}$ data *testing* setiap kategori atau status penerimaan.

BAB II

TINJAUAN PUSTAKA

2.1 Ujian Tulis Masuk Perguruan Tinggi Negeri

Ujian tulis atau tes tulis merupakan salah satu jalur pendaftaran Perguruan Tinggi Negeri di Indonesia. Tes tulis merupakan seleksi penerimaan mahasiswa baru melalui ujian tulis atau kombinasi hasil ujian tulis dengan ujian keterampilan. Tujuan diadakannya tes tulis adalah untuk menyeleksi calon mahasiswa yang diprediksi mampu menyelesaikan studi di PTN dengan baik dan tepat waktu serta memberikan kesempatan bagi calon mahasiswa untuk memilih lebih dari satu PTN lintas wilayah (SBMPTN, 2018).

Tes tulis dirancang untuk dapat mengukur kemampuan dasar yang dapat memprediksi keberhasilan calon mahasiswa di semua program studi, yakni kemampuan penalaran tingkat tinggi (*higher order thinking*), yang meliputi potensi akademik, penguasaan bidang studi dasar, bidang sains dan teknologi (saintek), serta bidang sosial dan humaniora (soshum).

Tabel 2.1 Jenis Materi Ujian Tes Tulis

Materi Ujian	Mata Uji
Tes Kemampuan dan Potensi Akademik (TKPA)	Matematika Dasar, Bahasa Indonesia, Bahasa Inggris, TPA Verbal, TPA Numerikal, dan TPA Figural
Tes Kemampuan Dasar Sains dan Teknologi (TKD Saintek)	Matematika IPA, Biologi, Kimia, dan Fisika
Tes Kemampuan Dasar Sosial dan Humaniora (TKD Soshum)	Sosiologi, Sejarah, Geografi, dan Ekonomi

Peserta tes tulis dapat memilih program studi sebanyak - banyaknya tiga program studi, dengan ketentuan.

- a. Peserta ujian yang hanya memilih satu program studi dapat memilih prodi di perguruan tinggi negeri manapun.

- b. Peserta yang memilih lebih dari satu prodi, salah satu pilihan prodi harus di PTN satu wilayah pendaftaran dengan tempat peserta mengikuti ujian, pilihan prodi lain dapat di PTN di luar wilayah pendaftaran.

Metode penilaian tes tulis 2018 menggunakan metode baru yakni dengan menerapkan *Item Response Theory* (IRT). Metode sebelumnya adalah dengan menggunakan penilaian skor 4 untuk jawaban yang benar, -1 untuk jawaban salah dan 0 untuk tidak dijawab. Metode penilaian pada dengan sistem IRT tidak hanya memperhitungkan jumlah soal yang dijawab benar atau salah tetapi juga memperhitungkan karakteristik setiap soal khususnya tingkat kesulitan. Metode penilaian dilakukan dengan 3 tahap berikut (SBMPTN, 2018).

- Tahap I, memberikan skor 1 pada setiap jawaban benar dan skor 0 untuk jawaban salah atau tidak dijawab.
- Tahap II, dengan pendekatan IRT maka setiap soal dilakukan analisis karakteristik berupa tingkat kesulitan soal, dengan output berupa dapat diketahui tingkat kesulitan soal (mudah, sedang, sulit).
- Tahap III, tingkat kesulitan soal digunakan untuk menghitung skor peserta. Soal yang relatif sulit akan mendapatkan nilai yang lebih tinggi.

Setiap soal dianalisis karakteristiknya dan dikategorikan. Soal yang relatif sulit akan mendapatkan nilai yang lebih tinggi. Setiap peserta yang dapat menjawab soal dengan jumlah yang sama dan benar, dapat memperoleh nilai yang sama atau berbeda. Hal ini tergantung pada soal mana saja yang mereka jawab dengan benar. Adanya penilaian dengan menggunakan sistem IRT dianggap lebih menguntungkan baik bagi peserta, maupun pihak PTN yang menerima, karena karakteristik dan kemampuan setiap peserta dapat diketahui lebih akurat.

2.2 *Classification Tree*

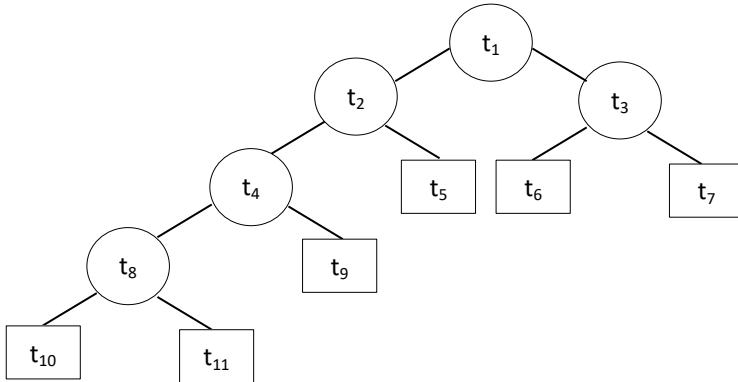
Pada penelitian ini akan dilakukan analisis dengan menggunakan metode *Classification tree*, dengan variabel respon

berupa status penerimaan peserta, yakni diterima atau tidak diterima. Variabel prediktor berupa nilai 10 mata uji tes tulis. *Classification tree* merupakan salah bagian dari metode *Classification and Regression Tree* (CART) dengan data variabel respon berupa data kategorik. Variabel yang berpengaruh dalam pohon klasifikasi adalah variabel yang menentukan adanya pemilahan (*splitting*) bagi ruang variabel prediktor (Lewis, 2000).

Pengklasifikasian dengan menggunakan metode CART memiliki beberapa kelebihan, antara lain (Lewis, 2000).

- a. Metode CART bersifat non parametrik sehingga tidak memerlukan asumsi yang mengikat seperti asumsi distribusi normal untuk variabel prediktor.
- b. CART mampu mengeksplorasi data berdimensi tinggi dengan komputasi yang efisien
- c. Metode ini memperhitungkan interaksi antar variabel prediktor yang berpengaruh. Hal ini berkaitan dengan pengambilan keputusan yang dilakukan secara bertahap dalam himpunan data yang kompleks
- d. Dapat menggunakan kombinasi data kontinyu maupun kategorik dapat digunakan pada metode ini.
- e. CART tidak hanya memberikan hasil klasifikasi, namun juga memberikan suatu nilai threshold pada setiap pemilah, dan hasil klasifikasi mudah diinterpretasikan.

Analisis CART memiliki sifat "*binary recursive partitioning*". Istilah "*binary*" mengimplikasikan bahwa sekelompok data yang terkumpul dalam suatu ruang direpresentasikan sebagai simpul atau *node* pada pohon keputusan, dimana *node* tersebut hanya dapat diklasifikasikan menjadi dua grup. Dengan kata lain, setiap *node* dapat diklasifikasikan menjadi dua simpul anak (*child nodes*), dan dalam hal ini simpul utama disebut sebagai *parent node* (Breiman, Friedman, Olshen, & Stone, 1993).



Gambar 2.1 Ilustrasi *Classification Tree*

Ilustrasi dari hasil pohon klasifikasi dapat dilihat pada Gambar 2.1. *Parent node* dinotasikan dengan t_1 yang merupakan simpul awal. Simpul dengan notasi t_2 , t_3 , dan t_4 disebut dengan simpul dalam atau *internal nodes*, sedangkan t_5 , t_6 , t_7 , t_9 , t_{10} , dan t_{11} merupakan simpul akhir atau *terminal nodes*. Menurut Segal (1988), terdapat empat komponen dalam proses pembentukan pohon klasifikasi. Komponen tersebut antara lain.

1. Himpunan pertanyaan dikotomis dengan bentuk pertanyaan “Apakah $x_i \in A$?”, x_i adalah amatan sampel dan $A \subset X$ (X merupakan ruang variabel independen). Hasil jawaban akan menentukan sekatan bagi ruang variabel independen. Data yang memiliki jawaban ya akan masuk ke anak ruang A , sedangkan data dengan jawaban tidak akan masuk kedalam ruang komplement A .
2. Kriteria pemilahan terbaik untuk alat evaluasi pemilahan menggunakan kriteria *goodness of split* ($\phi(s, t)$).
3. Penentuan ukuran pohon yang layak (*right sized tree*).
4. Statistik rata-rata variabel independen sebagai dugaan variabel dependen.

Analisis dengan menggunakan metode *Classification Tree*, data dibagi menjadi dua, yakni data *training* dan data *testing*. Menurut Witten (2005), Hand dan Kamber (2006), pada

umumnya perbandingan yang digunakan adalah 2:1 untuk data *training* dan *testing*, sehingga dalam penelitian ini menggunakan 2/3 untuk data *training* dan 1/3 data *testing*. Data *training* digunakan untuk membentuk pohon klasifikasi, dan data *testing* untuk evaluasi ketepatan klasifikasi. Tahapan yang terdapat dalam metode pohon klasifikasi terdiri dari tiga tahap, yakni pembentukan pohon klasifikasi, pemangkasan pohon klasifikasi, dan penentuan pohon klasifikasi optimum.

2.2.1 Proses Pembentukan Pohon Klasifikasi

Pohon klasifikasi terbentuk dari hasil pemilihan data ada setiap simpul induk menjadi dua simpul anak berdasarkan variabel terpenting. Kemudian kedua simpul anak tersebut dipilah kembali menjadi dua, dan selanjutnya. Terdapat tiga tahapan dalam proses pembentukan pohon klasifikasi, yakni pemilihan pemilah, penentuan simpul terminal, dan penandaan label kelas.

a. Pemilahan pemilah

Pemilahan pada sampel data *training* didasarkan pada aturan pemilahan dan kriteria *goodness of split*, dimana sampel data *training* yang digunakan masih bersifat heterogen. Aturan pemilahan adalah sebagai berikut (Breiman, Friedman, Olshen, & Stone, 1993).

1. Setiap pemilah bergantung pada satu nilai dari satu variabel prediktor.
2. Jika variabel prediktor berupa kontinu x_j , maka pemilahan berasal dari pertanyaan “Apakah nilai amatan $x_j \leq c$?”, c merupakan bilangan terkecil dari data amatan. Apabila ruang sampel berukuran N dan terdapat sebanyak-banyaknya n nilai amatan yang berbeda pada variabel x_j , maka terdapat $n-1$ pemilah berbeda yang dibentuk oleh gugus pertanyaan {“Apakah $x_j \leq c_i$?”} dengan $i = 1, 2, \dots, n-1$ dan c_i adalah nilai rata-rata dari variabel x_j berukuran berbeda.
3. Jika variabel prediktor adalah kategorik, pemilahan berasal dari semua kemungkinan pemilahan. Jika skala data berupa nominal dengan jumlah kategori sebanyak L , maka

pemilahan berjumlah $2^{L-1}-1$. Jika skala data berupa ordinal dengan jumlah kategori L , maka pemilahan berjumlah $L-1$.

Proses pemilahan ini diharuskan terjadi penurunan tingkat keheterogenan, artinya hasil pemilahan lebih homogen dibandingkan dengan simpul induknya. Tingkat keheterogenan tersebut diukur dengan menggunakan nilai *impurity* (i). Fungsi *impurity* (i) yang umum digunakan adalah indeks gini. Fungsi indeks gini dapat dituliskan dalam persamaan 2.1.

$$i(t) = \sum_{j \neq k} p(j|t)p(k|t) \quad (2.1)$$

keterangan

$i(t)$ = Indeks gini pada simpul t

$p(j|t)$ = Proporsi kelas j pada simpul t

$p(k|t)$ = Proporsi kelas k pada simpul t

Pemilahan terbaik merupakan pemilahan yang memaksimumkan ukuran kehomogenan dalam masing-masing simpul anak relatif terhadap simpul induk dan memaksimumkan ukuran pemisahan antara simpul anak kiri (t_L) dan kanan (t_R). Pemilahan terbaik menggunakan kriteria *Goodness of Split* ($\phi(s, t)$), yakni pemilahan yang dihitung berdasarkan selisih jumlah kuadrat sisaan antara simpul induk dengan kedua simpul anak pemilahnya.

$$\phi(s, t) = i(t) - \{p_L i(t_L) + p_R i(t_R)\} \quad (2.2)$$

$$\phi(s^*, t) = \max \phi(s, t) \quad (2.3)$$

keterangan

$i(t)$ = Indeks gini variabel respon pada simpul t

p_L = Proporsi pengamatan pada simpul kiri

$i(t_L)$ = Indeks gini variabel prediktor pada simpul anak kiri

p_R = Proporsi pengamatan pada simpul kanan

$i(t_R)$ = Indeks gini variabel prediktor pada simpul anak kanan

Pemilah terbaik akan menghasilkan nilai $\phi(s, t)$ tertinggi dan mampu mereduksi heterogenitas yang tinggi. Pemilah tersebut akan membentuk simpul dan kemudian akan melakukan pemilihan rekursif (berulang) sampai dihasilkan simpul akhir (*terminal nodes*).

Setiap variabel prediktor akan memberikan nilai yang menunjukkan seberapa besar atau tingkat kepentingan variabel prediktor tersebut berkontribusi dalam pembentukan pohon klasifikasi. Tingkat kepentingan variabel diperoleh dari penjumlahan nilai *goodness of split (improvement)* pada setiap node, dan kemudian diskalakan relatif terhadap variabel yang berkinerja terbaik (Steinberg, 2019).

$$skor = \sum_{i=1}^n \phi(s, t_i) \quad (2.4)$$

Variabel dengan skor 100 menunjukkan bahwa variabel tersebut memiliki nilai peningkatan tertinggi dan menjadi pemilah utama dalam pembentukan pohon klasifikasi. Suatu variabel dapat memperoleh nilai nol, apabila variabel tersebut tidak menjadi pemilah pada pohon klasifikasi.

b. Penentuan Simpul Terminal

Simpul terminal adalah simpul t yang tidak terdapat penurunan heterogenitas yang berarti, atau terdapat satu pengamatan di setiap simpul anak, atau terdapat batasan jumlah minimum n pengamatan di setiap simpul anak yang dihasilkan (Breiman, Friedman, Olshen, & Stone, 1993).

c. Penandaan Label Kelas

Penandaan label kelas berdasarkan aturan jumlah terbanyak yang ditunjukkan dalam persamaan 2.4.

$$p(j_i|t) = \max p(j_i|t) = \max \frac{N_j(t)}{N(t)} \quad (2.5)$$

keterangan:

$p(j_i|t)$ = Proporsi kelas j pada simpul t

$N_j(t)$ = Jumlah pengamatan kelas j pada simpul akhir t

$N(t)$ = Jumlah total pengamatan dalam simpul akhir.

2.2.2 Pemangkasan Pohon Klasifikasi (*Pruning*)

Pohon klasifikasi yang berukuran besar menyebabkan terjadi kasus *overfitting*, yakni nilai prediksi melebihi nilai sebenarnya. Untuk mengatasi kasus tersebut dilakukan pemangkasan pohon klasifikasi untuk mendapatkan ukuran pohon

yang layak, pemangkasan dilakukan dengan menggunakan ukuran *cost complexity pruning*.

Langkah pertama dalam tahap pemangkasan dilakukan pada subpohon dari pohon terbesar T_{\max} yang dinotasikan dengan T_1 . Nilai T_1 dari T_{\max} , didapatkan dari t_L dan t_R yang merupakan simpul anak kiri dan simpul anak kanan dari T_{\max} yang dihasilkan dari pemilahan pada setiap simpul induk t . Ketika terdapat dua simpul anak dan simpul induk yang memenuhi persamaan $R(t) = R(t_L) + R(t_R)$, simpul anak t_L dan t_R tersebut dipangkas. Proses ini akan terus diulangi sampai memberikan hasil tidak memungkinkannya lagi dilakukan pemangkasan. Pemangkasan dilakukan dengan persamaan berikut (Breiman, Friedman, Olshen, & Stone, 1993).

$$R_\alpha(T) = R(T) + \alpha |\tilde{T}| \quad (2.6)$$

dengan

α = parameter *complexity* (*Cost* bagi penambahan satu simpul akhir pada T)

$R(T)$ = penduga pengganti

$|\tilde{T}|$ = ukuran banyaknya *terminal node*

Parameter *complexity* menentukan pohon bagian $T(\alpha)$ yang dapat meminimumkan $R_\alpha(T)$ pada seluruh pohon bagian untuk setiap nilai α . Nilai parameter ini akan secara perlahan meningkat selama proses pemangkasan. Selanjutnya pencarian pohon bagian $\max T(\alpha) < T_{\max}$ yang dapat meminimumkan $R_\alpha(T_t)$ yakni sebagai berikut.

$$R_\alpha(T_t(\alpha)) = \min R_\alpha(T_t) \quad (2.7)$$

2.2.3 Penentuan Ukuran Pohon Klasifikasi Optimal

Semakin besar pohon klasifikasi maka tingkat kesalahan prediksinya juga akan semakin kecil, namun dapat menyebabkan nilai kompleksitas kesalahan yang tinggi. Sehingga perlu dipilih pohon klasifikasi optimal dengan berukuran sederhana namun memberikan nilai kesalahan prediksi yang cukup kecil (Therneau & Atkinson, 1997). Pohon klasifikasi optimal dinotasikan dengan T_t yaitu sebuah sub pohon terkecil dari T_{\max} yang meminimumkan

tingkat kesalahan prediksi. Struktur data yang digambarkan cenderung kompleks, maka perlu pemilihan pohon yang optimum dengan ukuran sederhana dan memberikan nilai penduga pengganti cukup kecil.

Salah satu metode yang dapat digunakan untuk mencari nilai penduga pengganti adalah *Test Sample Estimate*. Metode ini dilakukan dengan membagi sampel L menjadi dua himpunan, yaitu L_1 dan L_2 . Pengamatan L_1 merupakan data *training* digunakan untuk membentuk pohon T . Sedangkan pengamatan L_2 , data *testing* digunakan untuk menduga $R(T)$. Jika N_2 merupakan jumlah pengamatan L_2 , dan $X(\cdot)$ bernilai 0 jika pernyataan dalam kurung salah dan bernilai 1 jika pernyataan dalam tanda kurung benar. Penduga sampel uji dapat ditunjukkan dalam persamaan berikut.

$$R^{ts}(T_t) = \frac{1}{N_2} \sum_{(x_n, j_n) \in L_2}^N X(d(x_n) \neq j_n) \quad (2.8)$$

keterangan:

$R^{ts}(T_t)$ = Total proporsi kesalahan *test sample estimate*

N_2 = Jumlah pengamatan data *testing* (A_2)

$X(\cdot)$ = Bernilai 1 jika pernyataan $d(x_n) \neq j_n$ benar, dan bernilai 0 jika salah

Pohon klasifikasi dikatakan optimum apabila memenuhi kriteria $R^{ts}(T_t^*) = \min R^{ts}(T_t)$. Performa dari perbedaan ukuran pohon dapat diketahui dari nilai *relative cost*. *Relative cost* adalah nilai yang menunjukkan hubungan antara kesalahan klasifikasi dengan ukuran pohon. *Relative cost* bernilai antara 0 hingga 1. Pohon klasifikasi dengan nilai *relative cost* mendekati 0 merupakan pohon klasifikasi optimal (Anonim, 2018).

2.3 Evaluasi Ketepatan Klasifikasi

Penelitian ini menggunakan nilai akurasi (*total accuracy rate*), *sensitivity*, *specificity*, dan *Area Under Curve* (AUC) dalam melakukan evaluasi ketepatan klasifikasi. Nilai akurasi diperoleh dari proporsi observasi yang diprediksi secara tepat oleh fungsi

klasifikasi (Johnson & Winchern, 2007). *Sensitivity* menggambarkan proporsi kelas i atau proporsi yang benar-benar positif, dan *specificity* menggambarkan proporsi kelas j atau proporsi yang benar-benar negatif (Han & Kamber, 2006). Data yang digunakan merupakan data *unbalance* atau data memiliki jumlah pada setiap kelas yang berbeda jauh. Dalam perhitungannya, AUC mengakomodasi *sensitivity* dan *specificity* sehingga AUC dapat lebih baik untuk digunakan dalam mengukur ketepatan klasifikasi pada data *unbalance* daripada *total accuracy rate* (Bekkar, Djemaa, & Alitouche, 2013)

Nilai AUC digunakan untuk menunjukkan kekuatan hasil diagnostik atau klasifikasi. Kekuatan klasifikasi dikategorikan menjadi sangat buruk, buruk, sedang, baik, dan sangat baik (Dahlan, 2009). Tabel 2.2 menunjukkan rentang nilai dari masing-masing kategori.

Tabel 2.2 Kategori Kekuatan Nilai AUC

Rentang Nilai	Interpretasi
50-60%	Sangat Buruk
61-70%	Buruk
71-80%	Sedang
81-90%	Baik
91-100%	Sangat Baik

Persamaan yang digunakan dalam perhitungan nilai akurasi, *sensitivity*, *specificity*, dan AUC dapat dilihat pada persamaan 2.9-2.13.

Tabel 2.3 Crosstab Ketepatan Klasifikasi

Kelas Pengamatan Y	Kelompok Prediksi		Total
	1	2	
1	n_{11}	n_{12}	$N_{1.}$
2	n_{21}	n_{22}	$N_{2.}$
Total	$N_{.1}$	$N_{.2}$	N

$$Akurasi = \frac{n_{11} + n_{22}}{N} \quad (2.9)$$

$$Sensitivity = \frac{n_{11}}{N_{1.}} \quad (2.10)$$

$$Specificity = \frac{n_{22}}{N_2} \quad (2.11)$$

$$AUC = \frac{1}{2}(sensitivity + specificity) \quad (2.13)$$

Keterangan :

n_{11} = Jumlah pengamatan variabel Y kelas 1 tepat di klasifikasi sebagai variabel Y kelas 1

n_{21} = Jumlah pengamatan variabel Y kelas 2 salah di klasifikasi sebagai variabel Y kelas 1

n_{12} = Jumlah pengamatan variabel Y kelas 1 salah di klasifikasi sebagai variabel Y kelas 2

n_{22} = Jumlah pengamatan variabel Y kelas 2 tepat di klasifikasi sebagai variabel Y kelas 2

N_1 = Jumlah pengamatan dari variabel Y kelas 1

N_2 = Jumlah pengamatan dari variabel Y kelas 2

$N_{.1}$ = Jumlah prediksi dari variabel Y kelas 1

$N_{.2}$ = Jumlah prediksi dari variabel Y kelas 2

N = Jumlah total pengamatan / prediksi

(Halaman ini sengaja dikosongkan)

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder, yakni data tes tulis tahun 2018 yang diperoleh dari Kajibag SNMPTN-SBMPTN 2018. Tidak semua data peserta tes tulis yang digunakan, melainkan data peminat prodi Statistika Matematika, dan Aktuaria ITS, IPB, dan UI. Jumlah data pada masing-masing prodi setiap PTN dapat dilihat pada Tabel 3.1.

Tabel 3.1 Jumlah Data Penelitian

Program Studi	ITS	IPB	UI
Statistika	834	876	751
Matematika	603	505	831
Aktuaria	1066	1329	2025

Struktur data yang digunakan dalam penelitian dapat disajikan pada Tabel 3.2 dengan N adalah jumlah data pada setiap prodi.

Tabel 3.2 Struktur Data Penelitian

No.	X_1	X_2	X_3	...	X_{12}	Y
1	X_{11}	X_{21}	X_{31}	...	X_{121}	Y_1
2	X_{12}	X_{22}	X_{32}	...	X_{122}	Y_2
.
.
.
N	X_{1N}	X_{2N}	X_{3N}	...	X_{12N}	Y_N

3.2 Variabel Penelitian

Pada penelitian ini menggunakan 1 variabel respon, yakni status penerimaan peserta, yakni diterima atau tidak diterima. Variabel prediktor berupa nilai 10 mata uji dalam tes tulis masuk PTN rumpun saintek yang terdiri dari enam nilai TKPA dan empat nilai TKD Saintek. Nilai tersebut merupakan nilai yang telah diolah dengan menggunakan IRT. Secara keseluruhan variabel yang digunakan dapat disajikan pada Tabel 3.3.

Tabel 3.3 Variabel Penelitian untuk Analisis *Classification Tree*

Variabel	Nama Variabel	Skala	Keterangan
Y	Status Penerimaan Peserta	Nominal	0 = Tidak Diterima 1 = Diterima
X_1	Nilai Matematika Dasar	Rasio	-
X_2	Nilai Bahasa Indonesia	Rasio	-
X_3	Nilai Bahasa Inggris	Rasio	-
X_4	Nilai TPA Verbal	Rasio	-
X_5	Nilai TPA Numerikal	Rasio	-
X_6	Nilai TPA Figural	Rasio	-
X_7	Nilai Matematika IPA	Rasio	-
X_8	Nilai Biologi	Rasio	-
X_9	Nilai Kimia	Rasio	-
X_{10}	Nilai Fisika	Rasio	-

Selain variabel penelitian diatas, digunakan pula variabel lain untuk mengetahui karakteristik dan deskripsi peserta tes tulis masuk PTN 2018 peminat prodi Statistika dan Matematika ITS, IPB dan UI.

Tabel 3.4 Variabel Penelitian untuk Analisis Statistika Deskriptif

Variabel	Skala
Jenis Kelamin	Rasio
Asal SMA	Rasio

3.3 Langkah Analisis

Adapun langkah-langkah yang digunakan dalam penelitian ini adalah sebagai berikut.

1. Melakukan *screening* data, yakni melakukan pemilihan data dengan ketentuan peserta memilih program studi Statistika, Matematika, dan Aktuaria baik pada pilihan 1, 2, dan 3.
2. Pemecahan data dengan kategori Universitas berupa ITS, IPB, dan UI.
3. Mendeskripsikan data peserta tes tulis 2018 prodi Statistika, Matematika, dan Aktuaria ITS, IPB, dan UI dengan menggunakan statistika deskriptif untuk mengetahui

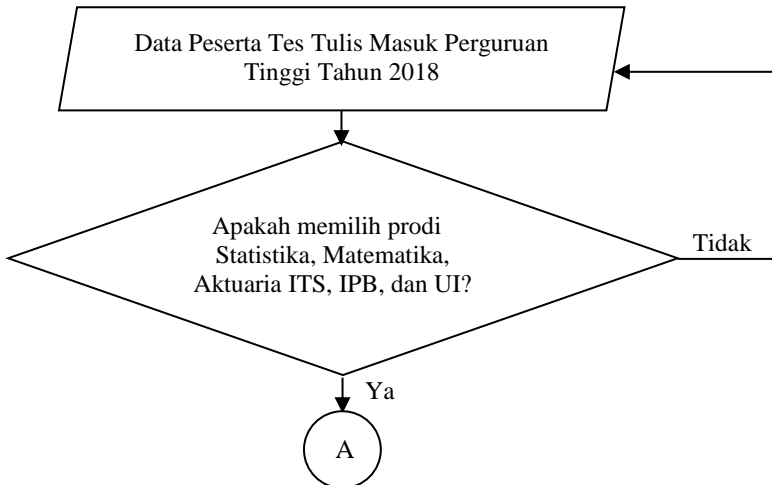
karakteristik peminat setiap program studi dan PTN berdasarkan status penerimaan, jenis kelamin, asal SMA dan nilai tes tulis.

4. Melakukan analisis batas nilai minimal mata uji yang berpengaruh signifikan dalam penerimaan peserta tes tulis masuk perguruan tinggi pada pilihan program studi Statistika, Matematika, dan Aktuaria ITS, IPB, dan UI dengan pendekatan *Classification tree*. Langkah-langkah analisis diladengan langkah-langkah sebagai berikut.
 - a. Membentuk pohon klasifikasi maksimal dengan menggunakan data *training* dengan jumlah $2/3$ dari data setiap prodi dan PTN. Tahapan dalam pembentukan pohon klasifikasi optimal adalah:
 - i. Melakukan pemilahan berdasarkan aturan pemilah indeks Gini, kemudian dilakukan evaluasi dengan menggunakan *goodness of split*. Tahapan dalam pemilahan adalah (1) Menghitung indeks Gini variabel respon (status penerimaan). (2) Menentukan nilai *threshold* masing-masing variabel prediktor, dengan cara mengurutkan nilai dari terendah hingga terbesar setiap variabel prediktor, kemudian menghitung nilai rata-rata dari kedua nilai yang bernilai berbeda pada data yang telah diurutkan. Jumlah kemungkinan pemilahan yang terjadi atau jumlah nilai *threshold* adalah $n-1$, dimana n merupakan perhitungan data dengan nilai yang berbeda. (3) Menghitung nilai indeks Gini dan *goodness of split* untuk simpul anak kanan dan kiri, pada masing-masing *threshold* setiap variabel prediktor. (4) Memilih nilai *goodness of split* tertinggi diantara semua kemungkinan yang ada, untuk dijadikan sebagai pemilah
 - ii. Menentukan *terminal nodes* dengan penghentian pembentukan pohon hingga mencapai jumlah pengamatan minimum pada setiap *terminal nodes*.

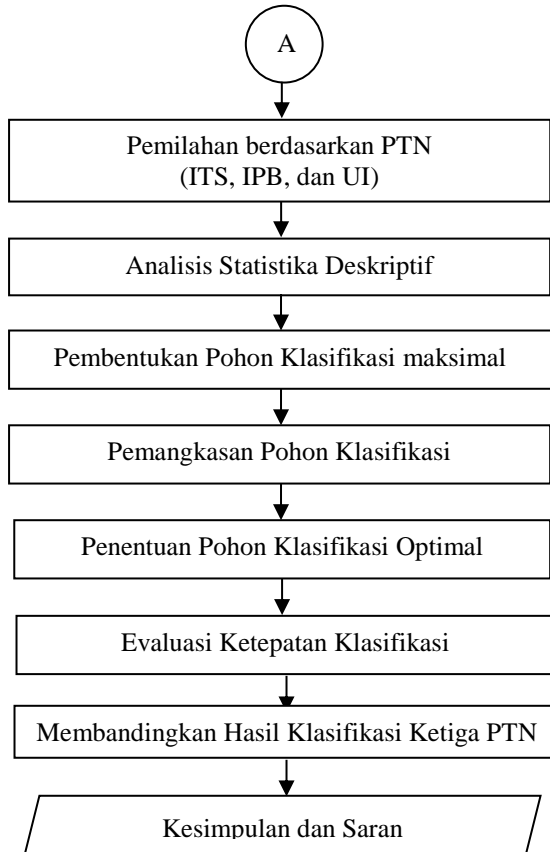
- iii. Melakukan penandaan label kelas pada setiap *terminal nodes* yang terbentuk.
 - b. Melakukan pemangkasan pohon (*pruning*) berdasarkan *criteria cost complexity minimum*
 - c. Menentukan pohon klasifikasi optimal
 - d. Mengevaluasi ketepatan klasifikasi dengan menggunakan ukuran nilai akurasi, *sensitivity*, *specificity*, dan *AUC* dengan menggunakan data *testing*.
5. Membandingkan batasan hasil *Classification Tree* diantara ketiga PTN pada prodi Statistika, Matematika, dan Aktuaria.
 6. Menarik kesimpulan dan saran.

3.4 Diagram Alir

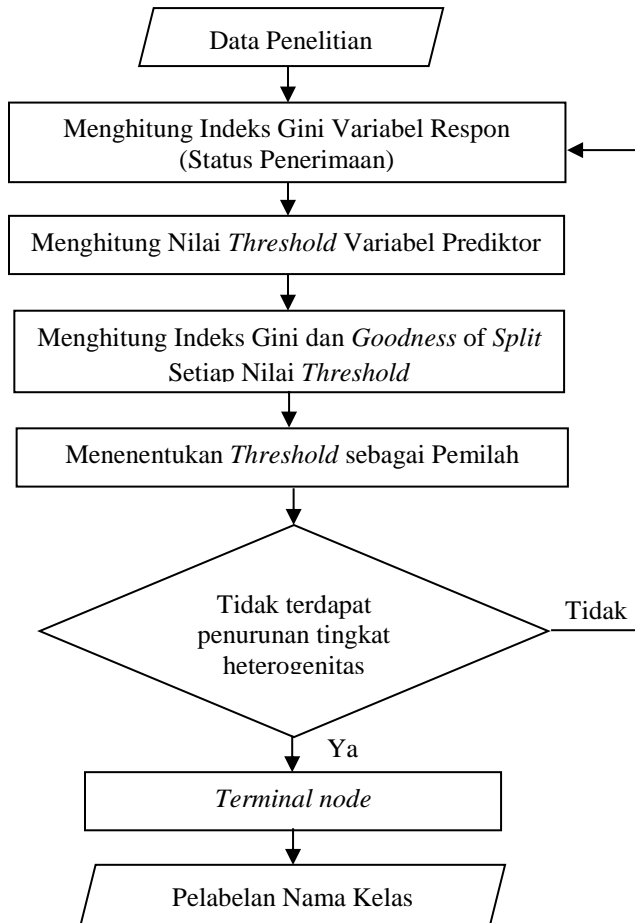
Berdasarkan langkah analisis yang telah dijelaskan sebelumnya, dapat digambarkan bentuk diagram alir penelitian pada Gambar 3.1.



Gambar 3.1 Diagram Alir Penelitian



Gambar 3.1 Diagram Alir Penelitian (lanjutan)



Gambar 3.2 Diagram Alir Pembentukan Pohon Klasifikasi Maksimal

BAB IV

ANALISIS DAN PEMBAHASAN

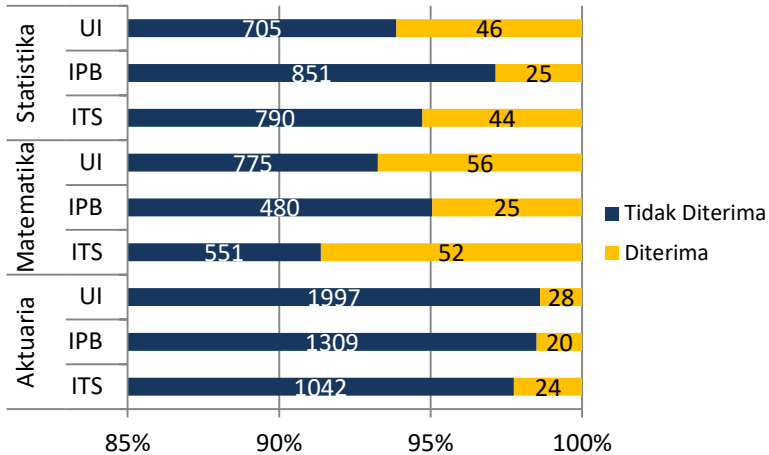
Penelitian ini bertujuan untuk mengetahui mata uji apa saja yang dapat menentukan status penerimaan peserta beserta dengan batas minimal nilai mata uji tes tulis masuk Perguruan Tinggi Negeri tahun 2018. Penentuan batasan dilakukan dengan menggunakan pendekatan *Classification Tree* dengan status penerimaan sebagai penentu klasifikasi.

4.1 Deskripsi Peminat Prodi Statistika, Matematika, dan Aktuaria di ITS, IPB dan UI

Karakteristik peminat program studi Statistika, Matematika, dan Aktuaria ITS, IPB dan UI dapat diketahui dari data peserta melalui eksplorasi data. Eksplorasi data dilakukan berdasarkan status penerimaan, jenis kelamin, dan asal sekolah.

4.1.1 Deskripsi Peminat Berdasarkan Status Penerimaan, Jenis Kelamin, dan Asal Sekolah

Perbandingan jumlah peminat pada setiap prodi dan PTN berdasarkan status penerimaan peserta tes tulis masuk PTN 2018 dapat diketahui pada Gambar 4.1. Jumlah peserta tes tulis masuk PTN tahun 2018 dengan pilihan program studi Statistika, Matematika, dan Aktuaria di Institut Teknologi Sepuluh Nopember (ITS), Institut Pertanian Bogor (IPB), dan Universitas Indonesia (UI). Diantara ketiga prodi, peminat dengan jumlah terbanyak adalah prodi Aktuaria. Banyaknya peminat Aktuaria disebabkan karena lulusan Aktuaria banyak dicari oleh perusahaan. Hal ini didukung dengan adanya Keputusan Menteri Keuangan Republik Indonesia No. 426/KMK.06/2003, dalam pasal 3 ayat 16 yang menyatakan bahwa setiap perusahaan asuransi jiwa harus mengangkat minimal 1 Aktuaris yang sudah tersertifikasi (Menkeu, 2003). Adanya keputusan tersebut membuat prospek kerja lulusan Aktuaria lebih menjamin dibandingkan dengan Statistika dan Matematika.

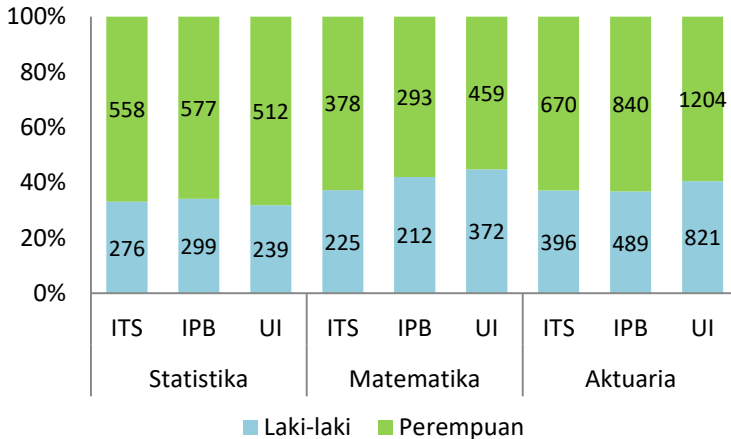


Gambar 4.1 Perbandingan Jumlah Peserta Diterima dan Tidak Diterima

Program studi Statistika IPB memiliki jumlah peminat paling banyak pada jalur tes tulis dibandingkan dengan ITS dan UI. Namun demikian, Statistika IPB memiliki tingkat ketetapan yang lebih tinggi dibandingkan lainnya. Hal ini dapat diketahui pada Gambar 4.1 yang menunjukkan bahwa Statistika IPB menerima 25 peserta dari 876 peminat, atau dengan kata lain 1 kursi (kuota penerimaan) direbutkan oleh 35 peserta. Ketetapan yang tinggi ditunjukkan pula pada prodi Matematika IPB, walaupun apabila ditinjau dari jumlah peminat Matematika IPB lebih rendah dibandingkan dengan ITS dan UI, Tingkat keketatan Matematika IPB adalah 1:20.

Pelaksanaan tes tulis masuk PTN tahun 2018, terdapat 4 PTN yang membuka prodi Aktuaria yakni IPB, UI, ITS, dan Universitas Padjajaran (UNPAD). Gambar 4.1 menunjukkan bahwa Aktuaria UI memiliki jumlah peminat paling banyak dibandingkan IPB dan ITS, dengan perbandingan antara jumlah peserta yang diterima dengan peminat adalah 1:73. Artinya bahwa 1 kuota penerimaan yang prodi Aktuaria UI diperebutkan oleh 73 peserta tes tulis pada tahun 2018.

Apabila jumlah peminat dilihat berdasarkan jenis kelamin perempuan dan laki-laki, perbandingan jumlah keduanya dapat dilihat pada Gambar 4.2

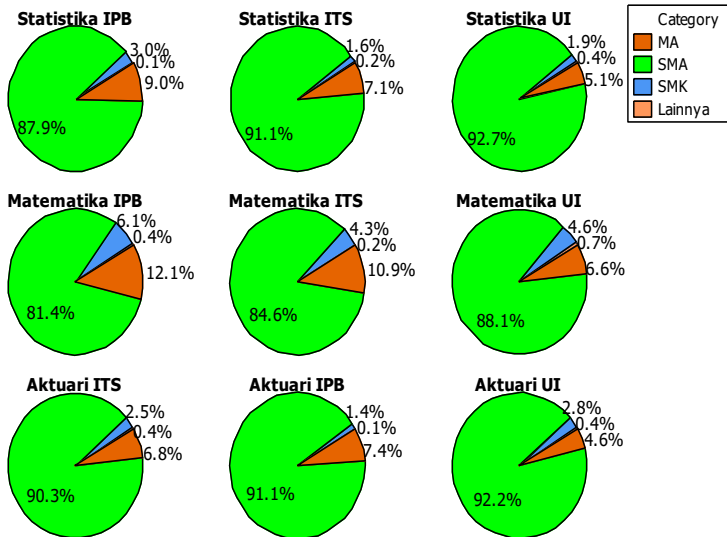


Gambar 4.2 Perbandingan Jumlah Peminat Prodi Statistika, Matematika, dan Aktuaria Berdasarkan Jenis Kelamin

Jumlah peminat prodi Statistika, Matematika, dan Aktuaria pada tes tulis masuk PTN 2018 berdasarkan jenis kelamin dapat diketahui pada Gambar 4.2. Peminat ketiga prodi secara keseluruhan didominasi oleh jenis kelamin perempuan di ketiga PTN. Prodi Statistika pada ketiga PTN memiliki presentase peminat dengan jenis kelamin perempuan hampir sama, yakni berada pada nilai 65%. Peminat Matematika ITS merupakan peminat dengan jenis kelamin perempuan tertinggi diantara IPB dan UI, yakni dengan persentase 63%. Peminat program studi Aktuaria ITS dan IPB memiliki nilai persentase peminat perempuan yang sama, yakni pada nilai 62%. Sedangkan peminat Aktuaria UI berjenis kelamin perempuan memiliki persentase dengan selisih 3% dari peminat Aktuaria ITS dan UI.

Penerimaan mahasiswa dibuka untuk semua asal sekolah Menengah Atas (SMA), Madrasah Aliyah (MA), Sekolah Menengah Kejuruan (SMK) baik negeri maupun swasta. Gambar

4.3 menyajikan persentase jumlah peminat Statistika, Matematika, dan Aktuaria ITS, IPB dan UI berdasarakan asal sekolah. Asal sekolah dikategorikan menjadi 4 yakni, MA, SMA, SMK, dan lainnya. Peserta yang termasuk asal sekolah lainnya merupakan peserta dengan keterangan berasal dari sekolah luar negeri dan paket C.



Gambar 4.3 Persentase Jumlah Peminat Berdasarkan Asal Sekolah

Peminat Statistika, Matematika, dan Aktuaria baik di ITS, IPB, maupun UI didominasi oleh peserta yang berasal dari SMA. Dominasi peminat yang berasal dari SMA disebabkan karena kedua prodi merupakan rumpun prodi Saintek, sehingga lebih banyak menerima siswa dengan asal sekolah yang terdapat jurusan IPA dibandingkan dengan sekolah kejuruan (SMK). Jurusan IPA dapat ditemui pada sekolah MA atau SMA. Namun demikian, jumlah SMA di Indonesia baik negeri maupun swasta lebih banyak dibandingkan dengan jumlah MA. Antusias peminat yang berasal dari SMK tetap ada dan tidak menutup kemungkinan bahwa peserta yang diterima berasal dari SMK. Selain berasal

dari SMA, MA, dan SMK, terdapat pula peminat yang berasal dari luar negeri dan peminat dari lulusan paket C yang termasuk dalam kategori lainnya dengan persentase kurang dari 1% pada masing-masing prodi dan PTN.

Tabel 4.1 Jumlah Peserta Diterima Berdasarkan Asal Sekolah

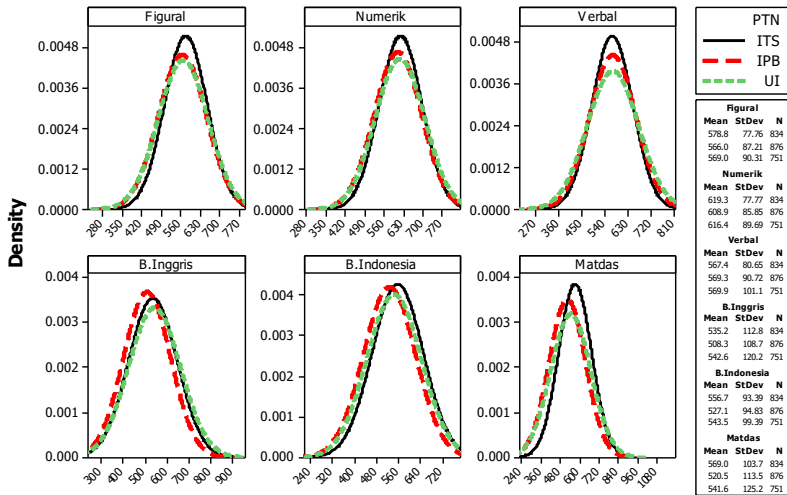
Prodi	Asal	ITS	IPB	UI
Statistika	MA	2	2	1
	SMA	42	23	45
	Lainnya	1	0	0
Matematika	MA	5	2	3
	SMA	47	22	53
	SMK	0	1	0
Aktuaria	MA	1	1	0
	SMA	23	19	28

Tabel 4.1 menunjukkan jumlah peserta tes tulis tahun 2018 yang diterima pada program studi Statistika, Matematika, dan Aktuaria ITS, IPB, dan UI berdasarkan asal sekolah. Peserta yang diterima pada ketiga prodi tersebut tetap didominasi oleh peserta yang berasal dari SMA. Ketiga prodi menunjukkan bahwa peserta yang diterima berasal dari MA dan SMA. Namun pada Statistika ITS terdapat satu peserta yang dikategorikan lainnya, dengan keterangan asal sekolah paket C dan prodi Matematika IPB menerima 1 peserta yang berasal dari SMK.

4.1.2 Karakteristik Nilai Peminat Statistika, Matematika, dan Aktuaria ITS, IPB, dan UI

1. Program Studi Statistika

Mata uji tes tulis masuk perguruan tinggi negeri pada kelompok Sainstek terdiri dari dua tes, yakni Tes Kemampuan dan Potensi Akademik (TKPA) dan Tes Kemampuan Dasar (TKD) Sainstek. Mata uji yang termasuk dalam TKPA adalah Tes Potensial Akademik (TPA), Bahasa Inggris, Bahasa Indonesia dan Matematika Dasar, sedangkan TKD Sainstek terdiri dari mata uji Matematika IPA, Biologi, Fisika, dan Kimia. Gambar 4.4 menunjukkan karakteristik nilai TKPA peminat Statistika.



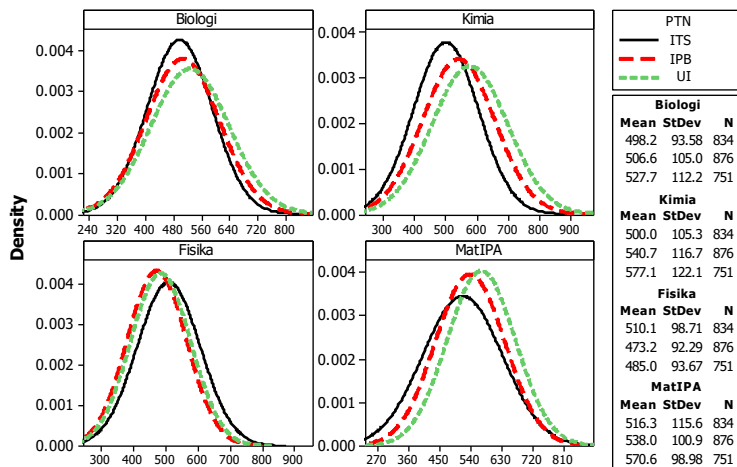
Gambar 4.4 Density Plot Nilai TKPA Peminat Statistika

TPA menunjukkan kemampuan pemahaman dan penalaran terbagi. TPA dibagi menjadi 3 tes, yakni Figural (gambar), Numerik (angka), dan Verbal (bahasa). Berdasarkan nilai rata-rata, Figural dan Numerik peminat Statistika lebih tinggi dibandingkan dengan IPB dan UI. Sedangkan untuk verbal, peminat Statistika UI memiliki nilai rata-rata lebih tinggi dibandingkan lainnya. Pada mata uji Figural terlihat bahwa nilai peminat ITS sedikit skewness kiri, artinya bahwa modus atau nilai yang sering muncul lebih besar dari nilai rata-rata, yang artinya bahwa nilai Figural peminat ITS cenderung tinggi. Kurva peminat Statistika UI pada ketiga nilai TPA menunjukkan kurva lebih rendah dan lebih lebar dibandingkan dari peminat ITS dan IPB. Artinya bahwa keragaman nilai TPA pada peminat Statistika UI lebih tinggi.

Pada nilai Bahasa Inggris terlihat bahwa ketiga kurva cenderung condong ke kanan (*skewness* kiri), yang berarti bahwa nilai Bahasa Inggris pada ketiga PTN memiliki modus yang bernilai lebih kecil dibandingkan rata-rata, atau dapat dikatakan bahwa nilai Bahasa Inggris cenderung rendah. Nilai Bahasa

Inggris, Bahasa Indonesia, dan Matematika Dasar pada peminat Statistika IPB menunjukkan kurva sedikit bergeser ke kiri dari kurva peminat ITS dan UI. Hal ini dapat diartikan bahwa rata-rata nilai Bahasa Inggris, Bahasa Indonesia dan Matematika Dasar peminat Statistika IPB pada tes tulis paling rendah dibandingkan ITS dan UI.

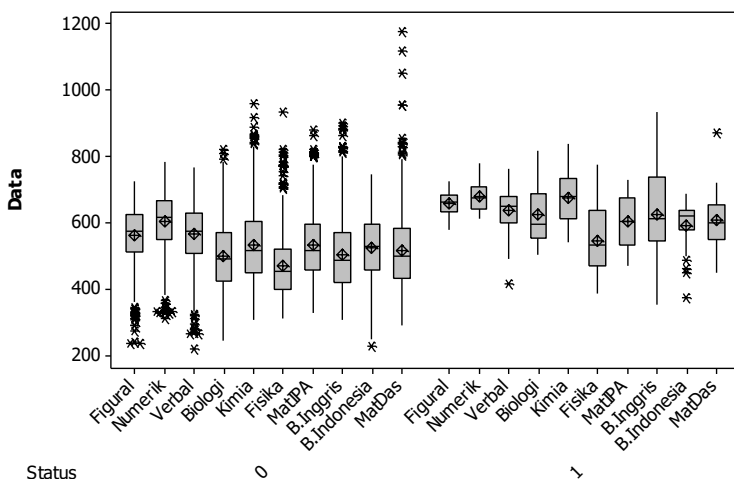
Selanjutnya akan dilakukan analisis nilai peminat Statistika ITS, IPB, dan UI berdasarkan nilai TKD Sainstek. Gambar 4.5 menunjukkan hasil plot dari nilai TKD Sainstek peminat Statistika pada tes tulis masuk PTN tahun 2018.



Gambar 4.5 Density Plot Nilai TKD Sainstek Peminat Statistika

Tes Kemampuan Dasar Sainstek terdiri dari mata uji Biologi, Kimia, Fisika, dan Matematika IPA. Plot pada variabel nilai Biologi dan Kimia menunjukkan bahwa kurva peminat Statistika UI (hijau) memiliki nilai rata-rata yang lebih tinggi dibandingkan kedua PTN lainnya. Namun kurva yang lebar dan rendah menunjukkan bahwa peminat Statistika UI memiliki nilai Kimia dan Biologi dengan keragaman yang tinggi. Nilai Fisika pada peminat Statistika IPB dan UI menunjukkan kurva yang hampir berhimpit, hal ini menunjukkan adanya sedikit kemiripan,

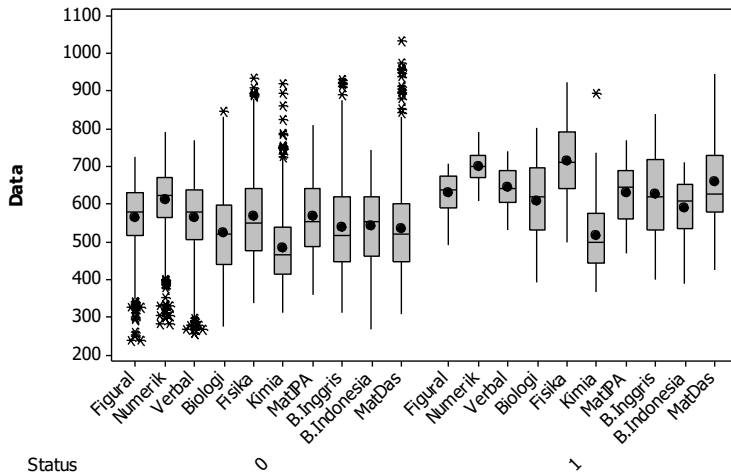
peserta yang tidak diterima menyebabkan nilai variansi pada setiap mata uji bernilai lebih besar daripada nilai peserta yang diterima. Hal tersebut dapat dilihat pada panjang *box* pada masing-masing nilai mata uji peserta yang tidak diterima cenderung lebih panjang dibandingkan dengan peserta yang diterima, kecuali untuk mata uji Fisika dan Bahasa Inggris peserta yang diterima memiliki nilai variansi yang lebih tinggi.



Gambar 4.7 Box Plot Nilai Mata Uji Tes Tulis Peminat Statistika IPB

Karakteristik nilai peminat Statistika IPB yang tidak diterima maupun diterima memiliki karakteristik yang hampir sama dengan nilai peminat Statistika ITS. Nilai rata-rata dan nilai median peserta yang diterima lebih tinggi dibandingkan dengan peminat yang tidak diterima. Hal ini dikarenakan nilai peminat yang diterima memiliki sedikit nilai *outlier* dibandingkan dengan yang tidak diterima. Selain itu, keberagaman nilai pada peminat yang diterima lebih terpusat, hal ini terlihat pada panjang *box* pada peminat yang diterima lebih pendek dibandingkan dengan peminat tidak diterima, kecuali pada mata uji Bahasa Inggris dan Fisika. Kedua mata uji tersebut memiliki *box* yang lebih

panjang pada peminat yang diterima. Nilai rata-rata tertinggi untuk peminat yang diterima berada pada nilai Numerik 682,02, nilai tersebut tidak berbeda jauh dengan peminat Statistika ITS yang diterima.



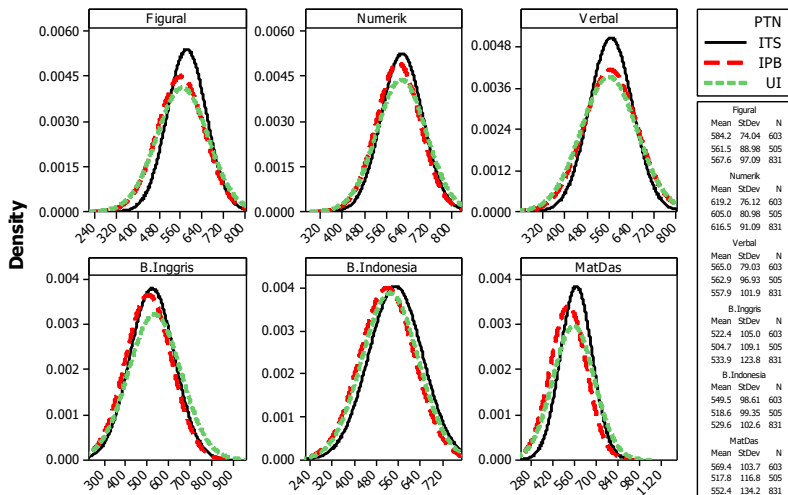
Gambar 4.8 Box Plot Nilai Mata Uji Tes Tulis Peminat Statistika UI

Karakteristik nilai pada peminat Statistika UI menunjukkan bahwa peminat yang diterima tidak terdapat nilai *outlier*, kecuali pada mata uji Kimia yang memiliki satu nilai *outlier*. Tidak adanya nilai *outlier* dapat diartikan bahwa nilai mata uji pada peminat yang diterima memiliki nilai yang hampir sama atau nilai antar peminat memiliki selisih yang tidak terlalu besar. Adanya outlier pada nilai mata uji Kimia peminat yang diterima, menyebabkan nilai rata-rata dan median bernilai paling rendah dibandingkan mata uji lain, yakni bernilai 516,7 dan 499,4. Nilai rata-rata tertinggi peminat yang diterima Statistika UI adalah nilai Fisika dengan nilai 712,4, berbeda dengan Statistika ITS dan IPB yang menunjukkan nilai rata-rata tertinggi pada nilai Numerik. Karakteristik peminat yang tidak diterima memiliki nilai median dan rata-rata yang hampir sama. Nilai tertinggi pada peminat yang tidak diterima berada pada nilai Matematika Dasar yang

mencapai nilai 1030,71. Nilai tertinggi yang ada pada Matematika Dasar menyebabkan nilai variansi yang tinggi dibandingkan pada mata uji lainnya. Nilai terendah peminat yang tidak diterima berada pada nilai Figural dengan nilai 239,76. Secara keseluruhan peminat yang tidak diterima memiliki keberagaman nilai yang lebih besar dibandingkan dengan peminat yang diterima, hal ini terlihat dari banyaknya jumlah *outlier* dan ukuran *box* yang lebih panjang.

2. Program Studi Matematika

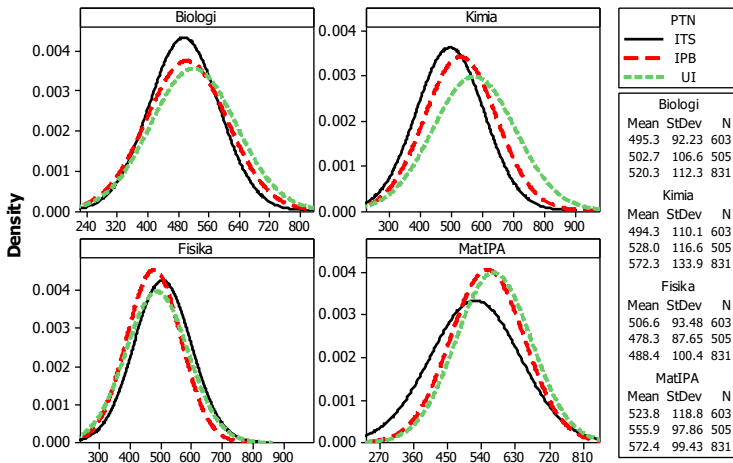
Penerimaan mahasiswa baru melalui jalur tes tulis masuk PTN berdasarkan pada nilai tes tulis yang telah dilaksanakan. Semakin tinggi atau semakin baik nilai dapat memengaruhi status penerimaan peserta tersebut. Nilai TKPA peminat Matematika ITS, IPB dan UI secara keseluruhan (peminat diterima dan tidak) dapat dilihat pada Gambar 4.9.



Gambar 4.9 Density Plot Nilai TKPA Peminat Matematika

Nilai TPA, yakni Figural, Numerik, dan Verbal pada peminat Matematika ITS memiliki nilai rata-rata yang lebih

tinggi. Selain itu kurva hitam pada ketiga nilai lebih tinggi dibandingkan dengan kurva merah (peminat Matematika IPB) dan kurva hijau (peminat UI). Hal ini menunjukkan bahwa peminat Matematika ITS dengan rentang nilai disekitar nilai rata-rata memiliki jumlah yang lebih banyak. Artinya bahwa peminat Matematika ITS memiliki nilai TPA atau kemampuan penalaran yang lebih tinggi dibandingkan IPB dan UI. Selajutnya apabila dilihat dari nilai Matematika Dasar dan Bahasa Indonesia peminat Matematika ITS memiliki nilai rata-rata yang lebih tinggi dengan keragaman yang rendah. Namun dalam hal kemampuan Bahasa Inggris, peminat Matematika ITS masih tertinggal dari peminat UI. Terlihat pula bahwa kurva peminat Matematika ITS pada nilai Bahasa Inggris cenderung condong kanan, yang dapat diartikan bahwa nilai yang sering diperoleh peminat Matematika ITS kurang dari nilai rata-rata.

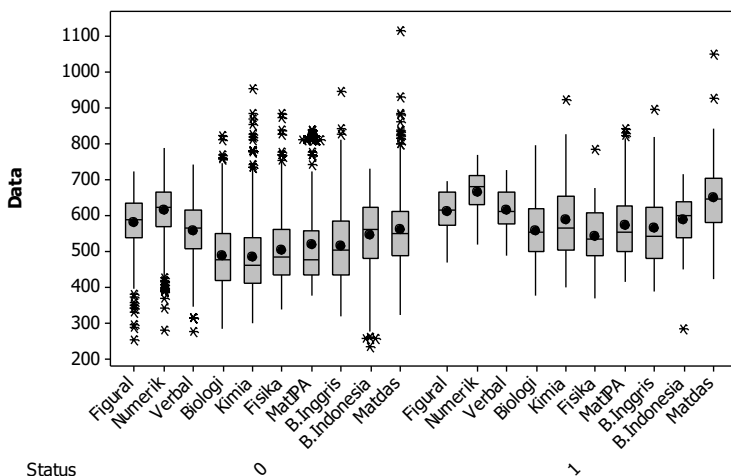


Gambar 4.10 *Density Plot* Nilai TKD Saintek Peminat Matematika

Apabila ditinjau dari nilai TKD Saintek peminat Matematika yang ditunjukkan pada Gambar 4.10, secara keseluruhan nilai TKD dengan rata-rata tertinggi berada pada

peminat Matematika UI, kecuali pada nilai Fisika. Diantara keempat nilai, nilai Kimia merupakan nilai dengan karakteristik yang berbeda diantara ketiga PTN. Hal ini terlihat bahwa ketiga kurva tidak saling berhimpit, melainkan saling bergeser.

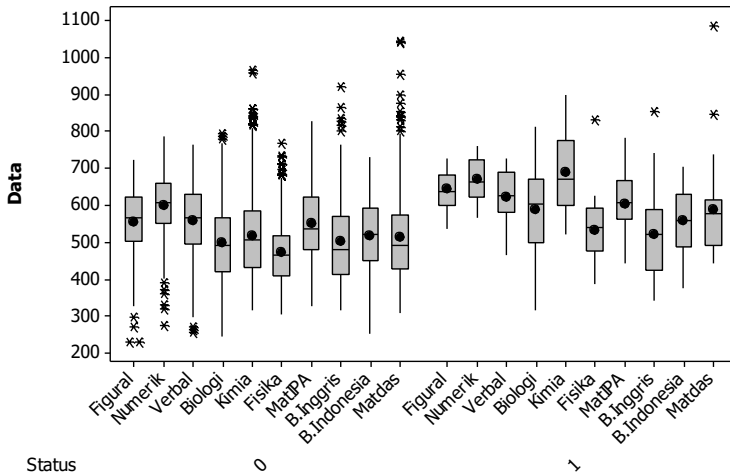
Karakteristik nilai 10 mata uji peminat program studi Matematika ITS, IPB dan UI pada tes tulis masuk PTN tahun 2018 berdasarkan status penerimaan dapat diuraikan sebagai berikut.



Gambar 4.11 Box Plot Nilai Mata Uji Tes Tulis Peminat Matematika ITS

Peminat Matematika ITS yang tidak diterima memiliki nilai median dan rata-rata yang hampir sama pada mata uji Biologi, Kimia, dan Fisika. Semua mata uji pada nilai peminat yang tidak diterima memiliki nilai *outlier* yang cukup banyak pada setiap mata uji, hal ini menyebabkan variansi atau penyebaran data bernilai besar. Berbeda dengan nilai peminat yang diterima, memiliki beberapa nilai *outlier* di semua mata uji. Bahkan untuk mata uji TPA dan Biologi memiliki tidak terdapat *outlier*. Artinya penyebaran data atau nilai peminat yang diterima pada setiap mata uji terpusat pada suatu nilai. Gambar 4.11 menunjukkan bahwa kedua status penerimaan memiliki nilai tertinggi pada

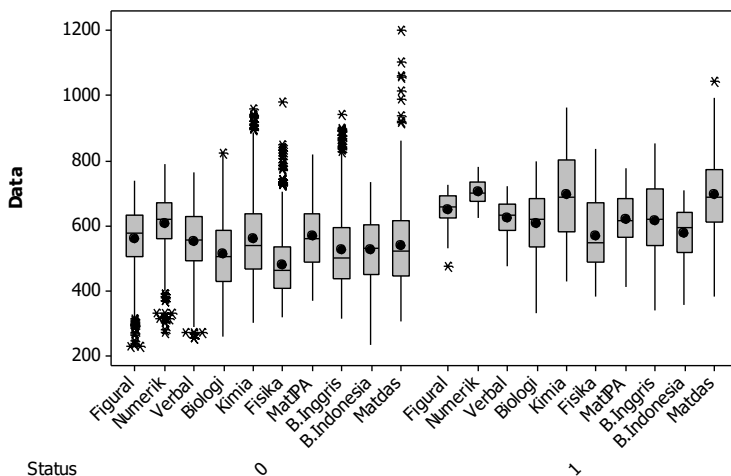
mata uji Matematika dasar dan nilai terendah adalah Bahasa Indonesia. Nilai Numerik merupakan mata uji dengan nilai rata tertinggi pada kedua status.



Gambar 4.12 Box Plot Nilai Mata Uji Tes Tulis Peminat Matematika IPB

Gambar 4.12 menunjukkan bahwa nilai peminat Matematika IPB yang tidak diterima memiliki nilai rata-rata dan median yang lebih rendah dibandingkan peminat yang diterima. Hal ini dapat dipengaruhi karena adanya *outlier* yang berjumlah cukup banyak pada nilai peminat yang tidak diterima. Selain itu variansi atau keberagaman data yang cukup besar yang ditunjukkan dengan ukuran *box* yang panjang dapat menyebabkan rendahnya nilai rata-rata pada peminat yang tidak diterima. Peminat yang diterima Matematika IPB memiliki karakteristik nilai rata-rata dan median yang hampir sama, yang dapat ditunjukkan pada Gambar 4.6 dimana simbol titik yang mendekati dan bahkan terletak diatas garis tengah pada *box*. Nilai rata-rata tertinggi peminat yang diterima berada pada mata uji Kimia dengan nilai 689,7 dan nilai terendah adalah mata uji Bahasa Inggris. Banyak peminat yang diterima memiliki nilai Kimia, Matematika IPA, Bahasa Inggris dan Matematika Dasar yang

lebih besar dari nilai rata-rata. Hal tersebut ditunjukkan pada garis *box* (*whisker*) diatas yang lebih panjang.

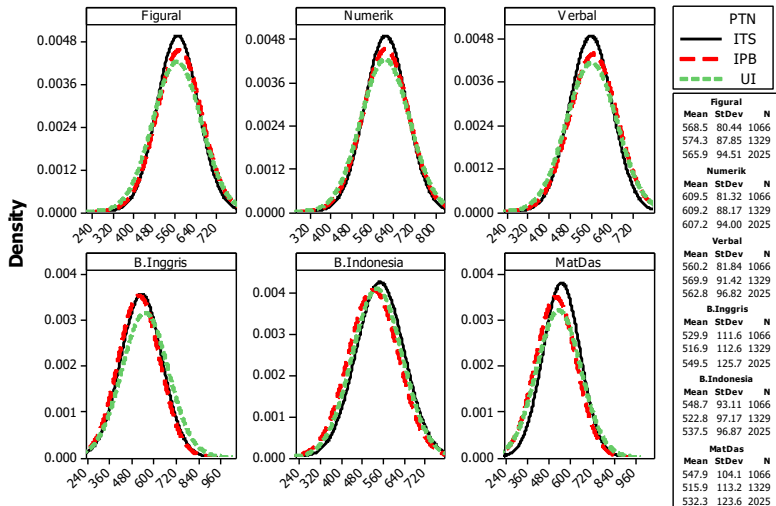


Gambar 4.13 Box Plot Nilai Mata Uji Tes Tulis Peminat Matematika UI

Peminat yang diterima pada Matematika UI memiliki nilai rata-rata tertinggi pada mata uji Numerik dengan nilai 704,91 dan terendah adalah nilai Fisika sebesar 571,2. Secara keseluruhan, peminat yang diterima cenderung memiliki nilai berada dibawah rata-rata, kecuali pada mata uji Fisika dan Matematika Dasar. Hal ini dapat ditunjukkan pada garis dibawah *box* pada masing-masing mata uji yang lebih panjang dibandingkan dengan garis diatas *box*. Penyebaran data yang cukup besar berada pada nilai Kimia. Peminat yang tidak diterima memiliki tingkat keberagaman yang hampir sama pada setiap mata uji, karena ukuran *box* yang hampir sama antara satu mata uji dengan yang lain, dan secara keseluruhan tingkat keberagaman peminat yang tidak diterima lebih besar dibandingkan dengan peminat yang diterima. Pada kedua status penerimaan menunjukkan nilai maksimal berada pada mata uji Matematika Dasar, dan nilai terendah adalah Fisika.

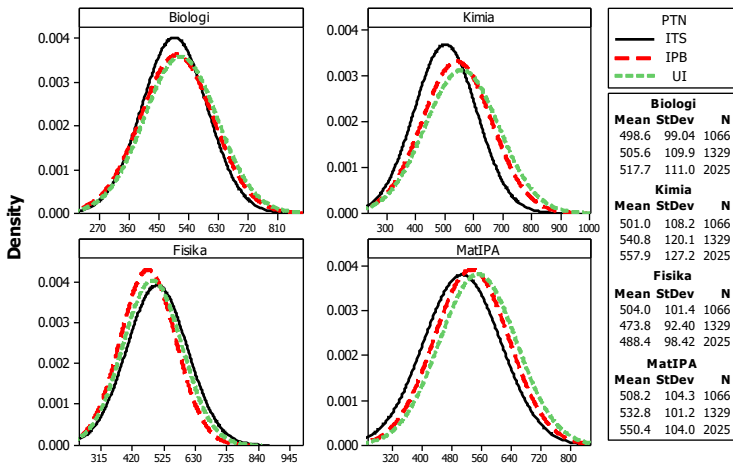
1. Program Studi Aktuaria

Karakteristik nilai tes tulis tahun 2018 peminat Aktuaria ITS, IPB dan UI pada kelompok nilai TKPA dapat dilihat pada Gambar 4.14.



Gambar 4.14 Density Plot Nilai TKPA Peminat Aktuaria

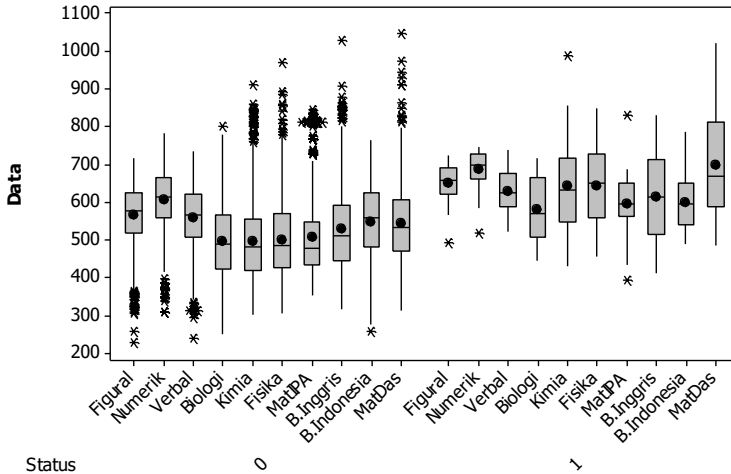
Karakteristik ketiga nilai TPA (Figural, Verbal, dan Numerik) pada ketiga peminat menunjukkan kemiripan karakteristik. Terlihat bahwa ketiga garis pada grafik Figural, Verbal maupun Numerik saling berhimpitan. Selain itu, peminat ITS menunjukkan grafik yang lebih tinggi dibandingkan IPB dan UI, artinya bahwa sebagian besar peminat Aktuaria memiliki nilai disekitar nilai rata-rata. Peminat UI memiliki keberagaman nilai yang tinggi, hal ini dapat diketahui dari grafik yang lebih lebar. Pada nilai Bahasa Indonesia, Bahasa Inggris, dan Matematika Dasar menunjukkan hal yang sama. Peminat ITS memiliki kemampuan dalam Bahasa Indonesia dan Matematika Dasar yang lebih tinggi dibandingkan IPB dan UI, sedangkan kemampuan Bahasa Inggris tertinggi berada pada peminat UI.



Gambar 4.15 Density Plot Nilai TKD Saintek Peminat Aktuaria

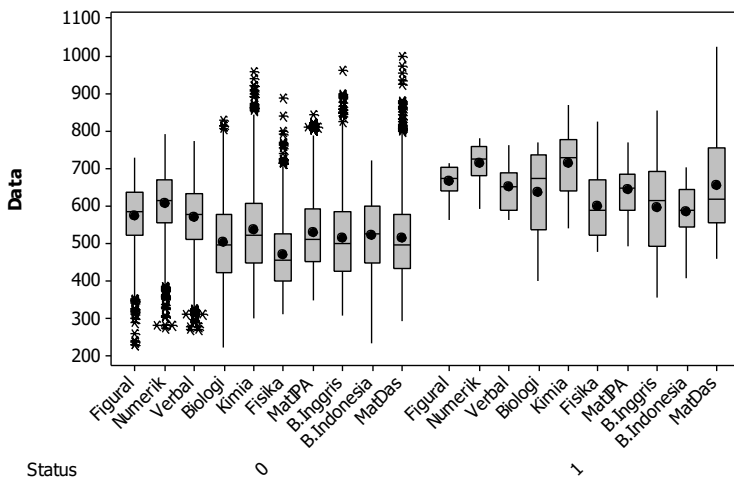
Gambar 4.15 menunjukkan kelompok nilai TKD Saintek peminat Aktuari ITS, IPB dan UI pada ujian tes tulis masuk PTN tahun 2018. Dapat dilihat bahwa pada nilai Biologi dan Kimia, peminat ITS memiliki jumlah peminat yang lebih tinggi dengan karakteristik nilai berada disekitar rata-rata dibandingkan dengan jumlah peminat IPB dan UI. Namun demikian, nilai rata-rata tertinggi kedua mata uji tersebut berada pada peminat UI. Grafik nilai Biologi dan Kimia pada peminat UI lebih lebar, artinya bahwa nilai Biologi dan Kimia peminat UI memiliki varian yang tinggi atau keberagaman nilai yang cukup besar. Mata uji Fisika dan Matematika IPA menunjukkan bahwa frekuensi peminat terbanyak dengan nilai berada disekitar rata-rata adalah peminat IPB. Karakteristik keempat mata uji pada peminat IPB dan UI menunjukkan kemiripan, hal ini terlihat dari grafik yang hampir berimpitan.

Karakteristik nilai peminat program studi Aktuaria ITS, IPB, dan UI berdasarkan status penerimaan, yakni tidak diterima (0) dan diterima (1), dapat dijelaskan pada uraian berikut.



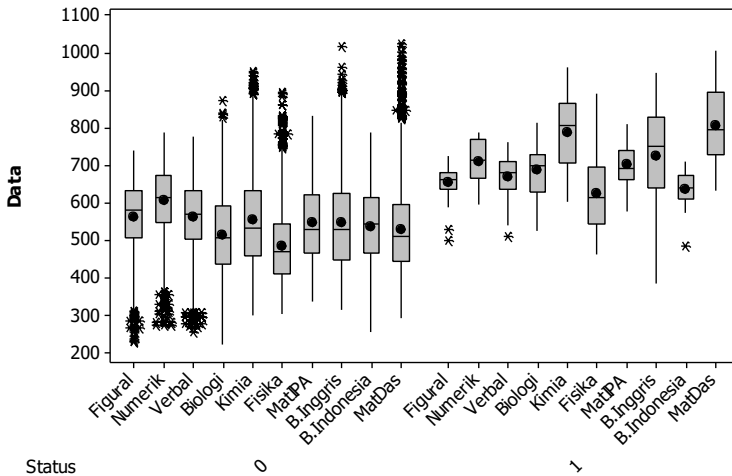
Gambar 4.16 Box Plot Nilai Mata Uji Tes Tulis Peminat Aktuarial ITS

Peminat Aktuarial ITS yang tidak diterima memiliki keberagaman nilai yang tinggi, hal ini ditunjukkan pada Gambar 4.16 dimana terdapat nilai *outlier* baik diatas maupun dibawah nilai rata-rata. Jumlah *outlier* yang cukup banyak pada semua mata uji, kecuali pada Biologi dan Bahasa Indonesia hanya terdapat satu data *outlier*. Nilai rata-rata Biologi, Kimia, Fisika, dan Matematika IPA memiliki nilai yang saling mendekati. Peminat yang diterima Aktuarial ITS cenderung memiliki nilai Kimia, Bahasa Inggris, dan Matematika dasar yang lebih besar dari nilai rata-ratanya. Peminat yang diterima juga memiliki karakteristik adanya nilai yang berada jauh di atas maupun dibawah dari nilai peminat lainnya. Nilai tertinggi peminat yang diterima berada pada nilai Matematika Dasar sebesar 1021,2 dan terendah berada pada Matematika IPA dengan nilai 395. Nilai rata-rata tertinggi berada pada nilai Matematika Dasar dengan nilai 698,1 dan terendah berada pada nilai Biologi sebesar 581,6.



Gambar 4.17 Box Plot Nilai Mata Uji Tes Tulis Peminat Aktuaria IPB

Gambar 4.17 menunjukkan perbedaan pada kedua status penerimaan. Peminat yang diterima memiliki nilai yang seragam sedangkan peminat yang tidak diterima memiliki nilai keberagaman yang tinggi pada semua mata uji, kecuali pada mata uji Bahasa Indonesia yang tidak menunjukkan adanya data *outlier*. Pada nilai TPA, peminat yang tidak diterima memiliki nilai outlier yang berada jauh dibawah nilai rata-rata. Karakteristik peminat Aktuaria IPB yang diterima memiliki kesamaan dengan peminat Aktuaria ITS, dimana sebagian besar peminat yang diterima memiliki karakteristik nilai lebih besar dibandingkan dengan nilai rata-rata pada mata uji Matematika Dasar dan Bahasa Inggris. Selain kedua mata uji tersebut, nilai Fisika juga menunjukkan bahwa peminat dengan nilai diatas rata-rata lebih banyak. Nilai rata-rata tertinggi peminat yang diterima adalah 714,5 pada Numerik dan terendah adalah Bahasa Inggris dengan nilai 584. Gambar 4.17 menunjukkan bahwa pada mata uji Verbal, Kimia, Fisika, Bahasa Inggris, dan Matematika Dasar sebagian besar peminat yang diterima memiliki karakteristik nilai diatas rata-rata.



Gambar 4.18 Box Plot Nilai Mata Uji Tes Tulis Peminat Aktuaria UI

Peminat Aktuaria UI yang diterima secara keseluruhan memiliki nilai rata-rata yang lebih tinggi dibandingkan dengan nilai rata-rata peminat Aktuaria ITS dan IPB. Nilai rata-rata tertinggi Matematika Dasar dengan nilai 806,9 dan terendah adalah Fisika dengan nilai 626,3. Walau memiliki nilai rata-rata tertinggi diantara Aktuaria ITS dan IPB, terdapat beberapa nilai yang berada jauh dari nilai lainnya atau adanya outlier pada nilai Figural, Verbal dan Bahasa Indonesia. *Outlier* yang ada pada ketiga mata uji tersebut berada jauh dibawah dari ketiga nilai rata-rata mata uji tersebut. Peminat yang tidak diterima memiliki nilai rata-rata dan median dibawah dari nilai peminat yang diterima. Rata-rata tertinggi peminat yang tidak diterima berada pada mata uji Numerik dengan nilai 605,74 dan terendah adalah Fisika dengan 486,5. Adanya nilai *outlier* pada setiap mata uji peminat yang tidak diterima, kecuali pada Matematika IPA dan Bahasa Indonesia, menyebabkan tingkat keragaman atau penyebaran data yang lebih besar dibandingkan peminat yang diterima. Hal tersebut diperkuat dengan visualisasi ukuran *box* pada status

peminat yang tidak diterima, secara keseluruhan lebih panjang dibandingkan dengan yang diterima.

4.2 Penentuan Mata Uji yang Signifikan dan Batas Nilai Minimal

Penentuan mata uji yang memberikan pengaruh signifikan terhadap status diterimanya peserta tes tulis dapat menggunakan analisis *Classification Tree*. Metode diterapkan pada data peserta peminat prodi Statistika dan Matematika ITS, IPB, dan UI berdasarkan status penerimaan peserta. Pengolahan data dengan analisis *Classification Tree* menggunakan pembagian data *training* dan *testing* sebesar $\frac{2}{3}$ dan $\frac{1}{3}$ dari data. Terdapat tiga tahapan analisis klasifikasi, yakni pembentukan pohon maksimal, pemangkasan pohon, dan penentuan pohon optimum.

4.2.1 Penentuan Mata Uji yang Signifikan dan Batas Nilai Minimal Program Studi Statistika

Analisis *Classification Tree* pada nilai 10 mata uji tes masuk PTN dengan prodi Statistika, selain menghasilkan klasifikasi dapat pula memberikan informasi mengenai mata uji yang memberikan pengaruh signifikan pada penentuan diterimanya peserta. Selain itu, dapat pula diperoleh informasi mengenai batasan nilai agar dapat masuk prodi Statistika, khususnya pada Statistika ITS. Batasan nilai didapatkan dari nilai *threshold* yang dijadikan sebagai pemilah pada analisis *Classification Tree*. Terdapat tiga tahapan dalam analisis dengan *Classification Tree*, yakni pembentukan pohon maksimal, Pemangkasan pohon, dan pemilihan pohon optimal.

a. Statistika Institut Teknologi Sepuluh Nopember

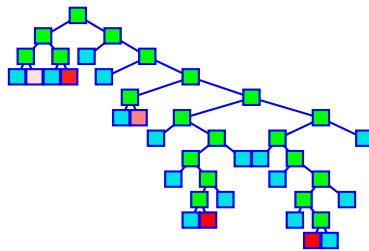
Pembentukan pohon maksimal adalah tahap pertama dalam analisis dengan *Classification Tree*. Pembentukan pohon maksimal diawali dengan pemilihan pemilah atau *classifier*, yakni dengan cara memilih variabel prediktor terpenting yang dijadikan sebagai pemilah awal. Proses pemilihan pemilah dilakukan berdasarkan kriteria *goodness of split*. Variabel dengan nilai

goodness of split tertinggi merupakan variabel terpenting dalam pemilahan, dimana pemilah tersebut mampu mereduksi heterogenitas yang tinggi. Tabel 4.2 menunjukkan variabel terpenting berdasarkan skor kontribusi terbesar yang dihasilkan pada setiap variabel prediktor.

Tabel 4.2 Skor Kontribusi Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Maksimal

Variabel	Skor	Turus
Numerik	100	
Kimia	56,85	
Biologi	42,11	
Figural	27,35	
Matdas	25,2	
Verbal	11,54	
Matipa	6,74	
B_Indonesia	3,07	
B_Inggris	1,05	
Fisika	1,04	

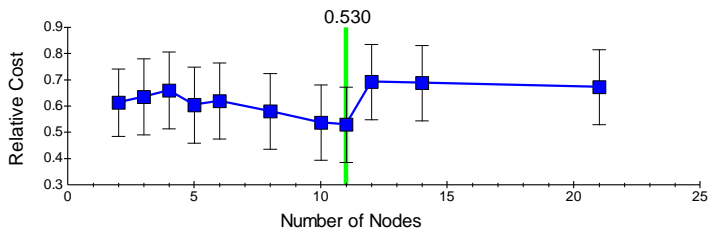
Skor pada Tabel 4.2 menunjukkan nilai kontribusi variabel dalam pembentukan pohon klasifikasi maksimal. Nilai kontribusi tertinggi berada pada variabel nilai Numerik. Artinya variabel nilai Numerik menjadi pemilah awal atau menjadi simpul induk (*parent nodes*). Nilai Numerik akan dilakukan proses pemilahan dengan menggunakan fungsi keheterogenan indeks gini. Pohon klasifikasi maksimal akan membentuk simpul terminal dengan terbanyak dengan tingkat kedalaman (*depth*) tertinggi. Hasil pembentukan pohon klasifikasi maksimal dapat dilihat pada Gambar 4.19.



Gambar 4. 19 Pohon Klasifikasi Maksimal Prodi Statistika ITS

Pohon klasifikasi maksimal pada Gambar 4.19 menunjukkan bahwa pohon membentuk simpul akhir atau *terminal node* yang berjumlah 21 nodes dengan tingkat kedalaman (*depth*) sebesar 12. Penamaan kelas klasifikasi dapat dilihat pada warna yang berbeda pada setiap *terminal nodes*. Warna merah menandakan bahwa klasifikasi termasuk dalam status peserta diterima (1), dan warna biru termasuk dalam status peserta tidak diterima (0). Semakin tajam warna yang dihasilkan, semakin besar perbedaan proporsi antara kategori status penerimaan.

Pohon klasifikasi maksimal seringkali terdapat kasus *overfitting* (akurasi yang dihasilkan melebihi kenyataan), namun jika diberikan batasan dalam proses pemilahan akan terjadi kasus *underfitting* (nilai akurasi yang dihasilkan kurang jika dibandingkan dengan nilai aktual). Sehingga diperlukan proses pemangkasan pohon untuk mendapatkan ukuran pohon klasifikasi yang optimal. Pemangkasan dilakukan berdasarkan pendekatan *cost complexity minimum*. Pohon klasifikasi yang optimal dapat diketahui dari hasil plot *relative cost*, dimana node dengan garis berwarna hijau merupakan ukuran pohon klasifikasi optimal



Gambar 4.20 Plot Relative cost

Gambar 4.20 menunjukkan bahwa nilai *relative cost* minimum berada pada nodes 11, dengan nilai sebesar 0,530. Hal tersebut menunjukkan bahwa pohon klasifikasi optimal memiliki jumlah *terminal nodes* sebanyak 11. Nilai *relative cost* yang kecil menunjukkan bahwa error atau kesalahan klasifikasi semakin kecil pula. Sehingga dalam penentuan pemangkasan pohon,

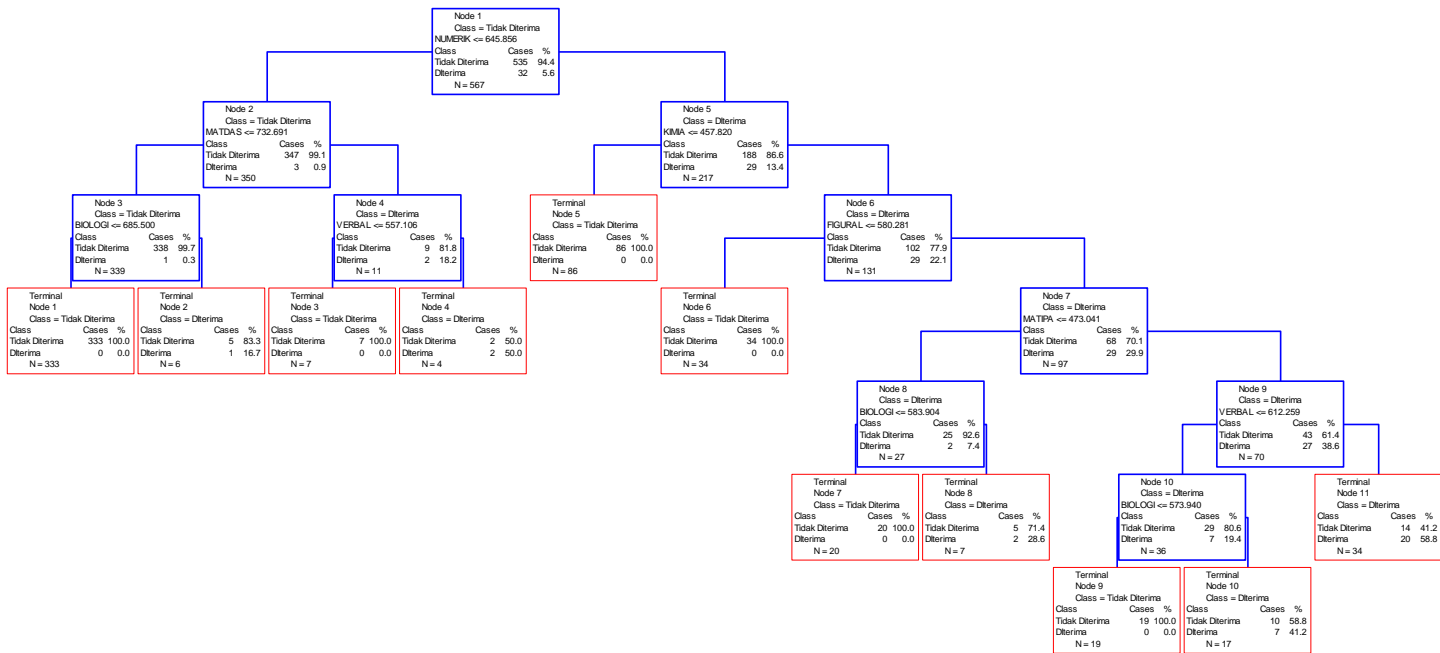
melihat nilai *relative cost* yang paling minimum. Secara lebih rinci nilai *relative cost* pada seluruh terminal node dapat ditunjukkan pada Lampiran 4. Berdasarkan nilai *relative cost* minimum, pohon klasifikasi optimal dari data peminat prodi Statistika ITS dapat dilihat pada Gambar 4.21. Sedangkan skor kontribusi variabel prediktor dalam pembentukan pohon klasifikasi optimal dapat dilihat pada Tabel 4.3

Tabel 4.3 Skor Kontribusi Variabel Prediktor dalam Pembentukan Pohon Klasifikasi Optimal

Variabel	Skor	Turus
Numerik	100,00	
Kimia	58,04	
Biologi	40,89	
Figural	27,93	
Matdas	23,58	
Verbal	8,83	
Matipa	6,88	

Tabel 4.3 menunjukkan bahwa variabel numerik merupakan variabel terpenting dalam pembentukan pohon klasifikasi optimal dengan skor kontribusi sebesar 100. Variabel nilai Bahasa Indonesia, Bahasa Inggris dan Figural tidak memberikan kontribusi dalam pembentukan pohon klasifikasi optimal. Proses pemilihan pemilah dalam membentuk pohon klasifikasi optimal menggunakan indeks gini memberikan hasil bahwa simpul terminal sebanyak 11 simpul dan tingkat kedalaman sebesar 7. Gambar 4.21 menampilkan topologi pohon klasifikasi optimal berdasarkan mata uji yang mempengaruhi status penerimaan peserta tes tulis pada prodi Statistika ITS.

Gambar 4.21 menunjukkan bahwa pohon klasifikasi optimal yang terbentuk menghasilkan 11 *terminal nodes*. Hal tersebut menunjukkan bahwa penentuan berdasarkan beberapa kriteria nilai, yang menyebabkan keketatan dalam penerimaan semakin ketat.



Gambar 4.21 Hasil Pohon Klasifikasi Optimal Peminat Prodi Statistika ITS

Berdasarkan hasil pohon klasifikasi optimal pada Gambar 4.21 diperoleh beberapa informasi karakteristik berikut.

1. *Terminal node* 1, 3, 5, 6, 7, dan 9 merupakan *terminal node* dengan klasifikasi peserta tidak diterima Statistika ITS. Aturan yang terbentuk dalam penentuan peserta yang tidak diterima Statistika ITS melalui jalur tes tulis adalah
 - a. Nilai Numerik $\leq 645,856$; Matematika Dasar $\leq 732, 691$; dan nilai Biologi $\leq 685,5$;
 - b. Nilai Numerik $\leq 645,856$; Matematika Dasar $> 732, 691$; dan nilai Verbal $\leq 557,106$;
 - c. Nilai Numerik $> 645,856$; Kimia $\leq 457,820$;
 - d. Nilai Numerik $> 645,856$; Kimia $> 457,820$; Figural $\leq 589,281$.
 - e. Nilai Numerik $> 645,856$; Kimia $> 457,820$; Figural $> 589,281$; Matematika IPA $\leq 473,041$; Biologi $\leq 573,904$
 - f. Nilai Numerik $> 645,856$; Kimia $> 457,820$; Figural $> 589,281$; Matematika IPA $> 473,041$; TPA Verbal $\leq 612,259$ dan Biologi $\leq 573,940$.
2. *Terminal node* 2, 4, 8, 10, 11, merupakan peserta yang diterima Statistika ITS melalui jalur tes tulis. Peserta memiliki nilai:
 - a. Numerik $\leq 645,856$; Matematika Dasar $\leq 732, 691$; dan nilai Biologi $> 685,5$;
 - b. Numerik $\leq 645,856$; Matematika Dasar $> 732, 691$; TPA Verbal $> 557,106$ dan nilai Biologi $> 685,5$;
 - c. Numerik $> 645,856$; Kimia $> 457,820$; Figural $> 589,281$; Matematika IPA $\leq 473,041$; dan Biologi $> 583,904$;
 - d. Numerik $> 645,856$; Kimia $> 457,820$; Figural $> 589,281$; Matematika IPA $> 473,041$; TPA Verbal $\leq 612,259$ dan Biologi $> 573,940$.
 - e. Numerik $> 645,856$; Kimia $> 457,820$; Figural; Matematika IPA $> 473,041$; dan TKP Verbal $> 612,259$.

Setelah mendapatkan pohon klasifikasi maksimal, akan dapat diketahui nilai ketepatan klasifikasi yang menandakan kebaikan model klasifikasi. Hasil klasifikasi status penerimaan peserta tes

tulis masuk PTN 2018, dengan peminat prodi Statistika ITS dapat diketahui pada Tabel 4.4.

Tabel 4.4 Evaluasi Ketepatan Klasifikasi Peminat Statistika ITS

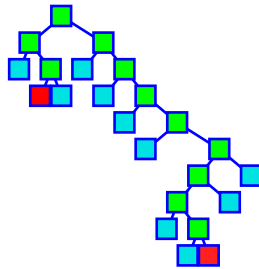
	Akurasi	Sensitivity	Specificity	AUC
<i>Training</i>	0,937	0,933	1,000	0,967
<i>Testing</i>	0,873	0,886	0,583	0,735

Berdasarkan Tabel 4.4 diketahui bahwa klasifikasi yang diberikan pada pohon klasifikasi optimal memiliki persentase ketepatan atau akurasi yang cukup tinggi, yakni 93,7% dengan kesalahan klasifikasi sebesar 6,3%. Kesalahan klasifikasi data dengan status tidak diterima namun diklasifikasikan diterima dapat disebabkan karena salah satu penentu nilai untuk tidak diterima adalah nilai Biologi $\leq 685,5$ sedangkan pada Gambar 4.6 menunjukkan bahwa nilai peminat yang tidak diterima terdapat beberapa peminat dengan nilai diatas 700. Sedangkan pada peminat yang diterima namun diklasifikasikan tidak diterima dapat disebabkan karena nilai terendah dari Figural adalah 533,75. Sedangkan pada aturan penerimaan menyebutkan bahwa nilai Figural $> 589,281$.

Hasil dari pohon klasifikasi yang optimal kemudian dilakukan evaluasi dengan menggunakan data *testing*. Hasil evaluasi menunjukkan bahwa tingkat ketepatan klasifikasi sebesar 87,3%, nilai tersebut lebih rendah dibandingkan dengan nilai akurasi pada data *training*. Artinya bahwa pada data testing terdapat kesalahan klasifikasi yang cukup banyak. Adanya kesalahan tersebut menyebabkan nilai AUC pada data *testing* bernilai 73,5%, nilai tersebut menunjukkan kekuatan yang sedang dalam pengklasifikasian.

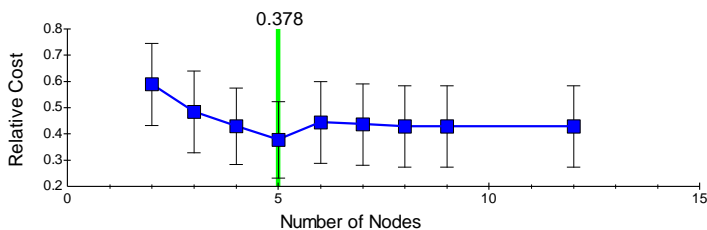
b. Statistika Institut Pertanian Bogor

Data *training* yang digunakan dalam analisis *classification tree* data peminat Statistika IPB sebesar 598 dan 278 data *testing*. Langkah dalam analisis yang digunakan sama halnya dengan langkah yang telah diuraikan pada Statistika ITS. Pohon klasifikasi maksimal yang terbentuk dapat dilihat pada Gambar 4.22.



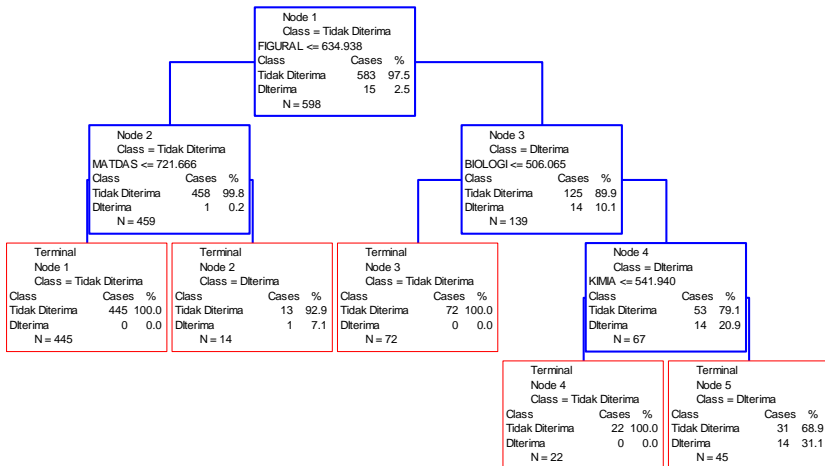
Gambar 4.22 Pohon Klasifikasi Maksimal Prodi Statistika IPB

Pohon klasifikasi maksimal yang ditunjukkan pada Gambar 4.22 memiliki jumlah *terminal nodes* sebanyak 12 dengan tingkat kedalaman pohon sebesar 10. Pada Lampiran 5, menunjukkan bahwa dalam pembentukan pohon klasifikasi, variabel nilai Figural memiliki skor kontribusi tertinggi. Hal ini dapat diartikan bahwa nilai Figural menjadi pemilah utama dalam pembentukan pohon klasifikasi. Selanjutnya akan dilakukan pemangkasan pohon dengan melihat nilai *relative cost* yang paling rendah pada setiap nodes yang ada.



Gambar 4.23 *Relative cost Plot* Data Peminat Statistika IPB

Gambar 4.23 menunjukkan bahwa nilai *cost complexity* terendah berada pada node 5 dengan nilai 0,378. Hal ini menandakan bahwa pohon klasifikasi yang optimal akan membentuk 5 *terminal node*, dengan kata lain perlu dilakukan pemangkasan pada pohon klasifikasi maksimal. Secara keseluruhan nilai relative cost pada setiap terminal nodes juga dapat dilihat pada Lampiran 5. Hasil pemangkasan pohon optimal dapat dilihat pada Gambar 4.24.



Gambar 4.24 Pohon Klasifikasi Optimal Data Peminat Statistika IPB

Hasil pohon klasifikasi optimal pada data peminat Statistika IPB pada tes tulis masuk PTN 2018 dapat diuraikan sebagai berikut.

1. *Terminal node 2* dan *5* menunjukkan bahwa klasifikasi termasuk dalam kelas diterima, dimana pada *terminal node 2* memiliki karakteristik nilai *Figural* $\leq 634,938$, *Matematika dasar* $> 721,666$. Sedangkan pada *terminal node 5*, menunjukkan bahwa jika peserta diterima prodi Statistika IPB jika nilai *figural* $> 634,938$, nilai *Biologi* $> 506,065$, dan nilai *Kimia* lebih dari $541,940$.
2. Peserta yang tidak diterima masuk dalam *terminal node 1, 3*, dan *4*. Aturan yang dihasilkan adalah
 - a. *Figural* $\leq 634,938$, *Matematika dasar* $\leq 721,666$;
 - b. *Figural* $> 634,938$, nilai *Biologi* $\leq 506,065$,
 - c. *Figural* $> 634,938$, nilai *Biologi* $> 506,065$, dan *Kimia* $> 541,940$.

Selanjutnya dilakukan pengukuran ketepatan klasifikasi dari pohon klasifikasi optimal yang telah terbentuk. Tabel 4.5 menunjukkan hasil klasifikasi pada data *training* dan data *testing*.

Tabel 4.5 Evaluasi Ketepatan Klasifikasi Peminat Statistika IPB

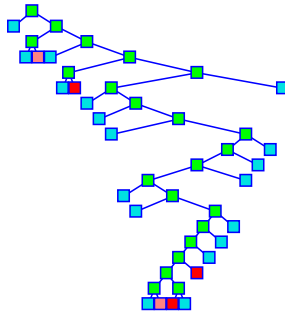
	Akurasi	Sensitivity	Specificity	AUC
<i>Training</i>	0,926	0,925	1,000	0,963
<i>Testing</i>	0,914	0,922	0,700	0,811

Tabel 4.5 menunjukkan bahwa nilai akurasi antara data *training* dan data *testing* memiliki nilai yang hampir sama. Data *training* memiliki nilai akurasi sebesar 92,6% dengan kesalahan klasifikasi pada kelas tidak diterima sebesar 44 peserta. Kesalahan klasifikasi pada peminat yang tidak diterima namun diklasifikasikan diterima dapat disebabkan karena salah satu ketentuan peserta tidak diterima adalah nilai Matematika dasar $\leq 721,666$. Dapat dilihat pada Gambar 4.7 yang menunjukkan bahwa nilai Matematika dasar peminat yang tidak diterima terdapat beberapa peserta dengan nilai Matematika dasar $> 721,666$ dan bahkan mencapai 1178.

Sedangkan hasil akurasi dari data *testing*, yang merupakan hasil validasi klasifikasi dari pohon optimal yang terbentuk memiliki selisih 1,2% lebih rendah dibandingkan dengan data *training*, yakni 91,4%. Pada data *testing* terdapat kesalahan klasifikasi peminat yang diterima menjadi tidak diterima dapat disebabkan karena nilai minimum pada mata uji matematika dasar peminat yang diterima adalah 449,3, sedangkan batas nilai Matematika dasar $> 721,666$. Nilai AUC pada data *testing* menunjukkan bahwa klasifikasi dengan menggunakan *Classification tree* memiliki kekuatan yang baik dalam mengklasifikasikan, hal ini terlihat bahwa nilai AUC data *testing* bernilai 81,1%.

c. Statistika Universitas Indonesia

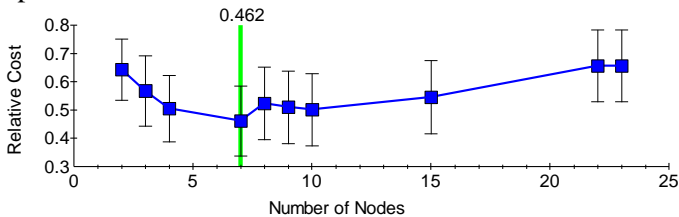
Pembagian data *training* dan *testing* pada data peminat Statistika UI adalah 511 untuk data *training* dan 240 untuk data *testing*. Hasil pembentukan pohon klasifikasi maksimal dengan menggunakan data *training* dapat dilihat pada Gambar 4.25.



Gambar 4.25 Pohon Klasifikasi Maksimal Data Peminat Statistika UI

Pohon klasifikasi maksimal pada data peminat Statistika UI menunjukkan bahwa *terminal nodes* yang terbentuk adalah 23 nodes dengan kedalaman sebesar 20. Variabel nilai Numerik menjadi pemilah utama dalam pembentukan pohon klasifikasi, hal ini dikarenakan selain memiliki tingkat keheterogenan yang tinggi, variabel nilai Numerik memiliki nilai kontribusi yang tinggi dalam pembentukan pohon. Hal ini dapat diketahui pada skor masing-masing variabel yang menunjukkan skor kontribusi dalam pembentukan pohon pada Lampiran 6.

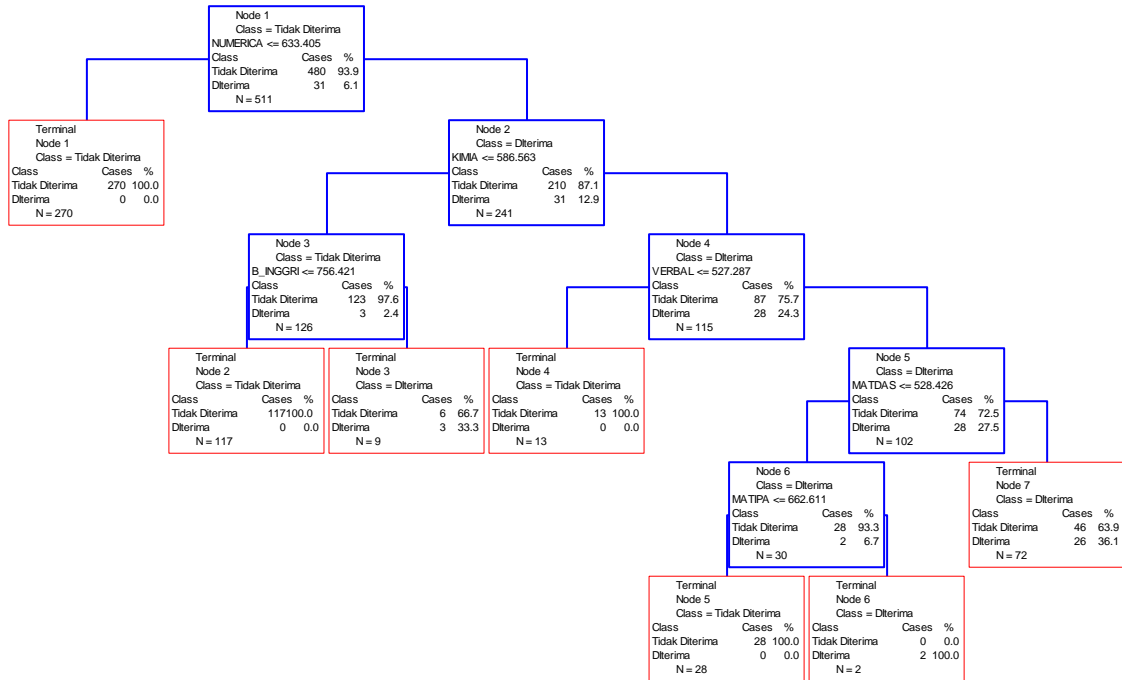
Pohon klasifikasi maksimal memiliki jumlah *terminal nodes* yang terlalu banyak, dan dimungkinkan akan terjadi kasus *overfitting*. Sehingga perlu dilakukan pemangkasan pohon. Pemangkasan pohon dilakukan dengan melihat nilai *relative cost* yang paling minimum pada setiap *terminal nodes*. Gambar 4.26 menunjukkan plot nilai *relative cost* pada masing-masing terminal node. Secara rinci nilai *relative cost* dapat dilihat pula pada Lampiran 6.



Gambar 4.26 *Relative cost Plot* Data Peminat Statistika UI

Gambar 4.26 menunjukkan bahwa pada *terminal node 7* memiliki nilai *relative cost* sebesar 0,462. Nilai tersebut merupakan nilai paling minimum diantara *terminal node* lainnya. Secara lebih rinci, nilai *relative cost* dapat dilihat pada Lampiran 6. Pemangkasan pohon dilakukan pada *terminal nodes 7*, dimana merupakan terminal node dengan kesalahan klasifikasi atau error yang lebih kecil dibandingkan pada *terminal node* lainnya.. Hasil pemangkasan pohon dapat dilihat pada Gambar 4.27. Berdasarkan hasil pohon klasifikasi optimal terdapat beberapa aturan dalam penerimaan peserta tes tulis masuk PTN tahun 2018 pada peminat Statistika UI. Aturan yang terbentuk adalah sebagai berikut.

1. *Terminal node 3, 6, dan 7* merupakan *terminal nodes* dengan klasifikasi berupa peserta tes tulis diterima di Statistika UI. Aturan penerimaan yang dihasilkan dari pohon klasifikasi optimal dapat diuraikan sebagai berikut.
 - a. Numerik $> 633,405$; Kimia $\leq 586,563$; Bahasa Inggris $> 756,421$
 - b. Numerik $> 633,405$; Kimia $> 586,563$; TPA Verbal $> 527,287$; Matematika Dasar $\leq 528,426$, dan Matematika IPA $> 662,611$
 - c. Numerik $> 633,405$; Kimia $> 586,563$; TPA Verbal $> 527,287$; dan Matematika Dasar $> 528,426$
2. *Terminal node 1, 2, 4, dan 5* merupakan *terminal node* dengan klasifikasi peserta tidak diterima di Statistika UI. Nilai peserta yang tidak diterima dapat digambarkan sebagai berikut:
 - a. Numerik $\leq 633,405$
 - b. Numerik $> 633,405$; Kimia $\leq 586,563$; Bahasa Inggris $\leq 756,421$
 - c. Numerik $> 633,405$; Kimia $> 586,563$; TPA Verbal $\leq 527,287$
 - d. Numerik $> 633,405$; Kimia $> 586,563$; TPA Verbal $> 527,287$; Matematika Dasar $\leq 528,426$; dan Matematika IPA $\leq 662,611$.



Gambar 4.27 Pohon Klasifikasi Optimal Data Peminat Statistika UI

Hasil klasifikasi pohon optimal kemudian dilakukan validasi apakah pohon yang dihasilkan dapat digunakan dalam penentuan status penerimaan pada data baru peminat statistika UI.

Tabel 4.6 Evaluasi Ketepatan Klasifikasi Peminat Statistika UI

	Akurasi	Sensitivity	Specificity	AUC
<i>Training</i>	0,898	0,892	1,000	0,946
<i>Testing</i>	0,858	0,871	0,667	0,769

Tabel 4.6 menunjukkan bahwa data *training* memiliki nilai akurasi sebesar 89,8%. Semua data peminat yang diterima Statistika UI dengan menggunakan aturan yang dihasilkan pohon klasifikasi optimal menunjukkan bahwa semua data tepat klasifikasi, sedangkan pada data yang tidak diterima terdapat kesalahan klasifikasi sebesar 52 data. Hasil akurasi dari data *testing*, menunjukkan nilai yang mendekati dengan nilai akurasi data *training*, yakni sebesar 85,8%. Kekuatan hasil klasifikasi pada data *testing* masuk dalam kategori sedang dengan nilai AUC sebesar 76,9%.

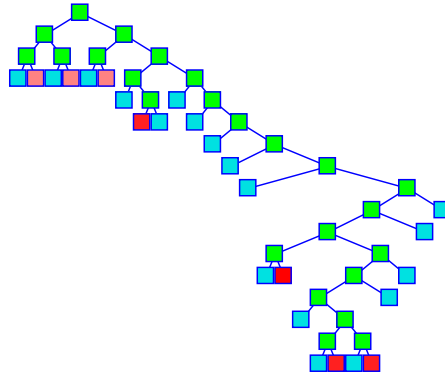
4.2.2 Penentuan Mata Uji yang Signifikan dan Batas Nilai Minimal Program Studi Matematika

a. Matematika Institut Teknologi Sepuluh Nopember

Program studi Matematika ITS merupakan prodi dengan predikat akreditasi A. Adanya predikat ini tentu saja dapat meningkatkan peminat peserta tes tulis masuk PTN 2018 untuk dapat masuk dalam prodi Matematika ITS. Tahun 2018 peminat mengalami sedikit penurunan dibandingkan tahun lalu, yakni dari 612 tahun 2018, menjadi 603. Hal ini dapat dikarenakan adanya program studi Aktuaria yang dibuka dan masuk dalam fakultas yang sama dengan Matematika dan Statistika. Hasil analisis klasifikasi penerimaan peserta tes tulis masuk PTN 2018 prodi Matematika ITS dapat dijelaskan sebagai berikut.

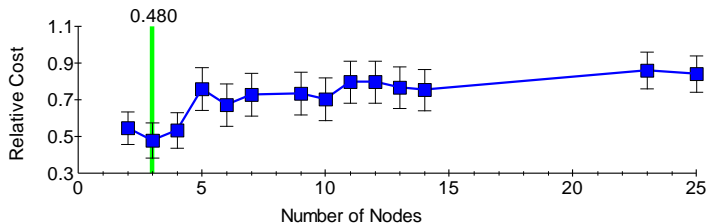
Pembentukan pohon maksimal dilakukan pada data *training* dengan jumlah $2/3$ tiap kategori dari data keseluruhan, yakni sebesar 32 data untuk peminat diterima dan 390 peminat yang tidak diterima. Total data *training* adalah 422 data. Gambar 4.28

merupakan hasil pohon klasifikasi maksimal dari data *training* peminat prodi Matematika ITS pada tes tulis masuk PTN tahun 2018.



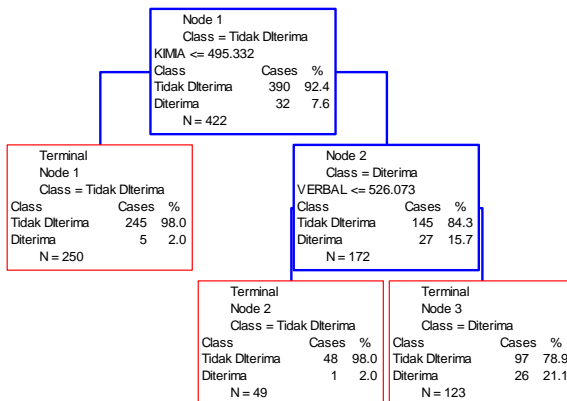
Gambar 4.28 Pohon Klasifikasi Maksimal Data Peminat Prodi Matematika ITS

Pohon klasifikasi maksimal yang terbentuk memiliki 25 *terminal nodes* dengan tingkat kedalaman sebesar 17. Pemilah utama dalam pohon klasifikasi yang terbentuk adalah variabel nilai Kimia. Variabel tersebut memiliki skor kontribusi tertinggi dalam pembentukan pohon (Lampiran 7). Pohon yang terbentuk memiliki ukuran yang terlalu besar dan dapat menyulitkan dalam mengambil keputusan. Selain itu, dikhawatirkan akan terjadi *overfitting* dalam penentuan kelas. Sehingga diperlukan pemangkasan pohon untuk mendapatkan ukuran pohon klasifikasi yang optimal.



Gambar 4.29 *Relative cost Plot* Data Peminat Matematika ITS

Gambar 4.29 menunjukkan bahwa nilai *relative cost* terendah berada pada *terminal nodes* 0,480. Dengan demikian pemangkasan pohon dilakukan pada *terminal nodes* ke 3. Hasil pemangkasan pohon atau pohon klasifikasi optimal dapat dilihat pada Gambar 4.30.



Gambar 4.30 Pohon Klasifikasi Optimal Data Peminat Matematika ITS

Pohon klasifikasi optimal memiliki tiga *terminal nodes* dengan parent node berupa variabel nilai kimia. Aturan yang terbentuk berdasarkan hasil klasifikasi optimal dapat dijelaskan sebagai berikut.

1. Peminat Matematika ITS yang tidak diterima dalam tes tulis masuk PTN tahun 2018, memiliki 2 aturan, yakni (1) Apabila peserta tes tulis memiliki nilai kimia $\leq 495,332$ dan (2) Apabila peserta memiliki nilai kimia $> 495,332$ dan nilai TPA verbal $\leq 526,073$.
2. Peminat Matematika ITS dapat diterima apabila memiliki nilai Kimia $> 495,332$ dan nilai TPA Verbal $> 526,073$.

Dari hasil aturan yang terbentuk, dapat dilakukan klasifikasi pada data *training* dan *testing*. Tabel 4.7 menyajikan hasil klasifikasi yang berdasarkan aturan yang diberikan pada pohon klasifikasi optimal.

Tabel 4.7 Evaluasi Ketepatan Evaluasi Peminat Matematika ITS

	Akurasi	Sensitivity	Specificity	AUC
<i>Training</i>	0,756	0,751	0,813	0,782
<i>Testing</i>	0,729	0,720	0,800	0,760

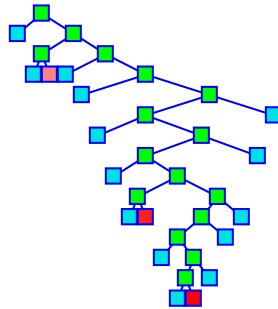
Pohon klasifikasi optimal terbentuk menghasilkan nilai akurasi pada data *training* sebesar 75,6% dengan kesalahan klasifikasi sebesar 97 peserta yang tidak diterima namun diprediksi diterima, dan 6 peserta yang diterima namun diprediksi tidak diterima. Adanya kesalahan klasifikasi ini disebabkan karena pada data peminat yang tidak diterima terdapat beberapa peminat dengan nilai TPA Verbal > 526,073, selain itu pada hasil *Boxplot* pada Gambar 4.11 menunjukkan sebaran data berada di rentang nilai 500 hingga 600. Sedangkan pada klasifikasi peminat yang diterima terjadi kesalahan klasifikasi karena nilai terendah untuk Nilai Kimia dengan status diterima adalah 402,9 dan nilai Verbal sebesar 487,93. Nilai tersebut tidak memenuhi aturan penerimaan peminat Matematika ITS berdasarkan hasil pohon klasifikasi optimal, sehingga diklasifikasikan kedalam status tidak diterima.

Akurasi yang dihasilkan pada data *testing* menunjukkan persentase yang lebih rendah dibandingkan dengan data *training*, yakni sebesar 72,9%. Penurunan akurasi disebabkan adanya data misklasifikasi pada kedua kategori. Nilai AUC pada data *testing* menunjukkan nilai sebesar 76% yang masuk dalam kategori kekuatan klasifikasi sedang.

b. Matematika Institut Pertanian Bogor

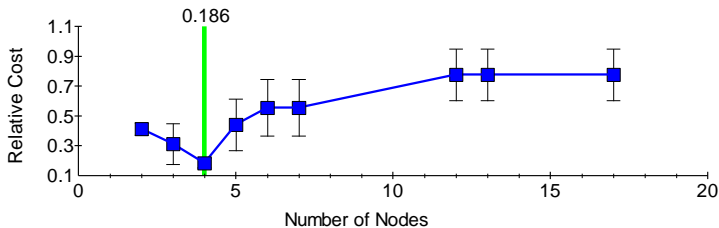
Peminat Matematika IPB lebih sedikit dibandingkan dengan peminat Matematika ITS dan UI. Hal ini disebabkan karena kedua PTN, yakni ITS dan UI memiliki prodi Matematika sejak tahun 1960-an, sedangkan IPB baru memiliki prodi ini pada tahun 1980-an. Umur prodi terkadang memberikan pengaruh pada peminat, seperti semakin tua umur prodi maka semakin banyak pula lulusan yang dihasilkan. Berbagai lulusan mahasiswa prodi tersebut dapat memberikan pengaruh kepada adik kelas sewaktu SMA agar mendaftarkan diri.

Klasifikasi dengan menggunakan analisis *classification tree*, dilakukan dengan 3 tahapan. Pertama, pembentukan pohon klasifikasi maksimal. Pembentukan dilakukan dengan menggunakan nilai indeks gini dalam penentuan pemilah utama. Hasil pembentukan pohon maksimal dapat diketahui pada Gambar 4.31.



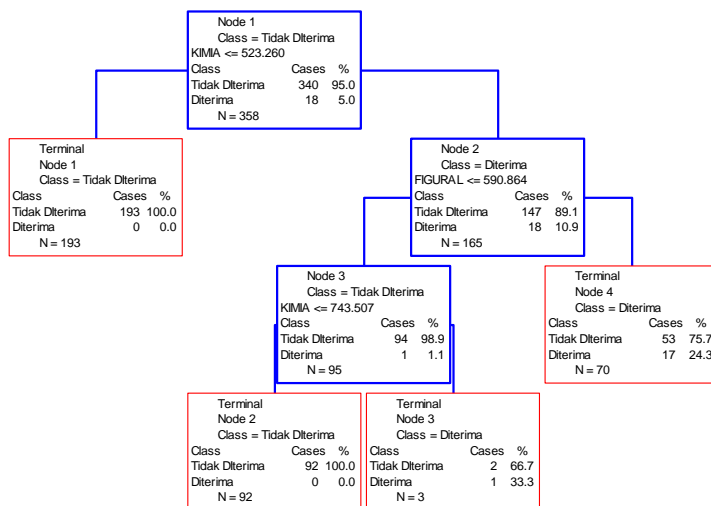
Gambar 4.31 Pohon Klasifikasi Maksimal Data Peminat Matematika IPB

Sama halnya dengan pohon klasifikasi maksimal ITS, peminat Statistika IPB melalui jalur tes tulis menunjukkan bahwa pohon yang terbentuk berukuran besar. *Terminal nodes* yang dihasilkan berjumlah 17 node dengan tingkat kedalaman pohon (*depth*) sebesar 15. Variabel nilai Kimia menjadi pemilah utama dalam pembentukan pohon. Ukuran pohon yang besar dapat memungkinkan terjadi *overfitting*. Untuk mencegah adanya kasus *overfitting* dapat dilakukan dengan pemangkasan pohon.



Gambar 4.32 *Relative cost Plot* Data Peminat Matematika IPB

Gambar 4.32 menunjukkan bahwa nilai minimum *relative cost* berada pada *terminal node* ke 4, yakni dengan nilai 0,186. Sedangkan pada pohon klasifikasi maksimal nilai *relative cost* berada diatas nilai *terminal node* ke-4. Pemangkasan dilakukan pada nilai *relative cost* yang paling minimum diantara *terminal node*. Dengan demikian pemangkasan dilakukan hingga membentuk pohon dengan jumlah *terminal node* 4. Secara lebih rinci, nilai *relative cost* setiap terminal node dapat dilihat pada Lampiran 8.



Gambar 4.33 Pohon Klasifikasi Optimal Data Peminat Matematika IPB

Pohon klasifikasi optimal pada Gambar 4.33 diketahui memiliki simpul akhir atau *terminal nodes* sebesar 4 dengan 2 simpul node diklasifikasikan sebagai peminat yang tidak diterima dan 2 node peminat yang diterima. Selain hasil klasifikasi, dapat diketahui pula karakteristik nilai dari setiap klasifikasi yang dihasilkan. Adapun karakteristik nilai peminat Matematika IPB dalam tes tulis masuk PTN 2018 dapat dijelaskan sebagai berikut.

1. Peminat yang tidak diterima memiliki nilai Kimia $\leq 523,260$. Jika terdapat peminat dengan nilai Kimia $> 523,260$ dan tidak lebih dari 743,507, serta nilai TPA figural $\leq 590,864$ maka peserta tersebut masuk dalam klasifikasi tidak diterima Matematika IPB.
2. Peminat yang diterima Matematika IPB memiliki nilai Kimia $> 523,260$ dan nilai Figural $> 590,864$. Selain itu apabila peserta memiliki nilai figural $\leq 590,864$, namun nilai Kimia dalam rentang 523,260 hingga 743,507 maka dapat diklasifikasikan sebagai peminat yang diterima.

Pohon klasifikasi optimal yang terbentuk selanjutnya dilakukan pengukuran ketepatan klasifikasi dengan menggunakan nilai akurasi. Tabel 4.8 menyajikan hasil akurasi dari data *training* dan data *testing*.

Tabel 4.8 Evaluasi Ketepatan Klasifikasi Peminat Matematika IPB

	Akurasi	Sensitivity	Specificity	AUC
<i>Training</i>	0,846	0,838	1,000	0,919
<i>Testing</i>	0,823	0,814	1,000	0,907

Hasil klasifikasi dengan menggunakan aturan yang ada pada hasil pohon klasifikasi optimal menunjukkan bahwa ketepatan klasifikasi pada data *training* adalah sebesar 84,6% dimana terdapat 55 peminat yang salah klasifikasi, yakni diklasifikasikan pada kelas diterima. Selanjutnya dilakukan validasi aturan yang telah terbentuk dengan menggunakan data *testing*. Hasil menunjukkan bahwa masih terdapat kesalahan klasifikasi, yakni sebesar 26 peminat diklasifikasikan sebagai peserta yang diterima.

Adanya kesalahan klasifikasi tersebut, dikarenakan terdapat beberapa peminat yang tidak diterima memiliki nilai Kimia lebih tinggi dari 743,507, sehingga dapat diklasifikasikan diterima berdasarkan aturan pohon klasifikasi optimal yang telah terbentuk. Sedangkan pada kelas yang diterima, terlihat bahwa semua peminat yang diterima telah tepat klasifikasi. Hal ini dikarenakan dapat diketahui pada Gambar 4.12, nilai Kimia pada peserta yang diterima paling rendah adalah 523,5. Nilai kimia

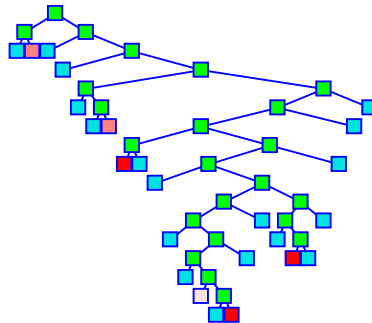
merupakan syarat diterimanya peminat Matematika IPB dengan nilai dalam rentang 523,260 hingga 743,507.

Nilai akurasi *testing* yang dihasilkan tidak jauh berbeda dari data *training*, yakni 82,3%. Persentase ketepatan klasifikasi yang dihasilkan masuk dalam kategori baik, hal ini ditunjukkan pula pada nilai AUC data *testing* yang bernilai 90,7%.

c. Matematika Universitas Indonesia

Prodi Matematika UI merupakan prodi yang lebih dahulu ada daripada prodi Matematika ITS dan IPB. Usia yang lebih tua 2 tahun dari Matematika ITS menyebabkan prodi ini tidak sepi dari peminat. Tercatat pula jumlah peminat Matematika UI pada tes tulis masuk PTN tahun 2018 lebih banyak dibandingkan kedua PTN. Hasil analisis klasifikasi dengan menggunakan *classification tree* dapat dilihat pada penjelasan dibawah.

Seperti pada analisis sebelumnya, pembentukan pohon klasifikasi maksimal merupakan tahap pertama. Tahap ini membentuk pohon dari data *training*, yakni 66,7% dari data keseluruhan, dimana data diambil berdasarkan kategori. Hasil pembentukan pohon klasifikasi maksimal dapat dilihat pada Gambar 4.34.

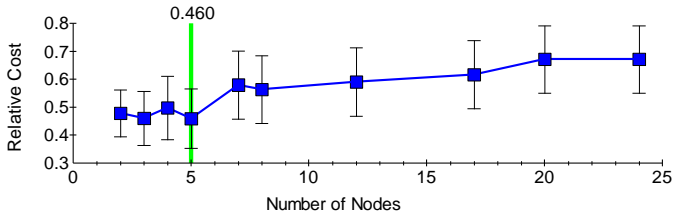


Gambar 4.34 Pohon Klasifikasi Maksimal Data Peminat Matematika UI

Pohon klasifikasi maksimal yang terbentuk berukuran cukup besar dengan tingkat kedalaman yang cukup dalam pula. Gambar 4.34 menunjukkan bahwa ukuran pohon klasifikasi

maksimal memiliki jumlah *terminal node* sebanyak 24 nodes dengan kedalaman (*depth*) pohon sebesar 17. Pada pohon klasifikasi maksimal variabel Numerik menjadi pemilah utama pohon dimana variabel tersebut memiliki skor tertinggi dalam variabel terpenting yang dapat dilihat pada Lampiran 9.

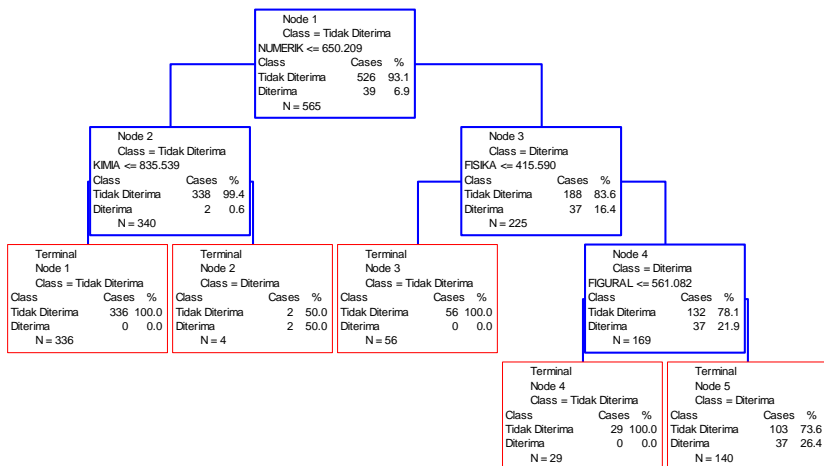
Hasil pohon klasifikasi memiliki ukuran yang besar. Diperlukan proses pemangkasan untuk mendapatkan pohon klasifikasi optimal. Pemangkasan dilakukan dengan melihat nilai minimum dari *relative cost* setiap *terminal nodes*. Gambar 4.35 menunjukkan nilai *relative cost* yang terbentuk pada setiap *terminal nodes*.



Gambar 4.35 *Relative cost Plot Data Peminat Matematika UI*

Gambar 4.35 menunjukkan bahwa nilai *relative cost* terendah pada *terminal node* 5 dengan nilai 0,460. Hal ini diartikan bahwa kesalahan klasifikasi terkecil berada pada pohon klasifikasi dengan jumlah *terminal node* sebanyak 5. Semakin kecil nilai *relative cost*, maka semakin kecil pula kesalahan klasifikasi yang dihasilkannya. Secara lebih rinci, nilai *relative cost* pada masing-masing *terminal node* dapat dilihat pada Lampiran 9. Dengan demikian, perlu dilakukan pemangkasan pohon hingga pohon klasifikasi menjadi optimal. Berdasarkan nilai *relative cost*, maka pohon klasifikasi optimal terbentuk pada pohon dengan jumlah *terminal nodes* sebanyak 5.

Pohon klasifikasi optimal pada Peminat Matematika UI pada tes tulis masuk perguruan tinggi negeri tahun 2018 dapat dilihat pada Gambar 4.36.



Gambar 4.36 Pohon Klasifikasi Optimal Peminat Matematika UI

Pohon klasifikasi optimal pada Gambar 4.36 memiliki jumlah *terminal nodes* sebanyak 5, dengan 3 nodes klasifikasi peminat yang tidak diterima dan 2 nodes peminat yang diterima. Aturan pengklasifikasian yang terbentuk dari pohon klasifikasi optimal dapat dijelaskan sebagai berikut.

- Peminat yang diklasifikasikan tidak diterima Matematika UI memiliki beberapa karakteristik nilai sebagai berikut.
 - Nilai Numerik $\leq 650,209$ dan nilai Kimia $\leq 835,539$
 - Nilai Numerik $> 650,209$ dan nilai Fisika $\leq 415,590$
 - Nilai Numerik $> 650,209$; nilai Fisika $> 415,590$; dan Nilai Figural $\leq 561,082$
- Peminat yang masuk dalam klasifikasi diterima Matematika UI melalui jalur tes tulis memiliki karakteristik nilai Numerik $> 650,209$ dan nilai Fisika $> 415,590$. Serta nilai Numerik $> 650,209$; nilai Fisika $> 415,590$; dan Nilai Figural $> 561,082$.

Hasil dari klasifikasi yang ada kemudian dilakukan pengukuran ketepatan klasifikasi. Hasil klasifikasi data peminat Matematika UI dapat dijelaskan pada Tabel 4.9.

Tabel 4.9 Evaluasi Ketepatan Klasifikasi Peminat Matematika UI

	Akurasi	Sensitivity	Specificity	AUC
<i>Training</i>	0,814	0,800	1,000	0,900
<i>Testing</i>	0,774	0,775	0,765	0,770

Hasil klasifikasi menunjukkan bahwa ketepatan akurasi pada data *training* adalah sebesar 81,4%. Terlihat bahwa pada data *training* terjadi kesalahan klasifikasi sebanyak 105 data yang diklasifikasikan ke dalam kelas diterima. Hasil akurasi data *testing* memiliki akurasi lebih rendah dibandingkan dengan nilai akurasi dari data *training*. Hal ini disebabkan karena adanya kesalahan klasifikasi baik pada kelas tidak diterima maupun diterima. Salah satu penyebab kesalahan klasifikasi pada kelas diterima adalah terdapat nilai Figural yang kurang dari batas minimal diterima nya (Gambar 4.13), yakni dengan nilai 477,26 sedangkan batas minimal adalah lebih dari 561,082. Sehingga peminat tersebut diklasifikasikan pada data yang tidak diterima.

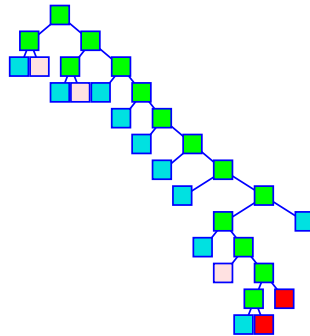
Kesalahan klasifikasi pada status peserta tidak diterima menjadi diterima dikarenakan pada hasil *Boxplot* nilai Kimia dan Fisika terdapat data *outlier* dengan jumlah yang banyak. Pada aturan hasil *classification tree* disebutkan bahwa nilai Kimia kurang dari 835,539, namun pada kenyataannya terdapat peserta dengan nilai melebihi batasan yang ada, dengan demikian peserta yang memiliki nilai $> 835,539$ diklasifikasikan pada kelas diterima. Selain itu batasan nilai Figural adalah kurang dari sama dengan 561,082, pada data peminat Matematika UI memiliki sebaran nilai berada pada rentang nilai 500 hingga 600 dan nilai maksimum berada pada nilai 739,26. Dengan demikian, peminat yang tidak diterima namun memiliki karakteristik nilai $> 561,082$ diklasifikasikan kedalam peminat yang diterima.

Nilai AUC menunjukkan kekuatan klasifikasi yang dihasilkan. AUC data *testing* menunjukkan nilai 77%. Artinya pohon klasifikasi yang terbentuk memiliki tingkat kekuatan sedang dalam mengklasifikasikan data.

4.2.3 Penentuan Mata Uji yang Signifikan dan Batas Nilai Minimal Program Studi Aktuarial

a. Aktuarial Institut Teknologi Sepuluh Nopember

Program studi Aktuarial merupakan prodi baru di ITS pada tahun 2018. Meskipun baru dibuka, prodi ini memiliki peminat yang tinggi dibandingkan prodi yang ada di FMKSD. Data yang digunakan penentuan mata uji yang signifikan dengan menggunakan *Classification Tree* dibagi menjadi dua yakni data *training* yang berjumlah 725 dan data *testing* sebesar 341. Hasil pembentukan pohon maksimal dari data *training* dapat dilihat pada Gambar 4.37.

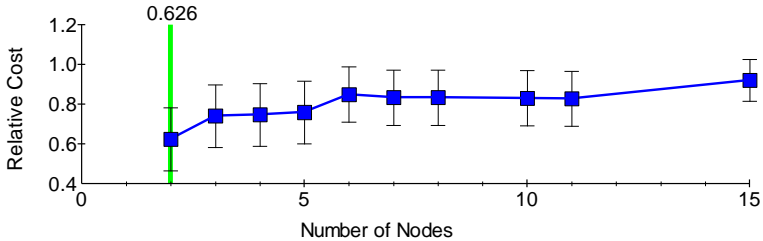


Gambar 4.37 Pohon Klasifikasi Maksimal Data Peminat Aktuarial ITS

Gambar 4.37 merupakan hasil pohon klasifikasi maksimal yang terbentuk dari data *training*. Dapat diketahui bahwa ukuran pohon yang terbentuk cukup besar. *Terminal node* yang terbentuk berjumlah 15 *nodes* dengan kedalaman (*depth*) pohon sebesar 13. Variabel yang menjadi pemilah utama dalam pohon klasifikasi maksimal pada Gambar 4.26 adalah variabel Nilai Fisika, dimana variabel tersebut memiliki skor tertinggi dalam variabel terpenting yang dapat dilihat pada Lampiran 10.

Pohon klasifikasi yang berukuran besar perlu dilakukan pemangkasan pohon agar dapat memperoleh pohon klasifikasi yang optimal. Pemangkasan dilakukan dengan melihat nilai minimum dari *relative cost* setiap *terminal nodes*. Gambar 4.38

menunjukkan nilai *relative cost* yang terbentuk pada setiap *terminal nodes*.



Gambar 4.38 *Relative cost Plot Data Peminat Aktuarial ITS*

Hasil perhitungan nilai *relative cost* menunjukkan bahwa nilai minimum terjadi pada saat pohon klasifikasi memiliki 2 *terminal node*, dengan nilai *relative cost* sebesar 0,626. Artinya bahwa pohon klasifikasi dengan 2 *terminal node* memiliki kesalahan klasifikasi yang lebih kecil dibandingkan dengan *terminal node* lainnya. Dengan demikian, dilakukan pemangkasan pohon dari pohon klasifikasi maksimal dengan jumlah *terminal node* sebesar 15 menjadi 2 *terminal node*. Nilai *relative cost* pada setiap *terminal node* dapat dilihat pada Lampiran 10. Hasil pohon klasifikasi optimal dapat dilihat pada Gambar 4.39.

Node 1			Terminal Node 1			Terminal Node 2		
Class = Tidak Diterima			Class = Tidak Diterima			Class = Diterima		
FISIKA <= 552.345			N = 510			N = 215		
Class	Cases	%	Class	Cases	%	Class	Cases	%
Tidak Diterima	710	97.9	Tidak Diterima	509	99.8	Tidak Diterima	201	93.5
Diterima	15	2.1	Diterima	1	0.2	Diterima	14	6.5
N = 725								

Gambar 4.39 Pohon Klasifikasi Optimal Data Peminat Aktuarial ITS

Pohon klasifikasi optimal pada Gambar 4.39 menunjukkan bahwa variabel Nilai Fisika menjadi nilai mata uji yang utama yang digunakan dalam memprediksi status penerimaan peserta. Simpul utama terdiri dari 725 peserta, kemudian dipilah menjadi

2 simpul anak yang menjadi *terminal node* pula. Pada *Terminal node* 1 sebanyak 510 peserta diklasifikasikan sebagai peserta yang tidak diterima prodi Aktuaria dalam tes tulis masuk PTN 2018. *Terminal node* 2 menunjukkan bahwa peserta diklasifikasikan sebagai peserta yang diterima. Karakteristik peserta yang tidak diterima adalah memiliki nilai Fisika $\leq 552,345$ dan sebaliknya peserta dengan nilai Fisika $> 552,345$ diklasifikasikan sebagai peserta yang diterima prodi Aktuaria pada tes tulis masuk PTN tahun 2018.

Hasil klasifikasi berdasarkan analisis dengan *classification tree* kemudian dilakukan evaluasi ketepatan klasifikasi. Hasil evaluasi ketepatan klasifikasi data peminat Aktuaria ITS dapat dijelaskan pada Tabel 4.10.

Tabel 4.10 Evaluasi Ketepatan Klasifikasi Peminat Aktuaria ITS

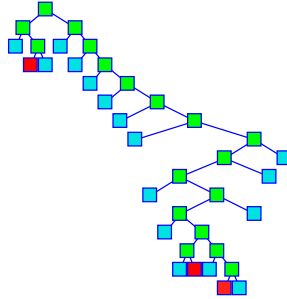
	Akurasi	Sensitivity	Specificity	AUC
<i>Training</i>	0,721	0,717	0,933	0,825
<i>Testing</i>	0,707	0,708	0,667	0,688

Hasil akurasi yang menunjukkan nilai yang cukup baik pada data *testing* maupun *training*, yakni 72,1% dan 70,7%. Nilai tersebut didapat karena ada nya kesalahan klasifikasi baik pada data *training* maupun *testing*. Aturan klasifikasi yang terbentuk adalah apabila nilai nilai Fisika $\leq 552,345$, maka peserta dinyatakan tidak diterima. Pada Gambar 4.16 menunjukkan bahwa pada peserta yang diterima terdapat nilai kurang dari batas tersebut, sehingga dengan menggunakan hasil *classification tree*, peminat tersebut masuk dalam peserta yang tidak diterima. Sama halnya dengan peminat yang diterima memiliki syarat nilai Fisika $> 552,345$. Gambar 4.16 menunjukkan bahwa peminat yang tidak diterima banyak yang memiliki nilai Fisika yang lebih besar dari 552,345.

Adanya kesalahan klasifikasi yang cukup banyak, menyebabkan nilai AUC pada data *testing* masuk dalam kategori lemah, yakni dengan nilai 68,8%.

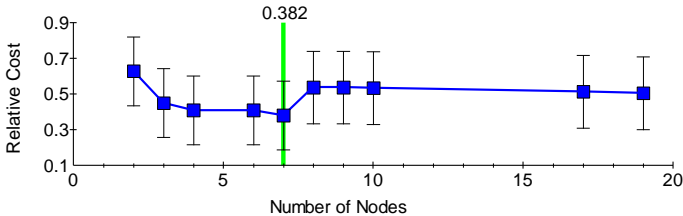
b. Aktuaria Institut Pertanian Bogor

Data *training* yang digunakan dalam analisis *classification tree* pada data peminat Statistika IPB sebesar 908 dan sisanya sebesar 421 merupakan data *testing*. Pembentukan pohon klasifikasi maksimal yang dibangun dari data *training* dapat dilihat pada Gambar 4.40.



Gambar 4. 40 Pohon Klasifikasi Maksimal Prodi Statistika IPB

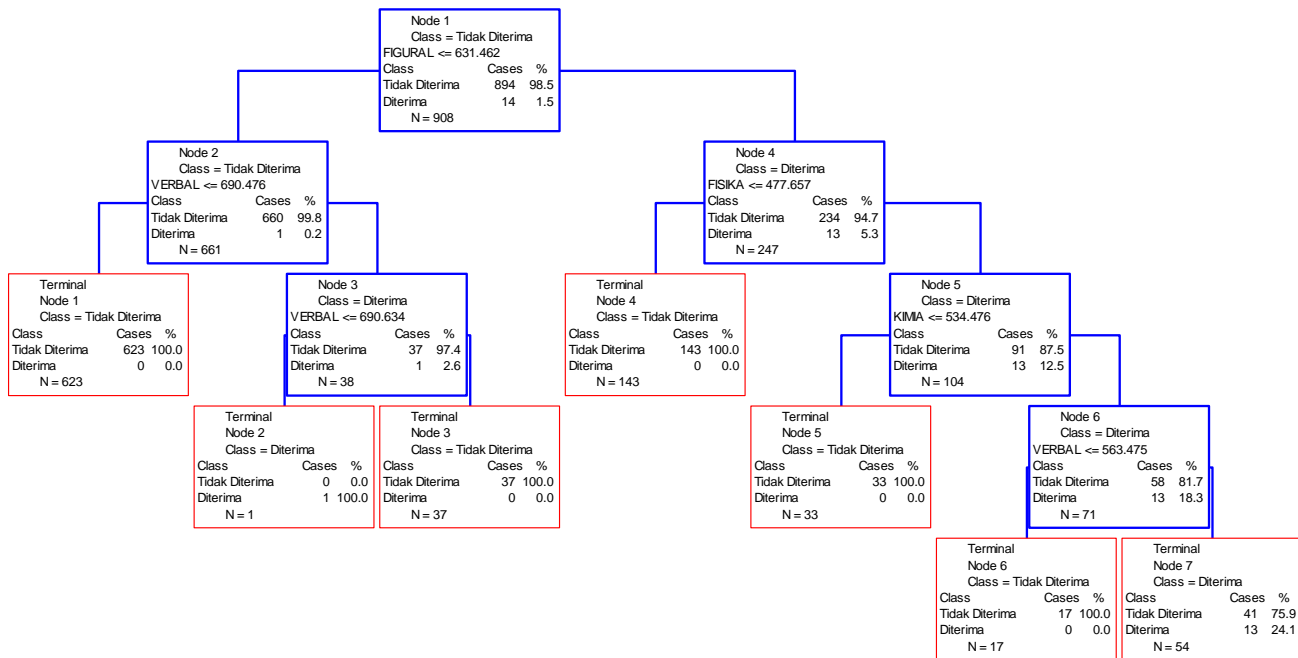
Pohon klasifikasi maksimal yang ditunjukkan pada Gambar 4.40 memiliki jumlah *terminal nodes* sebanyak 19 dengan tingkat kedalaman pohon sebesar 16. Variabel Nilai Figural menjadi pemilah utama dalam pembentukan pohon klasifikasi. Pada Lampiran 11 dapat diketahui skor masing-masing variabel yang menunjukkan kontribusi variabel dalam pembentukan pohon. Nilai Figural memiliki skor tertinggi yang menunjukkan bahwa nilai Figural memiliki kontribusi yang besar dalam pembentukan pohon klasifikasi dan menjadi pemilah utama. Ukuran pohon yang besar diperlukan pemangkasan Selanjutnya akan dilakukan pemangkasan pohon dengan melihat nilai *relative cost* yang paling rendah pada setiap nodes yang ada.



Gambar 4.41 *Relative cost Plot* Data Peminat Statistika IPB

Gambar 4.41 menunjukkan bahwa nilai *relative cost* terendah berada pada node ke-7 dengan nilai sebesar 0,382. Hal ini menunjukkan bahwa pemangkasan pohon menghasilkan pohon klasifikasi optimal dengan 7 *terminal node*. Pada Lampiran 11 dapat diketahui hasil nilai *relative cost* pada masing-masing *terminal node*. Gambar 4.42 merupakan hasil pohon klasifikasi optimal. Pohon klasifikasi optimal yang terbentuk memiliki 7 *terminal node*, yang terdiri dari 5 *terminal node* diklasifikasikan tidak diterima dan 2 *terminal node* diklasifikasikan diterima. Aturan dalam penerimaan peserta tes tulis dengan pilihan prodi Aktuaria IPB yang dihasilkan dari pohon klasifikasi optimal dapat diuraikan sebagai berikut.

1. *Terminal node* 1, 3, 4, 5 dan 6 merupakan *terminal node* dengan klasifikasi tidak diterima. Karakteristik pada masing-masing *terminal node* secara berurutan dijelaskan pada uraian berikut.
 - a. $\text{Figural} \leq 631,462$, dan $\text{Verbal} \leq 690,476$
 - b. $\text{Figural} \leq 631,462$, dan $\text{Verbal} > 690,634$
 - c. $\text{Figural} > 631,462$, dan $\text{Fisika} \leq 690,476$
 - d. $\text{Figural} > 631,462$, $\text{Fisika} > 690,476$, dan $\text{Kimia} \leq 534,476$
 - e. $\text{Figural} > 631,462$, $\text{Fisika} > 690,476$, $\text{Kimia} > 534,476$, dan $\text{Verbal} \leq 563,475$
2. Peserta yang diklasifikasikan diterima pada prodi Aktuaria IPB masuk dalam *terminal node* 2 dan 7. Aturan yang dihasilkan adalah
 - a. $\text{Figural} \leq 631,462$, dan $\text{Verbal} \leq 690,634$
 - b. $\text{Figural} > 631,462$, $\text{Fisika} > 690,476$, $\text{Kimia} > 534,476$, dan $\text{Verbal} > 563,475$



Gambar 4. 42 Pohon Klasifikasi Optimal Data Peminat Aktuaria IPB

Selanjutnya dilakukan pengukuran ketepatan klasifikasi dari pohon klasifikasi optimal yang telah terbentuk. Tabel 4.11 menunjukkan hasil klasifikasi pada data *training* dan data *testing*.

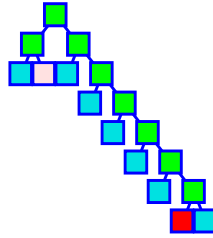
Tabel 4.11 Evaluasi Ketepatan Klasifikasi Peminat Aktuarial IPB

	Akurasi	Sensitivity	Specificity	AUC
<i>Training</i>	0,955	0,954	1,000	0,977
<i>Testing</i>	0,948	0,952	0,667	0,810

Akurasi yang dihasilkan bernilai cukup tinggi, hal ini dapat dilihat bahwa nilai akurasi mendekati nilai 1 baik pada data *testing* maupun data *training*. Namun demikian, nilai AUC pada data *testing* menunjukkan nilai yang lebih rendah. Penyebab rendahnya nilai AUC pada data *testing* adalah nilai *specificity* yang cukup rendah. *Specificity* menunjukkan ketepatan klasifikasi peminat yang diterima. Nilai 66,7% pada *specificity* disebabkan karena hasil aturan pada pohon klasifikasi maksimal menunjukkan bahwa peserta yang diterima memiliki nilai Figural $\leq 631,462$, dan Verbal $\leq 690,634$, serta Figural $> 631,462$, Fisika $> 690,476$, Kimia $> 534,476$, dan Verbal $> 563,475$. Dapat dilihat pada Gambar 4.17, pada peminat yang diterima banyak yang memiliki karakteristik nilai minimal pada mata uji Fisika adalah 540,4 dengan nilai rata-rata 601,4. Peminat yang diterima dengan nilai Fisika kurang dari 690,476 termasuk dalam anggota peserta yang tidak diterima. Namun demikian, kekuatan klasifikasi pada data *testing* masih tergolong pada kategori klasifikasi yang baik.

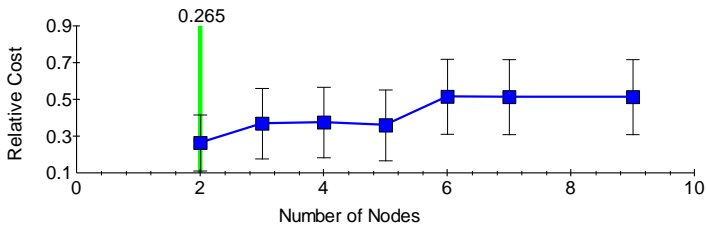
c. Aktuarial Universitas Indonesia

Pembagian data *training* dan *testing* pada data peminat Statistika UI adalah 1366 untuk data *training* dan 659 untuk data *testing*. Data *training* digunakan untuk pembentukan pohon, sedangkan data *testing* digunakan untuk melakukan evaluasi ketepatan klasifikasi. Hasil pembentukan pohon klasifikasi maksimal peminat Aktuarial UI pada tes tulis masuk PTN 2018 dapat dilihat pada Gambar 4.43.



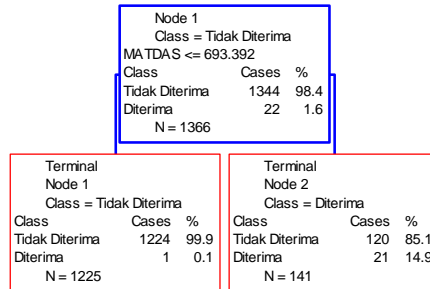
Gambar 4. 43 Pohon Klasifikasi Maksimal Data Peminat Aktuaria UI

Pohon klasifikasi maksimal pada data peminat Aktuaria UI menunjukkan bahwa *terminal nodes* yang terbentuk berjumlah 9 nodes dengan kedalaman sebesar 8. Pohon yang terbentuk menggunakan variabel nilai Matematika Dasar sebagai pemilah utama. Variabel Matematika dasar selain memiliki tingkat keheterogenan yang tinggi, juga memiliki skor kontribusi tertinggi dalam pembentukan pohon (Lampiran 12). Pohon klasifikasi apabila memiliki jumlah *terminal nodes* yang terlalu banyak dikhawatirkan dapat kasus *overfitting*. Sehingga perlu dilakukan pemangkasan pohon. Pemangkasan pohon dilakukan dengan melihat nilai *relative cost* yang paling minimum pada setiap *terminal nodes*.



Gambar 4. 44 *Relative cost Plot* Data Peminat Aktuaria UI

Nilai minimum *relative cost* berada pada *terminal node* ke-2 dengan nilai sebesar 0,265. Untuk lebih jelasnya, nilai *relative cost* pada masing-masing *terminal node* dapat dilihat pada Lampiran 12 Pohon klasifikasi optimal yang akan membentuk 2 *terminal nodes*. Hasil pemangkasan pohon dapat dilihat pada Gambar 4.45.



Gambar 4. 45 Pohon Klasifikasi Optimal Data Peminat Aktuarial UI

Gambar 4.45 adalah pohon klasifikasi optimal yang terbentuk dari data *training* peminat Aktuarial UI pada tes tulis masuk PTN 2018. Nilai Matematika Dasar menjadi pemilah utama dalam pengklasifikasian status diterimanya peserta. Pohon klasifikasi optimal membentuk 2 *terminal node*, dimana *terminal node* 1 merupakan klasifikasi peserta yang tidak diterima dan *terminal node* 2 termasuk dalam klasifikasi peserta diterima. Karakteristik peserta yang tidak diterima memiliki nilai Matematika Dasar $\leq 693,392$, sedangkan untuk peserta yang diklasifikasikan sebagai peserta yang diterima prodi Aktuarial UI memiliki nilai Matematika Dasar $> 693,392$.

Hasil klasifikasi pohon optimal kemudian dilakukan validasi dengan menggunakan data *testing*, dengan tujuan untuk mengetahui apakah pohon yang dihasilkan dapat digunakan dalam penentuan status penerimaan pada data peminat Aktuarial UI.

Tabel 4.12 Evaluasi Ketepatan Klasifikasi Peminat Aktuarial UI

	Akurasi	Sensitivity	Specificity	AUC
<i>Training</i>	0,911	0,911	0,955	0,933
<i>Testing</i>	0,901	0,902	0,833	0,868

Tabel 4.12 menunjukkan bahwa nilai akurasi yang dihasilkan cukup tinggi, yakni 91,1% untuk data *training* dan 90,1% untuk *testing*. Kesalahan klasifikasi yang terjadi tidak terlalu banyak, sehingga nilai akurasi yang dihasilkan cukup tinggi. Kesalahan klasifikasi disebabkan karena terdapat nilai

Matematika Dasar peminat yang tidak diterima lebih dari batas 693,392, bahkan Gambar 4.18 menunjukkan bahwa terdapat nilai mencapai 1022,91. Selain itu, terdapat pula nilai yang kurang dari batasan pada peminat yang diterima, yang menyebabkan terjadinya kesalahan klasifikasi. Nilai AUC pada data *testing* masuk dalam kategori kekuatan klasifikasi yang baik, yakni dengan nilai 86,8%.

4.3 Perbandingan Batasan Nilai Peminat yang Diterima pada Prodi Statistika dan Matematika

Perbandingan batasan nilai mata uji peminat yang diterima prodi Statistika, yang dihasilkan dari pohon klasifikasi optimal pada masing-masing Perguruan Tinggi dapat disajikan pada Tabel 4.13. Karakteristik nilai yang digunakan merupakan karakteristik nilai pada *terminal node* dengan peminat yang diterima memiliki jumlah paling banyak.

Tabel 4.13 Perbandingan Batasan Nilai Penerimaan Peserta Tes Tulis Masuk PTN 2018 Prodi Statistika

	ITS	IPB	UI
Pemilah Utama	Numerik	Figural	Numerik
Karakteristik Nilai	Numerik > 645,86		Numerik > 633,41
	Kimia > 457,82	Figural > 634,94	Kimia > 586,56
	Figural > 589,28	Biologi > 506,07	Verbal > 527,29
	MatIPA > 473,04	Kimia > 541,94	MatDas > 528,43
	Verbal > 612,28		

Tabel 4.13 menunjukkan bahwa ketiga PTN memiliki kesamaan dalam variabel terpenting yang digunakan dalam menentukan penerimaan peserta tes tulis masuk PTN tahun 2018, yakni pada mata uji Tes Potensial Akademik. TPA berupa Numerik, yakni kemampuan pemahaman dan penalaran dalam bentuk angka, merupakan variabel terpenting dalam penentuan diterimanya di Statistika ITS dan UI. Sedangkan untuk Statistika IPB menggunakan nilai Figural, yakni kemampuan penalaran dalam

bentuk gambar sebagai penentu utama dalam penerimaan peserta tes tulis masuk PTN 2018.

Dilihat dari karakteristik nilai ITS dan UI yang sama-sama menggunakan nilai Numerik sebagai penentu utama, memiliki kriteria atau batasan nilai mata uji yang lebih banyak dibandingkan dengan UI. Hal ini dapat diartikan bahwa keketatan dalam penentuan penerimaan semakin ketat, karena melihat nilai dari beberapa mata uji. Selain menggunakan TPA numerik, klasifikasi peminat Statistika ITS dan UI juga menggunakan nilai TPA Verbal, dimana Statistika ITS pada kedua mata uji ini memiliki nilai batasan lebih tinggi dibandingkan dengan Statistika UI. Artinya bahwa kemampuan pemahaman dan penalaran dalam bentuk angka dan numerik pada peminat Statistika ITS yang diterima lebih tinggi.

Pada ketiga PTN, nilai mata uji Kimia menjadi penentu pula dalam diterimanya ke dalam prodi Statistika. Mata uji Matematika merupakan mata uji yang dekat dengan rumpun Statistika, yang seharusnya menjadi salah satu penentu dalam penerimaan peserta. Pada Statistika ITS, penentuan menggunakan nilai Matematika IPA, Statistika UI menggunakan nilai Matematika Dasar, dan Statistika IPB tidak menggunakan nilai matematika dalam penentuan penerimaan peserta tes tulis tahun 2018.

Perbandingan hasil penentuan batas minimal untuk diterima prodi Matematika dengan menggunakan *Classification Tree* dapat dijelaskan pada Tabel 4.14.

Tabel 4.14 Perbandingan Batasan Nilai Penerimaan Peserta Tes Tulis Masuk PTN 2018 Prodi Matematika

	ITS	IPB	UI
Pemilah Utama	Kimia	Kimia	Numerik
Karakteristik Nilai	Kimia > 495,33 Verbal > 526,07	Kimia > 523,26 Figural > 590,86	Numerik > 650,21 Fisika > 415,59 Figural > 561,08

Hasil perbandingan menunjukkan bahwa variabel terpenting yang digunakan sebagai pemilah utama dalam analisis *classification tree* pada penerimaan peminat prodi Matematika melalui jalur tes tulis tahun 2018 pada ketiga PTN memiliki perbedaan. Penerimaan pada Matematika ITS dan IPB menggunakan nilai Kimia sebagai penentu utama, sedangkan Matematika UI menggunakan nilai TPA Numerik. Batas nilai kimia pada penentuan diterimanya Matematika ITS lebih rendah dibandingkan dengan Matematika IPB.

Karakteristik nilai batasan penerimaan peminat Matematika UI lebih banyak dibandingkan dengan ITS dan IPB. Artinya dalam menyeleksi penerimaan peserta, Matematika UI lebih ketat, karena menilai penerimaan berdasarkan 3 kriteria nilai. Dari ketiga PTN, terdapat salah satu mata uji TPA dalam penentuan penerimaan. Matematika ITS menggunakan nilai TPA Verbal, IPB menggunakan TPA Figural, dan UI menggunakan TPA Numerik dan Figural. Namun demikian, tidak ada mata uji Matematika dalam penentuan penerimaan di ketiga PTN tersebut.

Hasil perbandingan penentuan batas minimal nilai ujian tes tulis untuk diterima prodi Aktuaria dengan menggunakan *Classification Tree* dapat dijelaskan pada Tabel 4.15.

Tabel 4.15 Perbandingan Batasan Nilai Penerimaan Peserta Tes Tulis Masuk PTN 2018 Prodi Aktuaria

	ITS	IPB	UI
Pemilah Utama	Fisika	Figural	Matematika Dasar
Karakteristik Nilai	Fisika > 552,35	Figural > 631,46 Fisika > 690,48 Kimia > 534,48 Verbal > 563,48	MatDas > 693,39

Tabel 4.15 menunjukkan bahwa pemilah utama atau variabel dengan tingkat kepentingan tertinggi pada ketiga PTN berbeda. Hasil pohon klasifikasi optimal menunjukkan bahwa penerimaan pada Aktuaria ITS hanya menggunakan nilai Fisika sebagai penentu utama, sama halnya dengan Aktuaria UI yang menggunakan satu nilai mata uji, yakni Matematika Dasar.

Berbeda dengan Aktuaria IPB yang menggunakan nilai Figural, Fisika, Kimia, dan Verbal dalam mengklasifikasikan peminat yang diterima. Artinya bahwa tingkat seleksi penerimaan pada Aktuaria IPB lebih tinggi, karena memperhatikan beberapa nilai mata uji lainnya.

Secara keseluruhan, baik prodi Statistika, Matematika, dan Aktuaria pada ketiga PTN, dalam menentukan penerimaan peserta tes tulis masuk PTN 2018 menggunakan salah satu nilai dari nilai TPA. Nilai TPA sangat diperhatikan karena TPA merupakan salah satu tes yang digunakan dalam mengukur tingkat kemampuan dasar yang dapat menggambarkan atau memprediksi bagaimana calon mahasiswa akan tampil di tahun-tahun berikutnya selama masa perkuliahan. Namun pada prodi Aktuaria, dimana penentuan batasan penerimaan dengan menggunakan nilai TPA hanya Aktuaria IPB.

(Halaman ini sengaja dikosongkan)

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Kesimpulan yang dapat diperoleh dari hasil analisis penelitian ini dapat dijabarkan sebagai berikut.

1. Tingkat keketatan yang tinggi ditunjukkan pada prodi Statistika IPB, Matematika IPB, dan Aktuaria UI. Berdasarkan jenis kelamin, peminat ketiga prodi secara keseluruhan didominasi oleh jenis kelamin perempuan di ketiga PTN. Berdasarkan asal sekolah, peminat ketiga prodi baik peminat keseluruhan atau peminat yang diterima didominasi oleh peserta dari SMA. Secara keseluruhan nilai TPA, peminat Statistika, Matematika, dan Aktuaria ITS memiliki kemampuan yang lebih baik dimana banyaknya peminat dengan nilai yang berada disekitar nilai rata-rata. Nilai TKD yakni nilai Biologi, Kimia, dan Matematika Dasar pada ketiga prodi menunjukkan rata-rata tertinggi berada pada peminat UI, sedangkan nilai Fisika tertinggi berada pada peminat ITS.
2. Batasan nilai yang dihasilkan dari *Classification Tree* untuk dapat diterima adalah sebagai berikut.
 - a. Statistika ITS, batasan nilai : Numerik $> 645,856$; Kimia $> 457,820$; TPA Figural $> 589,281$; Matematika IPA $> 473,041$; dan TPA Verbal $> 612,259$.
 - b. Statistika IPB, adalah Figural $> 634,938$; Biologi $> 506,065$; Kimia $> 541,940$.
 - c. Statistika UI, memiliki batasan nilai Numerik $> 633,405$; Kimia $> 586,563$; TPA Verbal $> 527,287$; Matematika Dasar $> 528,426$.
 - d. Matematika ITS, berupa Kimia $> 495,332$; dan TPA Verbal $> 526,073$.
 - e. Matematika IPB, Kimia $> 523,260$ dan TPA Figural $> 590,864$.
 - f. Matematika UI, TPA Numerik $> 650,209$; Fisika $> 415,590$; TPA Figural $> 561,082$.

- g. Aktuaria ITS, Fisika > 552,35
 - h. Aktuaria IPB, Figural > 631,462, Fisika > 690,476, Kimia > 534,476, dan Verbal > 563,475
 - i. Aktuaria UI, Matematika Dasar > 693,39
3. Prodi Statistika, Matematika, dan Aktuaria pada ketiga PTN, dalam menentukan penerimaan peserta tes tulis masuk PTN 2018 menggunakan salah satu nilai dari nilai TPA. Nilai TPA sangat diperhatikan karena TPA merupakan salah satu tes yang digunakan dalam mengukur tingkat kemampuan dasar yang dapat menggambarkan atau memprediksi bagaimana calon mahasiswa akan tampil di tahun-tahun berikutnya selama masa perkuliahan. Namun pada prodi Aktuaria, dimana penentuan batasan penerimaan dengan menggunakan nilai TPA hanya Aktuaria IPB.

5.2 Saran

Saran yang dapat diberikan untuk penelitian selanjutnya adalah dapat dilakukan pengklasifikasian dengan menggunakan penerimaan pada pilihan keberapa peserta diterima sebagai klasifikasi. Hal ini dikarenakan prioritas dalam pilihan prodi merupakan salah satu pertimbangan dalam penerimaan peserta tes tulis masuk perguruan tinggi negeri, selain dengan menggunakan nilai.

DAFTAR PUSTAKA

- Anggreini, Y. (2015). *Pemodelan regresi Logistik Biner Terhadap Peminat ITS di Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN) 2014*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Anonim. (2018). *Introducing CART*. United States: Minitab Inc.
- Bekkar, M., Djemaa, H. K., & Alitouche, T. A. (2013). Evaluation Measures for Models Assessment Over Imbalanced Data Sets. *Journal of Information Engineering and Applications*, 27-38.
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1993). *Classification and Regression Tree*. New York: Chapman & Hall.
- Dahlan, M. S. (2009). *Penelitian Diagnostik*. Jakarta: Salemba Medikka.
- Ghasemzadeh, A., Hammit, B. E., Ahmed, M. M., & Young, R. K. (2018). Parametric Ordinal Logistic Regression and Non-Parametric Decision Tree Approaches for Assessing the Impact of Weather Conditions on Driver Speed Selection Using Naturalistic Driving Data. *Transportation Research Board*. 1-11.
- Han, J., & Kamber, M. (2006). *Data Mining: Concepts and Techniques*. San Fransisco: Morgan Kaufman Publisher.
- Johnson, R., & Winchern, D. (2007). *Applied Multivariate Statistical Analysis (6th Edition)*. New Jersey: Prentice Hall.
- Karpen, S. C., & Ellis. (2018). e Application of Classification Tree to Pharmacy School Admissions The Application of Classification Tree to Pharmacy School Admissions. *Pharmaceutical Education*, 742-745.
- Lewis, R. J. (2000). *An Intoduction to Classification and Regression Tree (CART) Analysis*. San Francisco, California: Anual Meeting Of Society For Academy Emergency Medice.

- Menkeu. (2003). Keputusan Menteri Keuangan Republik Indonesia Nomor 426/KMK.06/2003. Jakarta: Menteri Keuangan Republik Indonesia.
- Pratiwi, L. D. (2015). *Klasifikasi Nilai Peminat SBMPTN (Seleksi Bersama Masuk Perguruan Tinggi Negeri) ITS dengan Pendekatan Classification and Regression Tree (CART)*. Surabaya: Institut Teknologi Sepuluh Nopember.
- Ristekdikti. (2018). *Hasil Klasterisasi Perguruan Tinggi Non Vokasi 2018*. Jakarta: Ristekdikti.
- SBMPTN. (2018). *Seleksi Bersama Masuk Perguruan Tinggi Negeri (SBMPTN) 2018*. Retrieved September 25, 2018, from www.sbmptn.ac.id
- Segal, M. R. (1988). Regression Tress for Censored Data. *Biometrics*. 34-47.
- Steinberg, D. (2019). *What is the Variable Importance Measure?* Retrieved January 3, 2019, from <https://www.salford-systems.com>
- Therneau, T. M., & Atkinson. (1997). *An Introduction to Recursive Partitioning Using the RPART Routine*. Scottsdale AZ: Mayo Foundation.
- Uly, Y. A. (2018). *Hanya 19,8% Peserta SBMPTN 2018 yang Lolos Seleksi*. Retrieved September 24, 2018, from Okezone.com: <https://news.okezone.com/read/2018/07/03/65/1917344/hanya-19-8-peserta-sbmptn-2018-yang-lolos-seleksi>
- Webb, P., & Yohannes, Y. (1999). *Classification and Regression Tress (CART) : A User Manual for Identifying Indicators of Vulnerability to Famine and Chronic Food Insecurity*. Washington D.C: International Food Policy Research Institute.
- Witten, I. H., Frank, E., & Hall, M. A. (2005). *Data Mining Practical Machine (3rd edition)*. USA: Elsevier.

LAMPIRAN

Lampiran 1A. Data Peminat Prodi Statistika ITS

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	527,59	584,85	512,15	663,42	488,29	621,88	429,76	544,04	664,48	470,00	P	SMA	0
2	673,49	770,80	499,75	738,49	474,43	535,42	524,82	628,41	498,75	648,81	L	SMA	1
3	633,19	675,45	539,84	404,53	399,80	442,92	420,15	461,89	542,90	550,01	P	SMA	0
4	538,62	454,38	607,64	471,69	422,81	514,91	471,26	436,36	561,74	538,67	P	SMA	0
5	578,16	611,24	638,37	412,31	420,93	461,11	446,81	423,10	630,40	642,70	P	SMA	0
.
.
.
830	518,82	693,05	358,97	359,57	414,74	484,93	465,37	487,88	530,62	595,03	L	SMA	0
831	593,85	535,39	520,16	387,29	498,66	586,86	391,55	507,81	568,93	521,19	L	MA	0
832	586,67	620,63	651,63	357,93	553,19	388,44	442,64	443,67	493,16	503,37	L	SMA	0
833	657,53	745,68	569,84	579,10	740,10	451,97	473,82	683,55	654,31	480,70	P	SMA	1
834	564,83	624,24	679,14	852,24	654,26	410,13	734,21	589,63	563,56	529,95	L	SMA	0

Lampiran 1B. Data Peminat Prodi Statistika IPB

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	531,61	466,73	568,05	388,69	608,65	638,99	768,25	439,72	430,49	574,27	P	SMA	0
2	580,85	473,60	542,18	332,82	450,08	365,36	499,22	476,33	413,57	571,83	L	SMA	0
3	312,38	467,41	518,68	425,61	495,62	665,80	513,36	530,65	548,83	439,01	L	SMA	0
4	553,29	548,48	522,65	384,31	855,40	466,53	676,08	525,36	513,58	647,59	L	SMA	0
5	500,51	472,70	466,95	449,66	595,77	629,51	491,50	539,00	484,25	447,12	L	SMK	0
.
.
.
873	641,44	542,34	705,51	691,33	610,68	653,07	441,92	437,83	434,48	689,83	P	SMA	0
874	463,31	648,37	492,01	517,67	600,86	394,89	674,58	541,51	560,57	471,20	L	MA	0
875	581,03	517,73	574,30	309,73	508,36	453,93	416,99	412,66	521,50	563,83	P	SMA	0
876	362,62	696,98	703,84	444,20	462,72	362,71	555,93	541,02	501,69	452,27	P	SMA	0

Lampiran 1C. Data Peminat Prodi Statistika UI

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	626,23	658,69	641,13	576,80	479,73	410,83	420,70	345,34	658,97	593,55	P	SMA	0
2	435,77	546,30	345,03	519,83	458,16	566,81	730,28	614,97	491,46	514,27	P	SMA	0
3	370,17	576,65	693,37	574,76	539,57	413,67	705,88	431,26	420,33	557,69	P	SMA	0
4	526,77	572,47	422,45	576,47	465,68	592,22	421,44	515,54	363,11	681,52	P	SMK	0
5	418,58	384,10	268,17	274,69	368,42	439,42	768,25	377,15	643,94	383,95	P	SMK	0
.
.
.
747	574,14	648,61	679,62	448,08	529,37	449,52	539,80	646,63	688,86	508,35	P	SMA	0
748	591,69	675,20	548,65	462,63	733,46	482,12	555,50	538,55	335,02	739,85	P	SMA	0
749	422,20	611,09	630,49	538,87	358,43	436,92	473,35	491,02	562,27	712,55	P	SMA	0
750	534,89	770,27	660,89	330,30	818,27	445,07	564,14	601,68	680,09	749,22	P	SMA	0
751	606,07	627,81	620,75	348,00	494,54	692,42	720,24	431,63	466,72	577,59	P	SMA	0

Lampiran 2A. Data Peminat Prodi Matematika ITS

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	557,22	594,98	648,72	433,58	420,93	461,11	439,78	423,10	497,23	470,00	P	SMA	0
2	603,30	521,03	567,30	556,59	443,04	562,16	449,33	473,82	443,46	609,03	L	SMA	0
3	582,73	645,91	541,76	404,53	372,51	461,02	424,19	389,25	331,95	550,01	L	SMA	0
4	587,47	574,58	493,51	528,37	507,91	596,90	558,92	775,14	676,57	550,53	P	SMA	0
5	549,90	555,97	436,61	332,27	480,37	501,46	385,27	389,48	440,20	469,75	P	SMA	0
.
.
.
599	583,18	530,32	509,16	454,81	447,45	461,11	413,89	459,85	590,49	470,00	L	SMA	0
600	479,95	683,22	460,76	647,68	633,84	595,03	629,58	510,17	470,86	816,72	L	LAIN	0
601	539,62	519,73	723,68	408,63	685,41	487,32	535,15	828,05	540,53	401,39	L	SMA	0
602	686,13	671,92	617,16	500,58	489,19	591,36	476,81	321,21	657,87	425,19	P	SMA	0
603	425,55	601,22	685,92	550,63	543,09	395,21	526,80	535,90	631,87	617,88	P	MA	0

Lampiran 2B. Data Peminat Prodi Matematika IPB

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	604,44	587,94	409,66	474,63	406,74	483,72	447,56	520,48	469,21	484,97	P	MA	0
2	535,94	633,51	349,35	394,30	325,63	423,47	804,47	384,54	350,25	508,16	L	SMA	0
3	509,30	690,56	607,04	419,87	444,31	355,67	435,13	425,38	589,67	412,17	P	SMA	0
4	655,68	603,55	562,10	751,98	458,25	477,72	515,37	431,01	588,04	434,85	L	SMA	0
5	635,55	691,63	485,09	398,05	402,61	425,34	495,85	399,91	709,87	459,70	L	SMA	0
.
.
.
501	705,59	626,32	514,21	721,06	735,58	601,02	549,34	626,68	640,11	588,81	P	SMA	0
502	496,13	711,24	537,06	592,92	547,34	609,84	470,53	440,52	407,90	552,16	P	SMA	0
503	459,53	521,87	492,33	430,85	406,81	596,28	433,67	452,56	476,25	504,96	L	SMA	0
504	600,49	669,31	363,69	499,75	450,34	554,34	655,91	678,07	471,62	479,06	L	SMA	0
505	486,40	602,36	552,66	574,76	517,89	429,69	664,63	506,03	490,32	648,41	L	SMA	0

Lampiran 2C. Data Peminat Prodi Matematika UI

No	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	507,47	560,28	475,84	496,67	506,90	362,54	565,94	570,77	467,53	490,42	L	SMA	0
2	529,69	548,04	441,83	461,91	395,56	452,93	515,98	447,76	235,79	578,17	L	SMA	0
3	230,05	644,84	493,65	337,60	601,68	383,21	574,46	512,59	519,66	477,52	L	SMA	0
4	424,60	371,69	321,80	550,31	536,39	415,93	467,42	440,20	490,65	416,06	P	MA	0
5	491,51	546,26	517,26	553,33	442,37	520,55	512,21	545,36	284,51	502,58	L	SMA	0
.
.
.
827	550,22	662,52	529,18	499,46	517,42	329,25	442,79	443,50	460,12	589,55	L	SMA	0
828	551,76	552,32	648,98	430,50	441,51	429,45	465,32	348,55	426,96	451,85	L	SMA	0
829	559,39	565,15	462,60	444,00	420,75	495,26	564,74	472,97	480,46	633,68	L	SMK	0
830	487,35	637,32	441,45	420,09	480,34	425,83	449,26	396,15	483,01	391,79	L	SMA	0
831	520,05	631,84	519,18	745,59	594,15	787,35	689,40	443,09	410,86	750,20	P	SMA	0

Lampiran 3A. Data Peminat Prodi Aktuaria ITS

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	664,92	657,51	613,28	402,52	506,26	465,60	449,99	603,62	517,82	539,89	L	SMA	0
2	687,74	480,75	522,11	449,16	434,93	485,42	486,71	508,61	364,11	442,94	L	SMA	0
3	533,95	507,80	545,16	494,26	409,41	483,70	495,62	579,08	622,82	500,54	L	SMA	0
4	532,07	552,62	471,71	328,86	397,10	484,93	456,21	432,52	480,54	529,31	P	SMA	0
5	510,35	581,55	530,19	552,04	437,35	613,76	405,12	561,32	529,76	461,97	P	MA	0
.
.
.
1062	465,39	600,74	627,74	589,77	504,26	521,54	472,21	358,95	542,49	518,62	P	SMA	0
1063	447,61	623,96	366,34	594,12	631,89	564,56	530,90	588,42	549,36	581,05	L	SMA	0
1064	563,14	666,74	548,77	393,68	424,80	495,34	582,62	469,00	432,93	467,11	P	SMA	0
1065	526,33	555,38	586,10	436,97	529,60	511,65	437,54	779,99	559,60	619,90	P	SMA	0
1066	617,40	532,51	584,40	638,31	490,90	547,06	480,45	567,91	659,47	569,09	P	SMA	0

Lampiran 3B. Data Peminat Prodi Aktuaria IPB

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	601,66	581,66	525,46	501,49	620,45	427,75	768,25	668,00	649,11	531,28	P	SMA	0
2	414,43	729,83	623,13	744,81	473,65	467,62	484,44	434,51	512,74	445,11	L	SMA	0
3	489,50	557,67	612,40	503,34	583,33	525,49	481,18	566,04	516,47	343,05	P	SMA	0
4	568,36	591,83	644,47	444,92	332,59	408,42	543,69	555,77	529,26	633,66	P	SMA	0
5	616,74	548,88	689,55	363,47	409,98	383,96	480,77	487,13	363,27	411,01	L	SMA	0
.
.
.
1325	479,66	689,82	531,78	745,75	591,15	453,45	648,50	390,75	381,81	428,60	P	SMA	0
1326	677,75	638,78	674,72	579,11	457,82	389,33	511,28	506,47	604,37	460,82	P	SMA	0
1327	508,24	729,36	582,32	345,20	782,86	382,19	524,58	865,23	530,35	508,70	L	SMA	0
1328	541,78	594,10	590,40	454,25	507,51	480,98	591,79	782,73	559,88	448,80	P	SMA	0
1329	628,86	700,45	692,68	606,06	579,64	440,72	479,43	576,84	682,61	443,66	P	SMA	0

Lampiran 3C. Data Peminat Prodi Aktuaria UI

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	Y
1	455,00	534,89	523,23	603,10	414,87	449,46	491,66	414,74	561,73	518,72	P	SMA	0
2	480,89	357,80	587,31	734,91	490,31	383,53	663,39	518,45	544,87	585,58	P	SMA	0
3	554,95	529,07	518,41	588,61	465,99	431,94	587,71	479,04	387,15	453,81	P	SMA	0
4	553,90	580,93	488,57	492,94	567,02	610,73	430,20	400,30	302,61	560,05	P	SMA	0
5	641,83	745,94	749,26	372,67	661,59	512,96	408,94	580,72	461,44	501,22	L	SMA	0
.
.
.
2021	691,47	691,98	605,82	680,39	691,34	503,84	573,19	713,25	599,29	712,64	L	SMA	0
2022	601,32	619,32	674,99	499,03	421,84	502,99	541,96	511,30	562,12	374,91	L	SMA	0
2023	603,70	531,97	473,44	446,92	529,64	394,22	639,60	506,17	471,50	410,30	P	SMA	0
2024	675,97	535,83	698,00	419,63	739,43	589,59	494,19	529,93	552,35	405,25	L	SMA	0
2025	706,37	654,25	568,04	522,70	712,00	622,43	734,40	626,09	661,44	856,54	L	SMA	0

Lampiran 4. Output CART Hasil Analisis Data Prodi Statistika ITS

Output Relative Cost

Tree Number	Terminal Nodes	Test Set <i>Relative cost</i>	Resubstitution <i>Relative cost</i>	Complexity
1	21	0,674 ± 0,143	0,041	-1,000
2	14	0,689 ± 0,144	0,054	0,000956
3	12	0,693 ± 0,144	0,062	0,002
4**	11	0,530 ± 0,144	0,067	0,003
5	10	0,538 ± 0,144	0,080	0,007
6	8	0,581 ± 0,144	0,116	0,009
7	6	0,621 ± 0,145	0,153	0,009
8	5	0,605 ± 0,144	0,175	0,011
9	4	0,661 ± 0,146	0,221	0,023
10	3	0,636 ± 0,145	0,284	0,032
11	2	0,615 ± 0,129	0,445	0,080
12	1	1,000 ± 8,63E-005	1,000	0,277

** Pohon klasifikasi Optimal

Output Informasi Terminal Node Pohon Klasifikasi Optimal

```

*****
*           Node 1: NUMERIK           *
*           N: 567                     *
*****
*****
*           Node 2           *           *           Node 5           *
*           N: 350           *           *           N: 217           *
*****
Node 1 was split on NUMERIK
A case goes left if NUMERIK <= 645.856
Improvement = 0.165           Complexity Threshold = 0.277

Node      Cases  Wgt Counts      Cost  Class
1         567    567.00      0.500 Tidak Diterima
2         350    350.00      0.126 Tidak Diterima
5         217    217.00      0.279 Diterima

                          Weighted Counts
Class      Top      Left      Right
Tidak Diterima  535.00  347.00  188.00
Diterima       32.00   3.00    29.00

                          Within Node Probabilities
Class      Top      Left      Right
Tidak Diterima  0.500  0.874  0.279
Diterima       0.500  0.126  0.721

Surrogate           Split           Assoc.           Improve.
1 FIGURAL s        581.994           0.355           0.160
2 MATDAS s         479.349           0.186           0.070
3 KIMIA s          484.928           0.182           0.121
4 VERBAL s         499.592           0.134           0.030
5 BIOLOGI s        455.549           0.132           0.104

Competitor          Split           Improve.
1 FIGURAL           581.994           0.160
2 KIMIA             474.311           0.149
    
```


3 MATIPA	473.778	0.125
4 BIOLOGI	475.807	0.116
5 VERBAL	633.473	0.092

```
*****
*           Node 2: MATDAS           *
*           N: 350                    *
*****
```

```
*****
*           Node 3           * *           Node 4           *
*           N: 339           * *           N: 11           *
*****
```

Node 2 was split on MATDAS

A case goes left if MATDAS <= 732.691

Improvement = 0.039 Complexity Threshold = 0.023

Node	Cases	Wgt	Counts	Cost	Class
2	350		350.00	0.126	Tidak Diterima
3	339		339.00	0.047	Tidak Diterima
4	11		11.00	0.212	Diterima

Weighted Counts

Class	Top	Left	Right
Tidak Diterima	347.00	338.00	9.00
Diterima	3.00	1.00	2.00

Within Node Probabilities

Class	Top	Left	Right
Tidak Diterima	0.874	0.953	0.212
Diterima	0.126	0.047	0.788

Surrogate	Split	Assoc.	Improve.
1 MATIPA s	824.270	0.300	0.019
2 VERBAL s	682.279	0.229	0.014
3 B_INGGRI s	736.433	0.205	0.013
4 KIMIA s	657.985	0.135	0.010
5 NUMERIK s	645.551	0.024	.298917E-04

Competitor	Split	Improve.
1 B_INDONE	655.152	0.036
2 B_INGGRI	697.337	0.028
3 MATIPA	812.159	0.024
4 FIGURAL	652.287	0.023
5 BIOLOGI	528.357	0.016

```
*****
*           Node 3: BIOLOGI           *
*           N: 339                    *
*****
```

```
=====
=           Terminal node 1           =           Terminal node 2           =
=           N: 333                    =           N: 6                      =
=====
```

Node 3 was split on BIOLOGI

A case goes left if BIOLOGI <= 685.500

Improvement = 0.023 Complexity Threshold = 0.011

Node	Cases	Wgt	Counts	Cost	Class
3	339		339.00	0.047	Tidak Diterima
-1	333		333.00	0.000	Tidak Diterima
-2	6		6.00	0.230	Diterima

Weighted Counts

Class	Top	Left	Right
Tidak Diterima	338.00	333.00	5.00
Diterima	1.00	0.00	1.00

Within Node Probabilities

Class	Top	Left	Right
Tidak Diterima	0.953	1.000	0.230

Diterima	0.047	0.000	0.770		
Surrogate	Split			Assoc.	Improve.
1 B_INDONE	s 692.867			0.355	0.019
Competitor	Split				Improve.
1 B_INDONE	692.867				0.019
2 B_INGGRI	697.337				0.013
3 MATIPA	812.159				0.011
4 FIGURAL	656.424				0.011
5 FISIKA	406.240				0.007

 * Node 4: VERBAL *
 * N: 11 *

```

=====
= Terminal node 3 = = Terminal node 4 =
= N: 7 = = N: 4 =
=====
  
```

Node 4 was split on VERBAL
 A case goes left if VERBAL <= 557.106
 Improvement = 0.010 Complexity Threshold = 0.007

Node	Cases	Wgt	Counts	Cost	Class
4	11		11.00	0.212	Diterima
-3	7		7.00	0.000	Tidak Diterima
-4	4		4.00	0.056	Diterima

Weighted Counts

Class	Top	Left	Right
Tidak Diterima	9.00	7.00	2.00
Diterima	2.00	0.00	2.00

Within Node Probabilities

Class	Top	Left	Right
Tidak Diterima	0.212	1.000	0.056
Diterima	0.788	0.000	0.944

Surrogate	Split			Assoc.	Improve.
1 B_INDONE	s 654.309			0.714	0.010
2 BIOLOGI	s 507.728			0.714	0.007
3 FISIKA	s 546.338			0.714	0.007
4 B_INGGRI	s 506.538			0.571	0.005
5 MATDAS	r 767.487			0.571	0.008
Competitor	Split				Improve.
1 B_INDONE	654.309				0.010
2 MATDAS	767.487				0.008
3 BIOLOGI	507.728				0.007
4 FISIKA	546.338				0.007
5 B_INGGRI	566.983				0.007

 * Node 5: KIMIA *
 * N: 217 *

```

=====
= Terminal node 5 = * Node 6 *
= N: 86 = * N: 131 *
=====
  
```

Node 5 was split on KIMIA
 A case goes left if KIMIA <= 457.820
 Improvement = 0.096 Complexity Threshold = 0.080

Node	Cases	Wgt	Counts	Cost	Class
5	217		217.00	0.279	Diterima
-5	86		86.00	0.000	Tidak Diterima
6	131		131.00	0.174	Diterima

Weighted Counts				
Class	Top	Left	Right	
Tidak Diterima	188.00	86.00	102.00	
Diterima	29.00	0.00	29.00	
Within Node Probabilities				
Class	Top	Left	Right	
Tidak Diterima	0.279	1.000	0.174	
Diterima	0.721	0.000	0.826	

Surrogate	Split	Assoc.	Improve.
1 BIOLOGI s	458.228	0.096	0.065
2 B_INDONE s	443.124	0.070	0.022
3 MATIPA s	446.854	0.027	0.044
4 FIGURAL r	712.384	0.023	0.004
5 VERBAL s	384.783	0.023	0.004

Competitor	Split	Improve.
1 BIOLOGI	473.794	0.068
2 MATIPA	473.041	0.064
3 FIGURAL	581.385	0.063
4 VERBAL	612.426	0.032
5 MATDAS	478.928	0.029

```

*****
*                               *
*           Node 6: FIGURAL     *
*                               *
*           N: 131              *
*                               *
*****

```

```

=====
=           Terminal node 6     = *           Node 7           *
=           N: 34              = *           N: 97           *
=====
*****

```

Node 6 was split on FIGURAL
A case goes left if FIGURAL <= 580.281
Improvement = 0.046 Complexity Threshold = 0.032

Node	Cases	Wgt Counts	Cost	Class
6	131	131.00	0.174	Diterima
-6	34	34.00	0.000	Tidak Diterima
7	97	97.00	0.123	Diterima

Weighted Counts				
Class	Top	Left	Right	
Tidak Diterima	102.00	34.00	68.00	
Diterima	29.00	0.00	29.00	
Within Node Probabilities				
Class	Top	Left	Right	
Tidak Diterima	0.174	1.000	0.123	
Diterima	0.826	0.000	0.877	

Surrogate	Split	Assoc.	Improve.
1 MATDAS s	468.980	0.147	0.017
2 B_INDONE s	444.550	0.088	0.012
3 BIOLOGI s	312.331	0.059	0.003

Competitor	Split	Improve.
1 MATDAS	477.189	0.020
2 BIOLOGI	473.794	0.019
3 MATIPA	473.041	0.017
4 VERBAL	612.259	0.014
5 B_INDONE	444.550	0.012

```

*****
*           Node 7: MATIPA           *
*           N: 97                     *
*****
*****
*           Node 8           *           *           Node 9           *
*           N: 27           *           *           N: 70           *
*****
Node 7 was split on MATIPA
A case goes left if MATIPA <= 473.041
Improvement = 0.011           Complexity Threshold = 0.009

Node      Cases  Wgt Counts      Cost Class
7         97          97.00      0.123 Diterima
8         27          27.00      0.428 Diterima
9         70          70.00      0.087 Diterima

                          Weighted Counts
Class      Top      Left      Right
Tidak Diterima  68.00    25.00    43.00
Diterima      29.00     2.00    27.00

                          Within Node Probabilities
Class      Top      Left      Right
Tidak Diterima  0.123    0.428    0.087
Diterima      0.877    0.572    0.913

Surrogate      Split      Assoc.      Improve.
1 KIMIA        r      842.769    0.286    .487566E-03
2 FISIKA        s      340.288    0.286    .487536E-03
3 MATDAS        s      471.841    0.051     0.007
4 VERBAL        s      408.235    0.034     0.003
5 BIOLOGI       s      374.142    0.017     0.007

Competitor      Split      Improve.
1 BIOLOGI       387.155    0.010
2 VERBAL        612.259    0.009
3 MATDAS        477.189    0.009
4 B INGGRI     576.441    0.005
5 KIMIA         485.555    0.005

*****
*           Node 8: BIOLOGI           *
*           N: 27                     *
*****

=====
=           Terminal node 7           =           Terminal node 8           =
=           N: 20                     =           N: 7                     =
=====

Node 8 was split on BIOLOGI
A case goes left if BIOLOGI <= 583.904

Improvement = 0.019           Complexity Threshold = 0.019

Node      Cases  Wgt Counts      Cost Class
8         27          27.00      0.428 Diterima
-7        20          20.00      0.000 Tidak Diterima
-8         7           7.00      0.130 Diterima

                          Weighted Counts
Class      Top      Left      Right
Tidak Diterima  25.00    20.00     5.00
Diterima       2.00     0.00     2.00

                          Within Node Probabilities
Class      Top      Left      Right
Tidak Diterima  0.428    1.000    0.130
Diterima      0.572    0.000    0.870

Surrogate      Split      Assoc.      Improve.
1 KIMIA        s      600.566    0.950     0.017
2 FIGURAL      s      650.444    0.500     0.013

```

3	NUMERIK	s	689.546	0.450	0.008
4	MATDAS	r	640.129	0.300	0.006
5	B_INDONE	s	483.326	0.300	0.004

Competitor		Split	Improve.
1	KIMIA	630.525	0.019
2	FIGURAL	650.444	0.013
3	NUMERIK	693.328	0.011
4	B_INGGRI	509.467	0.010
5	FISIKA	341.682	0.008

```
*****
*                               *
*           Node 9: VERBAL      *
*                               *
*                               *
*           N: 70                *
*                               *
*****
```

```
*****
*                               *
*           Node 10              *
*                               *
*           N: 36                *
*                               *
*****
=====
```

Node 9 was split on VERBAL

A case goes left if VERBAL <= 612.259

Improvement = 0.005 Complexity Threshold = 0.009

Node	Cases	Wgt	Counts	Cost	Class
9	70		70.00	0.087	Diterima
10	36		36.00	0.199	Diterima
-11	34		34.00	0.040	Diterima

Weighted Counts

Class	Top	Left	Right
Tidak Diterima	43.00	29.00	14.00
Diterima	27.00	7.00	20.00

Within Node Probabilities

Class	Top	Left	Right
Tidak Diterima	0.087	0.199	0.040
Diterima	0.913	0.801	0.960

Surrogate		Split	Assoc.	Improve.
1	BIOLOGI	r	572.369	0.364 .285879E-04
2	MATIPA	r	822.482	0.135 .162996E-03
3	MATDAS	s	489.616	0.128 .143722E-04
4	KIMIA	s	491.762	0.128 .315264E-03
5	FISIKA	r	831.818	0.114 .244670E-03

Competitor		Split	Improve.
1	BIOLOGI	383.782	0.005
2	FISIKA	392.544	0.003
3	B_INDONE	440.096	0.003
4	MATDAS	477.189	0.003
5	B_INGGRI	575.273	0.002

```
*****
*                               *
*           Node 10: BIOLOGI    *
*                               *
*                               *
*           N: 36                *
*                               *
*****
```

```
=====
=           Terminal node 9      = =           Terminal node 10      =
=           N: 19                = =           N: 17                =
=====
```

Node 10 was split on BIOLOGI

A case goes left if BIOLOGI <= 573.940

Improvement = 0.026 Complexity Threshold = 0.018

Node	Cases	Wgt	Counts	Cost	Class
10	36		36.00	0.199	Diterima
-9	19		19.00	0.000	Tidak Diterima
-10	17		17.00	0.079	Diterima

Weighted Counts			
Class	Top	Left	Right
Tidak Diterima	29.00	19.00	10.00
Diterima	7.00	0.00	7.00
Within Node Probabilities			
Class	Top	Left	Right
Tidak Diterima	0.199	1.000	0.079
Diterima	0.801	0.000	0.921

Surrogate		Split	Assoc.	Improve.
1	VERBAL r	590.141	0.421	0.013
2	FISIKA s	447.276	0.263	0.009
3	KIMIA s	466.969	0.105	0.002
4	NUMERIK s	647.470	0.053	0.001
5	B_INDONE s	430.861	0.053	0.001

Competitor		Split	Improve.
1	VERBAL	583.926	0.016
2	FISIKA	447.276	0.009
3	MATDAS	740.698	0.006
4	NUMERIK	723.710	0.006
5	KIMIA	707.929	0.004

=====

TEST SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	Tidak Diteri	Diterima	
Tidak Diterima	226.00	29.00	255.00
Diterima	5.00	7.00	12.00
PRED. TOT.	231.00	36.00	267.00
TOT. CORRECT	0.873		
SENSITIVITY: 0.886 SPECIFICITY: 0.583 FALSE REFERENCE: 0.022 FALSE RESPONSE: 0.806 REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"			

=====

LEARNING SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	Tidak Diteri	Diterima	
Tidak Diterima	499.00	36.00	535.00
Diterima	0.00	32.00	32.00
PRED. TOT.	499.00	68.00	567.00
TOT. CORRECT	0.937		
SENSITIVITY: 0.933 SPECIFICITY: 1.000 FALSE REFERENCE: 0.000 FALSE RESPONSE: 0.529 REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"			

Lampiran 5. Output CART Hasil Analisis Data Prodi Statistika IPB

Output Variable Importance Pohon Klasifikasi Maksimal

Variabel	Skor	
FIGURAL	100	
BIOLOGI	34,64	
MATDAS	25,74	
KIMIA	12,94	
FISIKA	5,91	
NUMERIK	3,65	
MATIPA	0,62	

Output Relative Cost

Tree Number	Terminal Nodes	Test Set Relative cost	Resubstitution Relative cost	Complexity
1	12	0,430 ± 0,155	0,015	-1,000
2	9	0,430 ± 0,155	0,021	0,000879
3	8	0,430 ± 0,155	0,026	0,003
4	7	0,437 ± 0,155	0,036	0,005
5	6	0,445 ± 0,155	0,053	0,009
6**	5	0,378 ± 0,146	0,075	0,011
7	4	0,431 ± 0,146	0,113	0,019
8	3	0,486 ± 0,156	0,158	0,022
9	2	0,590 ± 0,157	0,281	0,062
10	1	1,000 ± 0,000	1,000	0,359

** Pohon klasifikasi Optimal

Output Informasi Terminal Node Pohon Klasifikasi Optimal

```

*****
*                               Node 1: FIGURAL                               *
*                               N: 598                                       *
*****
*****
*                               *                               *                               *
*                               Node 2                               *                               Node 3                               *
*                               N: 459                               *                               N: 139                               *
*****
Node 1 was split on FIGURAL
A case goes left if FIGURAL <= 634.938
Improvement = 0.264          Complexity Threshold = 0.359

Node   Cases   Wgt Counts      Cost Class
  1     598     598.00      0.500 Tidak Diterima
  2     459     459.00      0.078 Tidak Diterima
  3     139     139.00      0.187 Diterima

Weighted Counts
Class      Top      Left      Right
Tidak Diterima  583.00  458.00  125.00
Diterima       15.00   1.00   14.00

```

```

                                Within Node Probabilities
Class          Top      Left      Right
Tidak Diterima 0.500    0.922    0.187
Diterima       0.500    0.078    0.813

    Surrogate          Split          Assoc.      Improve.
1  KIMIA      s      542.584          0.447      0.202
2  NUMERIK   s      613.272          0.423      0.164
3  VERBAL    s      615.683          0.373      0.150
4  MATIPA    s      578.332          0.351      0.178
5  BIOLOGI   s      503.674          0.318      0.172

    Competitor        Split          Improve.
1  KIMIA          542.584          0.202
2  MATIPA         578.970          0.179
3  BIOLOGI        506.608          0.175
4  B_INGGRI       580.757          0.169
5  NUMERIK        613.272          0.164
*****
*                   Node 2: MATDAS          *
*                   N: 459                  *
*****
=====
=      Terminal node 1      =      Terminal node 2      =
=      N: 445                =      N: 14                  =
=====

Node 2 was split on MATDAS
A case goes left if MATDAS <= 721.666
Improvement = 0.045          Complexity Threshold = 0.022

Node      Cases  Wgt Counts      Cost Class
  2         459    459.00      0.078 Tidak Diterima
-1         445    445.00      0.000 Tidak Diterima
-2          14     14.00      0.251 Diterima

                                Weighted Counts
Class          Top      Left      Right
Tidak Diterima 458.00  445.00  13.00
Diterima        1.00    0.00    1.00

                                Within Node Probabilities
Class          Top      Left      Right
Tidak Diterima 0.922    1.000    0.251
Diterima       0.078    0.000    0.749

    Surrogate          Split          Assoc.      Improve.
1  VERBAL    s      690.373          0.229      0.033
2  MATIPA    s      811.200          0.019      .316724E-04

    Competitor        Split          Improve.
1  VERBAL          690.373          0.033
2  BIOLOGI         650.633          0.026
3  KIMIA           673.890          0.022
4  FISIKA          580.107          0.021
5  B_INGGRI        624.164          0.021
*****
*                   Node 3: BIOLOGI          *
*                   N: 139                  *
*****
=====
=      Terminal node 3      =      *      Node 4          *
=      N: 72                =      *      N: 67            *
=====
*****

Node 3 was split on BIOLOGI
A case goes left if BIOLOGI <= 506.065
Improvement = 0.092          Complexity Threshold = 0.062

```


Node	Cases	Wgt	Counts	Cost	Class
3	139		139.00	0.187	Diterima
-3	72		72.00	0.000	Tidak Diterima
4	67		67.00	0.089	Diterima

Weighted Counts			
Class	Top	Left	Right
Tidak Diterima	125.00	72.00	53.00
Diterima	14.00	0.00	14.00

Within Node Probabilities			
Class	Top	Left	Right
Tidak Diterima	0.187	1.000	0.089
Diterima	0.813	0.000	0.911

Surrogate		Split	Assoc.	Improve.
1	MATIPA s	480.518	0.292	0.042
2	KIMIA s	530.691	0.222	0.061
3	NUMERIK s	603.606	0.167	0.043
4	MATDAS r	711.012	0.111	0.016
5	FISIKA s	377.077	0.069	0.013

Competitor		Split	Improve.
1	KIMIA	541.940	0.071
2	NUMERIK	612.470	0.047
3	FISIKA	411.558	0.046
4	MATIPA	578.970	0.043
5	B_INGGRI	579.878	0.038


```

*****
*                               *
*           Node 4: KIMIA       *
*           N: 67               *
*                               *
*****
=====
=      Terminal node 4      =   =      Terminal node 5      =
=           N: 22           =   =           N: 45           =
=====

```

Node 4 was split on KIMIA
A case goes left if KIMIA <= 541.940
Improvement = 0.033 Complexity Threshold = 0.019

Node	Cases	Wgt	Counts	Cost	Class
4	67		67.00	0.089	Diterima
-4	22		22.00	0.000	Tidak Diterima
-5	45		45.00	0.054	Diterima

Weighted Counts			
Class	Top	Left	Right
Tidak Diterima	53.00	22.00	31.00
Diterima	14.00	0.00	14.00

Within Node Probabilities			
Class	Top	Left	Right
Tidak Diterima	0.089	1.000	0.054
Diterima	0.911	0.000	0.946

Surrogate		Split	Assoc.	Improve.
1	MATDAS s	430.174	0.227	0.016
2	NUMERIK s	605.146	0.136	0.019
3	FISIKA s	356.000	0.091	0.003
4	B_INGGRI s	353.513	0.091	0.003
5	MATIPA s	423.205	0.045	0.001

Competitor		Split	Improve.
1	FISIKA	411.558	0.025
2	NUMERIK	609.809	0.020
3	MATDAS	446.075	0.019
4	MATIPA	578.970	0.012
5	VERBAL	615.817	0.009

```

=====
TEST SAMPLE CLASSIFICATION TABLE
=====
Actual          Predicted Class          Actual
Class           Tidak Diteri           Diterima                Total
-----
Tidak Diterima    247.00                21.00                268.00
Diterima          3.00                  7.00                10.00
-----
PRED. TOT.        250.00                28.00                278.00
TOT. CORRECT      0.914
SENSITIVITY:      0.922                SPECIFICITY:           0.700
FALSE REFERENCE:  0.012                FALSE RESPONSE:        0.750
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"
-----

```

```

=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====
Actual          Predicted Class          Actual
Class           Tidak Diteri           Diterima                Total
-----
Tidak Diterima    539.00                44.00                583.00
Diterima          0.00                  15.00                15.00
-----
PRED. TOT.        539.00                59.00                598.00
TOT. CORRECT      0.926
SENSITIVITY:      0.925                SPECIFICITY:           1.000
FALSE REFERENCE:  0.000                FALSE RESPONSE:        0.746
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"
-----

```

Lampiran 6. Output CART Hasil Analisis Data Prodi Statistika UI

Output Variable Importance Pohon Klasifikasi Maksimal

Variabel	Skor	
Numerical	100	
Kimia	42,95	
B_Inggris	34,81	
Matipa	17,27	
Verbal	11,53	
Figural	10,94	
Matdas	8,51	
B_Indonesia	5,02	
Biologi	4,37	
Fisika	2,88	

Output Relative Cost

Tree Number	Terminal Nodes	Test Set Relative cost	Resubstitution Relative cost	Complexity
1	23	0,658 ± 0,127	0,027	-1,000
2	22	0,658 ± 0,127	0,029	0,001
3	15	0,547 ± 0,130	0,056	0,002
4	10	0,502 ± 0,128	0,077	0,002
5	9	0,511 ± 0,128	0,085	0,004
6	8	0,524 ± 0,128	0,096	0,005
7**	7	0,462 ± 0,124	0,108	0,006
8	4	0,507 ± 0,118	0,194	0,014
9	3	0,569 ± 0,125	0,278	0,042
10	2	0,644 ± 0,108	0,438	0,080
11	1	1,000 ± 0,000	1,000	0,281

Output Informasi Terminal Node Pohon Klasifikasi Optimal

```

*****
*           Node 1: NUMERICA           *
*           N: 511                     *
*****
=====
= Terminal node 1 = * Node 2 *
= N: 270          = * N: 241 *
=====
Node 1 was split on NUMERICA
A case goes left if NUMERICA <= 633.405
Improvement = 0.196      Complexity Threshold = 0.281

Node   Cases  Wgt  Counts      Cost  Class
  1     511    511.00  0.500 Tidak Diterima
 -1     270    270.00  0.000 Tidak Diterima
  2     241    241.00  0.304 Diterima

                          Weighted Counts
Class      Top      Left      Right
Tidak Diterima  480.00  270.00  210.00
Diterima       31.00   0.00    31.00

                          Within Node Probabilities
Class      Top      Left      Right
Tidak Diterima  0.500   1.000   0.304
Diterima       0.500   0.000   0.696

Surrogate      Split      Assoc.  Improve.
1 FIGURAL s    518.100  0.326   0.077
2 VERBAL s    530.578  0.248   0.102
3 MATDAS s    426.783  0.207   0.061
4 KIMIA s    486.232  0.178   0.083
5 MATIPA s    515.001  0.120   0.091

Competitor     Split      Improve.
1 KIMIA        587.248  0.166
2 MATDAS       551.510  0.152
3 FIGURAL      572.369  0.112
4 VERBAL       530.578  0.102
5 BIOLOGI      502.848  0.098

```

```

*****
*           Node 2: KIMIA           *
*           N: 241                   *
*****
*****
*           Node 3           *   *   *           Node 4           *
*           N: 126           *   *   *           N: 115           *
*****

```

Node 2 was split on KIMIA
A case goes left if KIMIA <= 586.563
Improvement = 0.083 Complexity Threshold = 0.080

Node	Cases	Wgt Counts	Cost	Class
2	241	241.00	0.304	Diterima
3	126	126.00	0.274	Tidak Diterima
4	115	115.00	0.167	Diterima

Class	Weighted Counts		
	Top	Left	Right
Tidak Diterima	210.00	123.00	87.00
Diterima	31.00	3.00	28.00

Class	Within Node Probabilities		
	Top	Left	Right
Tidak Diterima	0.304	0.726	0.167
Diterima	0.696	0.274	0.833

Surrogate	Split	Assoc.	Improve.
1 MATIPA s	517.741	0.189	0.037
2 VERBAL s	529.191	0.130	0.052
3 FISIKA r	605.228	0.080	.169873E-05
4 BIOLOGI s	391.939	0.071	0.027
5 B_INGGRI r	823.594	0.068	.290275E-04

Competitor	Split	Improve.
1 MATDAS	551.510	0.067
2 BIOLOGI	501.959	0.061
3 MATIPA	516.057	0.053
4 VERBAL	529.191	0.052
5 B_INGGRI	576.580	0.028

```

*****
*           Node 3: B_INGGRI           *
*           N: 126                       *
*****

```

```

=====
=           Terminal node 2           =   =   =           Terminal node 3           =
=           N: 117                     =   =   =           N: 9                       =
=====

```

Node 3 was split on B_INGGRI
A case goes left if B_INGGRI <= 756.421
Improvement = 0.059 Complexity Threshold = 0.042

Node	Cases	Wgt Counts	Cost	Class
3	126	126.00	0.274	Tidak Diterima
-2	117	117.00	0.000	Tidak Diterima
-3	9	9.00	0.114	Diterima

Class	Weighted Counts		
	Top	Left	Right
Tidak Diterima	123.00	117.00	6.00
Diterima	3.00	0.00	3.00

Class	Within Node Probabilities		
	Top	Left	Right
Tidak Diterima	0.726	1.000	0.114
Diterima	0.274	0.000	0.886

Surrogate	Split	Assoc.	Improve.
1 FISIKA s	641.828	0.533	0.036

2	NUMERICA	s	752.420		0.514	0.034
3	BIOLOGI	s	625.253		0.343	0.023
4	B_INDONE	s	685.248		0.200	0.009
5	MATIPA	s	734.752		0.181	0.010

Competitor			Split		Improve.
1	FISIKA		641.828		0.036
2	NUMERICA		752.420		0.034
3	BIOLOGI		625.253		0.023
4	MATDAS		551.510		0.019
5	VERBAL		594.405		0.015

 * Node 4: VERBAL *
 * N: 115 *

```

=====
= Terminal node 4 = * Node 5 *
= N: 13 = * N: 102 =
=====
    
```

Node 4 was split on VERBAL
 A case goes left if VERBAL <= 527.287
 Improvement = 0.019 Complexity Threshold = 0.014

Node	Cases	Wgt	Counts	Cost	Class
4	115		115.00	0.167	Diterima
-4	13		13.00	0.000	Tidak Diterima
5	102		102.00	0.146	Diterima

Weighted Counts					
Class	Top	Left	Right		
Tidak Diterima	87.00	13.00	74.00		
Diterima	28.00	0.00	28.00		
Within Node Probabilities					
Class	Top	Left	Right		
Tidak Diterima	0.167	1.000	0.146		
Diterima	0.833	0.000	0.854		

Surrogate			Split		Assoc.	Improve.
1	KIMIA	r	927.114		0.077	0.001
2	FISIKA	s	319.991		0.077	0.001

Competitor			Split		Improve.
1	MATDAS		551.527		0.018
2	BIOLOGI		501.959		0.017
3	B_INGGRI		397.237		0.015
4	MATIPA		466.721		0.013
5	B_INDONE		385.300		0.012

 * Node 5: MATDAS *
 * N: 102 *

```

=====
* Node 6 * = Terminal node 7 =
* N: 30 * = N: 72 =
=====
    
```

Node 5 was split on MATDAS
 A case goes left if MATDAS <= 528.426
 Improvement = 0.015 Complexity Threshold = 0.015

Node	Cases	Wgt	Counts	Cost	Class
5	102		102.00	0.146	Diterima
6	30		30.00	0.475	Diterima
-7	72		72.00	0.103	Diterima

Weighted Counts					
Class	Top	Left	Right		
Tidak Diterima	74.00	28.00	46.00		
Diterima	28.00	2.00	26.00		

		Within Node Probabilities				
Class		Top	Left	Right		
Tidak Diterima		0.146	0.475	0.103		
Diterima		0.854	0.525	0.897		
		Surrogate	Split	Assoc.	Improve.	
1	FISIKA s	388.671		0.610	.371039E-05	
2	B INGGRI r	818.661		0.161	.832707E-03	
3	BIOLOGI r	822.968		0.017	0.002	
4	MATIPA s	412.884		0.017	0.002	
5	B_INDONE s	296.722		0.017	0.002	
		Competitor	Split	Improve.		
1	BIOLOGI	501.959		0.013		
2	MATIPA	466.721		0.012		
3	B_INDONE	385.300		0.012		
4	B_INGGRI	821.292		0.009		
5	FIGURAL	508.095		0.008		

 * Node 6: MATIPA *
 * N: 30 *

```

=====
= Terminal node 5 = = Terminal node 6 = =
= N: 28 = = N: 2 = =
=====

```

Node 6 was split on MATIPA

A case goes left if MATIPA <= 662.611

Improvement = 0.031

Complexity Threshold = 0.029

Node	Cases	Wgt	Counts	Cost	Class
6	30		30.00	0.475	Diterima
-5	28		28.00	0.000	Tidak Diterima
-6	2		2.00	0.000	Diterima

Weighted Counts

Class	Top	Left	Right
Tidak Diterima	28.00	28.00	0.00
Diterima	2.00	0.00	2.00

Within Node Probabilities

Class	Top	Left	Right
Tidak Diterima	0.475	1.000	0.000
Diterima	0.525	0.000	1.000

		Surrogate	Split	Assoc.	Improve.	
1	B_INGGRI s	698.235		0.964	0.029	
2	FISIKA r	384.360		0.857	0.023	
3	B_INDONE s	561.686		0.607	0.014	
4	MATDAS r	467.527		0.536	0.012	
5	KIMIA s	672.699		0.536	0.012	

		Competitor	Split	Improve.		
1	B_INGGRI	698.235		0.029		
2	FISIKA	384.360		0.023		
3	B_INDONE	561.686		0.014		
4	KIMIA	672.699		0.012		
5	MATDAS	467.527		0.012		

=====

TEST SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	Tidak Diteri	Diterima	
Tidak Diterima	196.00	29.00	225.00

```

Diterima                5.00          10.00          15.00
-----
PRED. TOT.              201.00          39.00          240.00
TOT. CORRECT            0.858
  SENSITIVITY:          0.871      SPECIFICITY:          0.667
FALSE REFERENCE:        0.025      FALSE RESPONSE:      0.744
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"
-----

```

```

=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class      Actual
Class       Tidak Diteri         Diterima             Total
-----
Tidak Diterima  428.00          52.00          480.00
Diterima       0.00            31.00          31.00
-----
PRED. TOT.     428.00          83.00          511.00
TOT. CORRECT   0.898
  SENSITIVITY: 0.892      SPECIFICITY:          1.000
FALSE REFERENCE: 0.000      FALSE RESPONSE:      0.627
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"
-----

```

Lampiran 7. Output CART Hasil Analisis Data Prodi Matematika ITS

Output Variable Importance Pohon Klasifikasi Maksimal

Variable	Skor	
Verbal	100	
Kimia	99,49	
B_Indonesia	47,52	
B_Ingggris	33,07	
Figural	31,97	
Fisika	30,6	
Numerik	29,91	
MatDas	6,76	
Biologi	3,51	
MatIPA	0	

Output Relative Cost

Tree Number	Terminal Nodes	Test Set Relative cost	Resubstitution Relative cost	Complexity
1	25	0,843 ± 0,099	0,036	-1,000
2	23	0,862 ± 0,100	0,041	0,001
3	14	0,755 ± 0,113	0,092	0,003
4	13	0,768 ± 0,113	0,100	0,004
5	12	0,799 ± 0,114	0,110	0,005
6	11	0,799 ± 0,114	0,128	0,009
7	10	0,705 ± 0,116	0,149	0,010
8	9	0,736 ± 0,117	0,177	0,014

Tree Number	Terminal Nodes	Test Set <i>Relative cost</i>	Resubstitution <i>Relative cost</i>	Complexity
9	7	0,730 ± 0,117	0,234	0,014
10	6	0,673 ± 0,115	0,268	0,017
11	5	0,761 ± 0,117	0,307	0,020
12	4	0,535 ± 0,097	0,350	0,021
13**	3	0,480 ± 0,096	0,436	0,043
14	2	0,548 ± 0,089	0,528	0,046
15	1	1,000 ± 6,10E-005	1,000	0,236

** Pohon klasifikasi Optimal

Output Informasi Terminal Node Pohon Klasifikasi Optimal

```

*****
*                               Node 1: KIMIA                               *
*                               N: 422                                    *
*****
=====
=      Terminal node 1      =      *      Node 2      *
=      N: 250              =      *      N: 172      *
=====
Node 1 was split on KIMIA
A case goes left if KIMIA <= 495.332
Improvement = 0.117          Complexity Threshold = 0.236

Node   Cases  Wgt Counts      Cost Class
  1     422    422.00      0.500 Tidak Diterima
 -1     250    250.00      0.199 Tidak Diterima
  2     172    172.00      0.306 Diterima

Weighted Counts
Class      Top      Left      Right
Tidak Diterima 390.00  245.00  145.00
Diterima      32.00    5.00   27.00

Within Node Probabilities
Class      Top      Left      Right
Tidak Diterima 0.500    0.801    0.306
Diterima      0.500    0.199    0.694

Surrogate      Split      Assoc.      Improve.
  1 BIOLOGI   s      472.513      0.300      0.102
  2 FISIKA    s      484.754      0.297      0.091
  3 MATIPA    s      454.167      0.182      0.037
  4 B_INGGRI s      423.059      0.105      0.048
  5 B_INDONE s      470.831      0.092      0.062

Competitor      Split      Improve.
  1 FISIKA      483.152      0.108
  2 BIOLOGI     490.347      0.105
  3 MATDAS      643.227      0.077
  4 VERBAL      527.463      0.076
  5 NUMERIK     626.427      0.065

*****
*                               Node 2: VERBAL                               *
*                               N: 172                                    *
*****
=====
=      Terminal node 2      =      =      Terminal node 3      =
=      N: 49                =      =      N: 123                =
=====
Node 2 was split on VERBAL
A case goes left if VERBAL <= 526.073
Improvement = 0.043          Complexity Threshold = 0.046

```


Node	Cases	Wgt	Counts	Cost	Class
2	172		172.00	0.306	Diterima
-2	49		49.00	0.202	Tidak Diterima
-3	123		123.00	0.234	Diterima

Class	Weighted Counts		
	Top	Left	Right
Tidak Diterima	145.00	48.00	97.00
Diterima	27.00	1.00	26.00

Class	Within Node Probabilities		
	Top	Left	Right
Tidak Diterima	0.306	0.798	0.234
Diterima	0.694	0.202	0.766

Surrogate	Split	Assoc.	Improve.
1 B_INDONE r	708.635	0.219	.472188E-03
2 NUMERIK s	508.131	0.066	0.013
3 FIGURAL s	417.707	0.050	0.006
4 B_INGGRI s	368.495	0.033	0.005
5 MATIPA r	833.952	0.017	0.001

Competitor	Split	Improve.
1 B_INDONE	471.454	0.040
2 FISIKA	484.000	0.039
3 NUMERIK	538.135	0.029
4 BIOLOGI	472.199	0.027
5 _INGGRI	421.994	0.023

=====
TEST SAMPLE CLASSIFICATION TABLE
=====

Actual Class	Predicted Class		Actual Total
	Tidak Diteri	Diterima	
Tidak Diterima	116.00	45.00	161.00
Diterima	4.00	16.00	20.00
PRED. TOT.	120.00	61.00	181.00
TOT. CORRECT	0.729		
SENSITIVITY:	0.720	SPECIFICITY:	0.800
FALSE REFERENCE:	0.033	FALSE RESPONSE:	0.738
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"			

=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====

Actual Class	Predicted Class		Actual Total
	Tidak Diteri	Diterima	
Tidak Diterima	293.00	97.00	390.00
Diterima	6.00	26.00	32.00
PRED. TOT.	299.00	123.00	422.00
TOT. CORRECT	0.756		
SENSITIVITY:	0.751	SPECIFICITY:	0.813
FALSE REFERENCE:	0.020	FALSE RESPONSE:	0.789
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"			

Lampiran 8. Output CART Hasil Analisis Data Prodi Matematika IPB

Output Variable Importance Pohon Klasifikasi Maksimal

Variable	Skor
Kimia	100
Figural	48,94
Verbal	14,4
Numerik	11,92
Fisika	8,74
B_Indonesia	3,38
MatDas	2,07
MatIPA	1,3
B_Inggris	1,14
Biologi	0

Output Relative Cost

Tree Number	Terminal Nodes	Test Set <i>Relative cost</i>	Resubstitution <i>Relative cost</i>	Complexity
1	17	0,779 ± 0,172	0,024	-1,000
2	13	0,779 ± 0,172	0,035	0,001
3	12	0,779 ± 0,172	0,041	0,003
4	7	0,557 ± 0,189	0,091	0,005
5	6	0,557 ± 0,189	0,103	0,006
6	5	0,443 ± 0,173	0,126	0,012
7**	4	0,186 ± 0,033	0,162	0,018
8	3	0,314 ± 0,136	0,211	0,025
9	2	0,414 ± 0,042	0,432	0,110
10	1	1,000 ± 0,000	1,000	0,284

** Pohon klasifikasi Optimal

Output Informasi Terminal Node Pohon Klasifikasi Optimal

```

*****
*           Node 1: KIMIA           *
*           N: 358                   *
*****
=====
=   Terminal node 1   =   *           Node 2           *
=   N: 193            =   *           N: 165           *
=====
*****

Node 1 was split on KIMIA
A case goes left if KIMIA <= 523.260
Improvement = 0.198           Complexity Threshold = 0.284

Node   Cases   Wgt Counts           Cost Class
  1     358     358.00           0.500 Tidak Diterima
 -1     193     193.00           0.000 Tidak Diterima
  2     165     165.00           0.302 Diterima
    
```

Weighted Counts			
Class	Top	Left	Right
Tidak Diterima	340.00	193.00	147.00
Diterima	18.00	0.00	18.00

Within Node Probabilities			
Class	Top	Left	Right
Tidak Diterima	0.500	1.000	0.302
Diterima	0.500	0.000	0.698

Surrogate	Split	Assoc.	Improve.
1 VERBAL s	512.188	0.249	0.084
2 MATDAS s	442.206	0.197	0.086
3 NUMERIK s	599.420	0.181	0.140
4 FIGURAL s	529.584	0.176	0.101
5 MATIPA s	438.616	0.166	0.035

Competitor	Split	Improve.
1 FIGURAL	590.864	0.174
2 NUMERIK	605.485	0.154
3 MATIPA	574.481	0.124
4 VERBAL	609.256	0.106
5 MATDAS	547.168	0.104

```

*****
*                               *
*           Node 2: FIGURAL     *
*                               *
*           N: 165              *
*                               *
*****

```

```

*****
*                               *
*           Node 3               *
*                               *
*           N: 95                *
*                               *
*****
=====
*                               *
*           Node 3               *
*                               *
*           N: 95                *
*                               *
*                               *
*           = Terminal node 4   *
*                               *
*           N: 70                *
*                               *
*                               *
=====

```

Node 2 was split on FIGURAL
A case goes left if FIGURAL <= 590.864
Improvement = 0.122 Complexity Threshold = 0.110

Node	Cases	Wgt Counts	Cost	Class
2	165	165.00	0.302	Diterima
3	95	95.00	0.167	Tidak Diterima
-4	70	70.00	0.142	Diterima

Weighted Counts			
Class	Top	Left	Right
Tidak Diterima	147.00	94.00	53.00
Diterima	18.00	1.00	17.00

Within Node Probabilities			
Class	Top	Left	Right
Tidak Diterima	0.302	0.833	0.142
Diterima	0.698	0.167	0.858

Surrogate	Split	Assoc.	Improve.
1 NUMERIK s	599.637	0.372	0.094
2 MATDAS s	492.782	0.310	0.039
3 VERBAL s	521.066	0.151	0.047
4 FISIKA s	365.398	0.035	0.015
5 MATIPA s	401.331	0.027	0.007

Competitor	Split	Improve.
1 NUMERIK	605.048	0.104
2 MATIPA	574.481	0.055
3 MATDAS	443.235	0.049
4 VERBAL	521.066	0.047
5 FISIKA	482.277	0.042

```

*****
*           Node 3: KIMIA           *
*           N: 95                   *
*****

```

```

=====
=      Terminal node 2      =      Terminal node 3      =
=      N: 92                =      N: 3                =
=====

```

Node 3 was split on KIMIA
A case goes left if KIMIA <= 743.507
Improvement = 0.041 Complexity Threshold = 0.025

Node	Cases	Wgt	Counts	Cost	Class
3	95		95.00	0.167	Tidak Diterima
-2	92		92.00	0.000	Tidak Diterima
-3	3		3.00	0.096	Diterima

Class	Weighted Counts		
	Top	Left	Right
Tidak Diterima	94.00	92.00	2.00
Diterima	1.00	0.00	1.00

Class	Within Node Probabilities		
	Top	Left	Right
Tidak Diterima	0.833	1.000	0.096
Diterima	0.167	0.000	0.904

Surrogate	Split	Assoc.	Improve.
1 VERBAL s	688.232	0.521	0.030
2 B_INDONE s	609.462	0.186	0.020

Competitor	Split	Improve.
1 VERBAL	688.232	0.030
2 B_INDONE	609.462	0.020
3 FISIKA	540.930	0.018
4 B_INGCRI	597.805	0.016
5 BIOLOGI	571.628	0.015

=====

TEST SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	Tidak Diteri	Diterima	
Tidak Diterima	114.00	26.00	140.00
Diterima	0.00	7.00	7.00

PRED. TOT. 114.00 33.00 147.00

TOT. CORRECT 0.823

SENSITIVITY: 0.814 SPECIFICITY: 1.000

FALSE REFERENCE: 0.000 FALSE RESPONSE: 0.788

REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"

=====

LEARNING SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	Tidak Diteri	Diterima	
Tidak Diterima	285.00	55.00	340.00
Diterima	0.00	18.00	18.00

PRED. TOT. 285.00 73.00 358.00

TOT. CORRECT 0.846

SENSITIVITY: 0.838 SPECIFICITY: 1.000

FALSE REFERENCE: 0.000 FALSE RESPONSE: 0.753

REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"

Lampiran 9. Output CART Hasil Analisis Data Prodi Matematika UI

Output Variable Importance Pohon Klasifikasi Maksimal

Variabel	Skor	
Numerik	100	
Fisika	31,41	
Kimia	31,01	
Figural	20,91	
B_Indonesia	15,13	
Biologi	2,95	
B_Ingggris	2,67	
Verbal	2,2	
MatDas	2,13	
MatIPA	0,8	

Output Relative Cost

Tree Number	Terminal Nodes	Test Set <i>Relative cost</i>	Resubstitution <i>Relative cost</i>	Complexity
1	24	0,673 ± 0,121	0,023	-1,000
2	20	0,673 ± 0,121	0,030	0,000972
3	17	0,618 ± 0,122	0,048	0,003
4	12	0,591 ± 0,123	0,086	0,004
5	8	0,564 ± 0,122	0,124	0,005
6	7	0,580 ± 0,122	0,137	0,007
7**	5	0,460 ± 0,106	0,200	0,016
8	4	0,499 ± 0,113	0,247	0,024
9	3	0,462 ± 0,097	0,302	0,028
10	2	0,479 ± 0,084	0,409	0,053
11	1	1,000 ± 0,000	1,000	0,296

** Pohon klasifikasi Optimal

Output Informasi Terminal Node Pohon Klasifikasi Optimal

```

*****
*           Node 1: NUMERIK           *
*           N: 565                     *
*****
*****
*           Node 2           *           *           Node 3           *
*           N: 340           *           *           N: 225           *
*****
Node 1 was split on NUMERIK
A case goes left if NUMERIK <= 650.209
Improvement = 0.193           Complexity Threshold = 0.296

Node      Cases  Wgt Counts           Cost  Class

```

```

1      565      565.00      0.500 Tidak Diterima
2      340      340.00      0.074 Tidak Diterima
3      225      225.00      0.274 Diterima

```

		Weighted Counts		
Class		Top	Left	Right
Tidak Diterima		526.00	338.00	188.00
Diterima		39.00	2.00	37.00

		Within Node Probabilities		
Class		Top	Left	Right
Tidak Diterima		0.500	0.926	0.274
Diterima		0.500	0.074	0.726

```

Surrogate      Split      Assoc.      Improve.
1 FIGURAL  s      596.049      0.481      0.140
2 VERBAL   s      557.368      0.286      0.107
3 KIMIA    s      568.959      0.237      0.142
4 MATDAS   s      496.088      0.195      0.065
5 MATIPA   s      518.030      0.129      0.039

```

```

Competitor      Split      Improve.
1 MATDAS        612.410      0.148
2 KIMIA         570.144      0.145
3 FIGURAL       596.412      0.141
4 VERBAL        559.432      0.112
5 BIOLOGI       578.560      0.106

```

```

*****
*           Node 2: KIMIA           *
*           N: 340                   *
*****

```

```

=====
=      Terminal node 1      =      =      Terminal node 2      =
=      N: 336               =      =      N: 4                   =
=====

```

Node 2 was split on KIMIA
A case goes left if KIMIA <= 835.539
Improvement = 0.044 Complexity Threshold = 0.024

Node	Cases	Wgt	Counts	Cost	Class
2	340		340.00	0.074	Tidak Diterima
-1	336		336.00	0.000	Tidak Diterima
-2	4		4.00	0.069	Diterima

		Weighted Counts		
Class		Top	Left	Right
Tidak Diterima		338.00	336.00	2.00
Diterima		2.00	0.00	2.00

		Within Node Probabilities		
Class		Top	Left	Right
Tidak Diterima		0.926	1.000	0.069
Diterima		0.074	0.000	0.931

```

Surrogate      Split      Assoc.      Improve.
1 BIOLOGI  s      722.488      0.620      0.033
2 B_INDONE s      710.042      0.362      0.016

```

```

Competitor      Split      Improve.
1 BIOLOGI       722.488      0.033
2 MATIPA        700.883      0.018
3 B_INDONE      710.042      0.016
4 NUMERIK       627.898      0.014
5 MATDAS        617.253      0.014

```

```
*****
*                               *
*           Node 3: FISIKA      *
*           N: 225              *
*                               *
*****
```

```
=====
= Terminal node 3 = * Node 4 *
= N: 56 = * N: 169 *
=====
```

Node 3 was split on FISIKA
 A case goes left if FISIKA <= 415.590
 Improvement = 0.061 Complexity Threshold = 0.053

Node	Cases	Wgt Counts	Cost	Class
3	225	225.00	0.274	Diterima
-3	56	56.00	0.000	Tidak Diterima
4	169	169.00	0.209	Diterima

Class	Weighted Counts		
	Top	Left	Right
Tidak Diterima	188.00	56.00	132.00
Diterima	37.00	0.00	37.00

Class	Within Node Probabilities		
	Top	Left	Right
Tidak Diterima	0.274	1.000	0.209
Diterima	0.726	0.000	0.791

Surrogate	Split	Assoc.	Improve.
1 MATDAS s	391.864	0.071	0.008
2 KIMIA s	390.076	0.054	0.009
3 FIGURAL s	337.682	0.018	0.001
4 BIOLOGI s	302.394	0.018	0.003
5 B_INDONE r	698.943	0.018	0.007

Competitor	Split	Improve.
1 KIMIA	569.863	0.053
2 BIOLOGI	499.685	0.042
3 FIGURAL	561.082	0.038
4 MATDAS	612.257	0.035
5 B_INGGRI	532.736	0.034

```
*****
*                               *
*           Node 4: FIGURAL    *
*           N: 169              *
*                               *
*****
```

```
=====
= Terminal node 4 = = Terminal node 5 =
= N: 29 = = N: 140 =
=====
```

Node 4 was split on FIGURAL
 A case goes left if FIGURAL <= 561.082
 Improvement = 0.036 Complexity Threshold = 0.028

Node	Cases	Wgt Counts	Cost	Class
4	169	169.00	0.209	Diterima
-4	29	29.00	0.000	Tidak Diterima
-5	140	140.00	0.171	Diterima

Class	Weighted Counts		
	Top	Left	Right
Tidak Diterima	132.00	29.00	103.00
Diterima	37.00	0.00	37.00

Class	Within Node Probabilities		
	Top	Left	Right
Tidak Diterima	0.209	1.000	0.171
Diterima	0.791	0.000	0.829

Surrogate	Split	Assoc.	Improve.
1 KIMIA s	404.859	0.069	0.007
2 BIOLOGI r	811.849	0.034	0.001

3	MATIPA	r	793.448		0.034	0.004
4	B_INGGRI	r	922.577		0.034	0.001
	Competitor		Split			Improve.
1	KIMIA		569.863			0.035
2	BIOLOGI		495.503			0.023
3	MATDAS		671.522			0.018
4	B_INGGRI		532.503			0.016
5	VERBAL		621.380			0.015

=====

TEST SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	Tidak Diteri	Diterima	
Tidak Diterima	193.00	56.00	249.00
Diterima	4.00	13.00	17.00

PRED. TOT.	197.00	69.00	266.00
TOT. CORRECT	0.774		
SENSITIVITY:	0.775	SPECIFICITY:	0.765
FALSE REFERENCE:	0.020	FALSE RESPONSE:	0.812
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"			

=====

LEARNING SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	Tidak Diteri	Diterima	
Tidak Diterima	421.00	105.00	526.00
Diterima	0.00	39.00	39.00

PRED. TOT.	421.00	144.00	565.00
TOT. CORRECT	0.814		
SENSITIVITY:	0.800	SPECIFICITY:	1.000
FALSE REFERENCE:	0.000	FALSE RESPONSE:	0.729
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"			

Lampiran 10. Output CART Hasil Analisis Data Prodi Aktuaria ITS

Output Variable Importance Pohon Klasifikasi Maksimal

Variabel	Skor
Fisika	100
Matipa	42,13
Biologi	23,8
Numerik	21,6
B_Indonesia	14,15
Verbal	3,42
B_Ingggris	2,9
Matdas	1,78
Kimia	1,76
Figural	0

Output Relative Cost

Tree Number	Terminal Nodes	Test Set Relative cost	Resubstitution Relative cost	Complexity
1	15	0,922 ± 0,105	0,021	-1,000
2	11	0,829 ± 0,139	0,028	0,0009
3	10	0,832 ± 0,139	0,031	0,001
4	8	0,835 ± 0,139	0,039	0,002
5	7	0,835 ± 0,139	0,046	0,004
6	6	0,850 ± 0,139	0,055	0,004
7	5	0,760 ± 0,158	0,092	0,018
8	4	0,748 ± 0,158	0,148	0,028
9	3	0,742 ± 0,158	0,211	0,031
10**	2	0,626 ± 0,159	0,350	0,069
11	1	1,000 ± 0,000	1,000	0,325

** Pohon klasifikasi Optimal

Output Informasi Terminal Node Pohon Klasifikasi Optimal

```

*****
*                               *
*           Node 1: FISIKA      *
*           N: 725              *
*                               *
*****
=====
=           Terminal node 1     =           Terminal node 2     =
=           N: 510              =           N: 215              =
=====
Node 1 was split on FISIKA
A case goes left if FISIKA <= 552.345
Improvement = 0.222           Complexity Threshold = 0.325

Node   Cases  Wgt  Counts           Cost  Class
  1     725    725.00    0.500  Tidak Diterima
 -1     510    510.00    0.085  Tidak Diterima
 -2     215    215.00    0.233  Diterima

                          Weighted Counts
Class           Top           Left           Right
Tidak Diterima  710.00    509.00    201.00
Diterima        15.00     1.00     14.00

                          Within Node Probabilities
Class           Top           Left           Right
Tidak Diterima  0.500     0.915     0.233
Diterima        0.500     0.085     0.767

Surrogate           Split           Assoc.           Improve.
1 MATIPA  s         544.537           0.462           0.190
2 B_INGGRI s         587.550           0.411           0.141
3 MATDAS  s         583.075           0.404           0.115
4 KIMIA   s         533.266           0.373           0.212
5 BIOLOGI s         495.265           0.328           0.135

Competitor          Split           Improve.
1 KIMIA             534.281           0.213
2 MATIPA            558.432           0.205
3 NUMERICA          626.260           0.147
4 B_INGGRI          589.538           0.143
5 BIOLOGI           495.265           0.135

```

```

=====
TEST SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class      Actual
Class       Tidak Diteri        Diterima             Total
-----
Tidak Diterima  235.00             97.00                332.00
Diterima        3.00                6.00                 9.00
-----
PRED. TOT.      238.00             103.00               341.00
TOT. CORRECT    0.707
SENSITIVITY:    0.708             SPECIFICITY:         0.667
FALSE REFERENCE: 0.013           FALSE RESPONSE:     0.942
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"
-----

```

```

=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class      Actual
Class       Tidak Diteri        Diterima             Total
-----
Tidak Diterima  509.00             201.00               710.00
Diterima        1.00                14.00                15.00
-----
PRED. TOT.      510.00             215.00               725.00
TOT. CORRECT    0.721
SENSITIVITY:    0.717             SPECIFICITY:         0.933
FALSE REFERENCE: 0.002           FALSE RESPONSE:     0.935
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"
-----

```

Lampiran 11. Output CART Hasil Analisis Data Prodi Aktuaria IPB

Output Variable Importance Pohon Klasifikasi Maksimal

Variabel	Skor	
Figural	100	
Fisika	48,7	
Verbal	35,41	
Kimia	14,34	
MatIPA	6,31	
Numerik	4,43	
B_Indonesia	2,29	
MatDas	0,94	
B_Inggris	0,47	
Biologi	0,46	


```

*****
*           Node 2: VERBAL           *
*           N: 661                   *
*****
=====
=      Terminal node 1      =      *           Node 3           *
=      N: 623              =      *           N: 38            *
=====

```

Node 2 was split on VERBAL
A case goes left if VERBAL <= 690.476
Improvement = 0.039 Complexity Threshold = 0.018

Node	Cases	Wgt	Counts	Cost	Class
2	661		661.00	0.088	Tidak Diterima
-1	623		623.00	0.000	Tidak Diterima
3	38		38.00	0.367	Diterima

Weighted Counts

Class	Top	Left	Right
Tidak Diterima	660.00	623.00	37.00
Diterima	1.00	0.00	1.00

Within Node Probabilities

Class	Top	Left	Right
Tidak Diterima	0.912	1.000	0.367
Diterima	0.088	0.000	0.633

Surrogate	Split	Assoc.	Improve.
1 B_INGGRI s	679.621	0.246	0.036
2 NUMERICA s	703.539	0.167	0.032
3 MATIPA s	685.585	0.147	0.031

Competitor	Split	Improve.
1 B_INGGRI	679.621	0.036
2 NUMERICA	703.539	0.032
3 MATIPA	685.585	0.031
4 FISIKA	588.154	0.025
5 B_INDO	617.332	0.020

```

*****
*           Node 3: VERBAL           *
*           N: 38                   *
*****
=====
=      Terminal node 2      =      =      Terminal node 3      =
=      N: 1                =      =      N: 37                =
=====

```

Node 3 was split on VERBAL
A case goes left if VERBAL <= 690.634
Improvement = 0.026 Complexity Threshold = 0.021

Node	Cases	Wgt	Counts	Cost	Class
3	38		38.00	0.367	Diterima
-2	1		1.00	0.000	Diterima
-3	37		37.00	0.000	Tidak Diterima

Weighted Counts

Class	Top	Left	Right
Tidak Diterima	37.00	0.00	37.00
Diterima	1.00	1.00	0.00

Within Node Probabilities

Class	Top	Left	Right
Tidak Diterima	0.367	0.000	1.000
Diterima	0.633	1.000	0.000

Surrogate	Split	Assoc.	Improve.
1 B_INGGRI r	670.514	0.919	0.023
2 MATIPA r	680.437	0.892	0.022
3 NUMERICA r	702.830	0.892	0.022

```

4 FISIKA r 579.280 0.865 0.021
5 B_INDO r 615.856 0.811 0.019

```

```

Competitor Split Improve.
1 B_INGGRI 670.514 0.023
2 NUMERICA 702.830 0.022
3 MATIPA 680.437 0.022
4 FISIKA 579.280 0.021
5 B_INDO 615.856 0.019

```

```

*****
* Node 4: FISIKA *
* N: 247 *
*****

```

```

=====
= Terminal node 4 = * Node 5 *
= N: 143 = * N: 104 *
=====

```

Node 4 was split on FISIKA
A case goes left if FISIKA <= 477.657
Improvement = 0.112 Complexity Threshold = 0.080

Node	Cases	Wgt	Counts	Cost	Class
4	247		247.00	0.220	Diterima
-4	143		143.00	0.000	Tidak Diterima
5	104		104.00	0.099	Diterima

Weighted Counts

Class	Top	Left	Right
Tidak Diterima	234.00	143.00	91.00
Diterima	13.00	0.00	13.00

Within Node Probabilities

Class	Top	Left	Right
Tidak Diterima	0.220	1.000	0.099
Diterima	0.780	0.000	0.901

```

Surrogate Split Assoc. Improve.
1 MATDAS s 472.077 0.357 0.058
2 KIMIA s 539.641 0.329 0.086
3 NUMERICA s 626.571 0.217 0.058
4 BIOLOGI s 397.733 0.203 0.031
5 MATIPA s 468.358 0.175 0.046

```

```

Competitor Split Improve.
1 KIMIA 539.641 0.086
2 MATIPA 556.522 0.065
3 MATDAS 481.093 0.064
4 NUMERICA 627.027 0.059
5 VERBAL 564.376 0.049

```

```

*****
* Node 5: KIMIA *
* N: 104 *
*****

```

```

=====
= Terminal node 5 = * Node 6 *
= N: 33 = * N: 71 *
=====

```

Node 5 was split on KIMIA
A case goes left if KIMIA <= 534.476
Improvement = 0.031 Complexity Threshold = 0.018

Node	Cases	Wgt	Counts	Cost	Class
5	104		104.00	0.099	Diterima
-5	33		33.00	0.000	Tidak Diterima
6	71		71.00	0.065	Diterima

Weighted Counts			
Class	Top	Left	Right
Tidak Diterima	91.00	33.00	58.00
Diterima	13.00	0.00	13.00
Within Node Probabilities			
Class	Top	Left	Right
Tidak Diterima	0.099	1.000	0.065
Diterima	0.901	0.000	0.935

Surrogate	Split	Assoc.	Improve.
1 MATIPA s	460.776	0.212	0.016
2 MATDAS s	479.382	0.182	0.017
3 BIOLOGI s	377.327	0.091	0.005
4 B_INGGRI s	338.098	0.091	0.003
5 NUMERICA s	541.716	0.061	0.002

Competitor	Split	Improve.
1 MATIPA	493.266	0.029
2 VERBAL	563.475	0.025
3 NUMERICA	627.027	0.023
4 MATDAS	479.382	0.017
5 B_INDO	409.888	0.011

```

*****
*           Node 6: VERBAL           *
*           N: 71                    *
*****

```

```

=====
=      Terminal node 6      =      Terminal node 7      =
=      N: 17                =      N: 54                =
=====

```

Node 6 was split on VERBAL
 A case goes left if VERBAL <= 563.475
 Improvement = 0.017 Complexity Threshold = 0.010

Node	Cases	Wgt Counts	Cost	Class
6	71	71.00	0.065	Diterima
-6	17	17.00	0.000	Tidak Diterima
-7	54	54.00	0.047	Diterima

Weighted Counts			
Class	Top	Left	Right
Tidak Diterima	58.00	17.00	41.00
Diterima	13.00	0.00	13.00

Within Node Probabilities			
Class	Top	Left	Right
Tidak Diterima	0.065	1.000	0.047
Diterima	0.935	0.000	0.953

Surrogate	Split	Assoc.	Improve.
1 NUMERICA s	577.333	0.118	0.004
2 BIOLOGI s	381.259	0.118	0.002
3 MATDAS s	423.329	0.118	0.002
4 MATIPA r	792.018	0.059	.978358E-03
5 B_INGGRI s	353.953	0.059	.978358E-03

Competitor	Split	Improve.
1 MATIPA	491.335	0.015
2 NUMERICA	627.027	0.014
3 B_INDO	409.888	0.008
4 MATDAS	471.336	0.006
5 BIOLOGI	631.060	0.004

=====

TEST SAMPLE CLASSIFICATION TABLE

=====

Actual Class	Predicted Class		Actual Total
	Tidak Diteri	Diterima	

```

-----
Tidak Diterima      395.00      20.00      415.00
Diterima            2.00        4.00        6.00
-----
PRED. TOT.         397.00      24.00      421.00
TOT. CORRECT       0.948
  SENSITIVITY:      0.952      SPECIFICITY:      0.667
FALSE REFERENCE:   0.005     FALSE RESPONSE:   0.833
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"
-----

```

```

=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====

```

```

Actual          Predicted Class          Actual
Class           Tidak Diteri    Diterima          Total
-----
Tidak Diterima      853.00      41.00      894.00
Diterima            0.00       14.00       14.00
-----
PRED. TOT.         853.00      55.00      908.00
TOT. CORRECT       0.955
  SENSITIVITY:      0.954      SPECIFICITY:      1.000
FALSE REFERENCE:   0.000     FALSE RESPONSE:   0.745
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"
-----

```

Lampiran 12. Output CART Hasil Analisis Data Prodi Aktuaria UI

Output Variable Importance Pohon Klasifikasi Maksimal

Variabel	Skor
MatDas	100
MatIPA	12,41
B_Ingggris	10,46
Biologi	5,07
B_Indonesia	2,13
Kimia	0,58
Numerik	0,39
Verbal	0
Figural	0
Fisika	0

Output Relative Cost

Tree Number	Terminal Nodes	Test Set <i>Relative cost</i>	Resubstitution <i>Relative cost</i>	Complexity
1	9	0,515 ± 0,204	0,009	-1,000
2	7	0,515 ± 0,204	0,012	0,000765
3	6	0,517 ± 0,204	0,014	0,001
4	5	0,361 ± 0,193	0,021	0,003
5	4	0,376 ± 0,193	0,041	0,010
6	3	0,370 ± 0,193	0,082	0,021
7**	2	0,265 ± 0,153	0,135	0,026
8	1	1,000 ± 0,000	1,000	0,433

** Pohon klasifikasi Optimal

Output Informasi Terminal Node Pohon Klasifikasi Optimal

```

*****
*                               *
*           Node 1: MATDAS      *
*                               *
*           N: 1366             *
*                               *
*****
=====
=      Terminal node 1      =      Terminal node 2      =
=      N: 1225              =      N: 141                =
=====
Node 1 was split on MATDAS
A case goes left if MATDAS <= 693.392
Improvement = 0.375          Complexity Threshold = 0.433

Node      Cases  Wgt Counts      Cost Class
1         1366   1366.00      0.500 Tidak Diterima
-1        1225   1225.00      0.048 Tidak Diterima
-2         141   141.00       0.086 Diterima

                          Weighted Counts
Class      Top      Left      Right
Tidak Diterima  1344.00  1224.00  120.00
Diterima       22.00    1.00    21.00

                          Within Node Probabilities
Class      Top      Left      Right
Tidak Diterima  0.500   0.952   0.086
Diterima       0.500   0.048   0.914

Surrogate      Split      Assoc.      Improve.
1 MATIPA      s      623.915    0.681      0.304
2 KIMIA       s      639.455    0.670      0.267
3 BIOLOGI     s      609.328    0.644      0.285
4 FISIKA      s      524.161    0.595      0.156
5 B_INDO     s      572.982    0.553      0.215

Competitor     Split      Improve.
1 MATIPA      623.915    0.304
2 BIOLOGI     609.328    0.285
3 KIMIA       607.665    0.271
4 B_INDO     572.982    0.215
5 FIGURAL    635.447    0.198

```



```

=====
TEST SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class      Actual
Class      Tidak Diteri      Diterima      Total
-----
Tidak Diterima      589.00      64.00      653.00
Diterima            1.00      5.00      6.00
-----
PRED. TOT.          590.00      69.00      659.00
TOT. CORRECT        0.901
  SENSITIVITY:      0.902      SPECIFICITY:      0.833
FALSE REFERENCE:    0.002      FALSE RESPONSE:    0.928
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"
-----

```

```

=====
LEARNING SAMPLE CLASSIFICATION TABLE
=====
Actual      Predicted Class      Actual
Class      Tidak Diteri      Diterima      Total
-----
Tidak Diterima      1224.00      120.00      1344.00
Diterima            1.00      21.00      22.00
-----
PRED. TOT.          1225.00      141.00      1366.00
TOT. CORRECT        0.911
  SENSITIVITY:      0.911      SPECIFICITY:      0.955
FALSE REFERENCE:    .816327E-03      FALSE RESPONSE:    0.851
REFERENCE = "Tidak Diterima", RESPONSE = "Diterima"
-----

```

(Halaman ini sengaja dikosongkan)

Lampiran 13. Surat Keterangan Pengambilan Data**SURAT PERNYATAAN**

Saya yang bertanda tangan dibawah ini, mahasiswa Departemen Statistika FMKSD ITS:

Nama : Rhavida Anniza Andyani

NRP : 06211540000006

menyatakan bahwa data yang digunakan dalam Tugas Akhir/Thesis ini merupakan data sekunder yang diambil dari penelitian / buku / Tugas Akhir / Thesis/ publikasi lainnya yaitu:

Sumber : Kajibag SNMPTN-SBMPTN 2018

Keterangan : Data Peserta SBMPTN Tahun 2018 yang Diterima dan Tidak Diterima

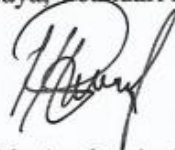
Surat Pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Mengetahui,
Pembimbing Tugas Akhir



Dr. Dra. Ismaini Zain, M.Si
NIP. 19600525 198803 2 001

Surabaya, 7 Januari 2019



Rhavida Anniza Andyani
NRP. 06211540000006

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS



Rhavida Anniza Andyani merupakan anak ketiga dari pasangan R. Agus Julianto dan Sri Rahayu. Penulis dilahirkan pada tanggal 12 Agustus 1996 di Kota Madiun. Sebelum memasuki masa perkuliahan di Statistika ITS tahun 2015, penulis telah menamatkan pendidikan formal di MI Islamiyah 01 Madiun, SMP Negeri 1 Madiun dan SMA Negeri 3 Madiun. Pada tahun kedua dan ketiga masa perkuliahan, penulis aktif di Himpunan Mahasiswa Statistika ITS (HIMASTA-ITS) sebagai staff departemen Kewirausahaan periode 2016/2017 dan Sekretaris departemen Kewirausahaan periode 2017/2018. Dalam bidang akademik, penulis berkesempatan mengikuti beberapa lomba sampai dengan tahapan semifinalis pada *Statistics National Challenge* (NSC) tahun 2017 yang diselenggarakan oleh Statistika Universitas Brawijaya, serta pada *Statistics Team Competition* (STC) yang diselenggarakan prodi Matematika Universitas Indonesia tahun 2018. Selain itu, penulis juga aktif mengikuti kepanitiaan di beberapa kegiatan. Apabila pembaca ingin melakukan diskusi mengenai tugas akhir ini atau memberikan segala kritik dan saran untuk penulis, dapat disampaikan melalui email rhavida.96@gmail.com.