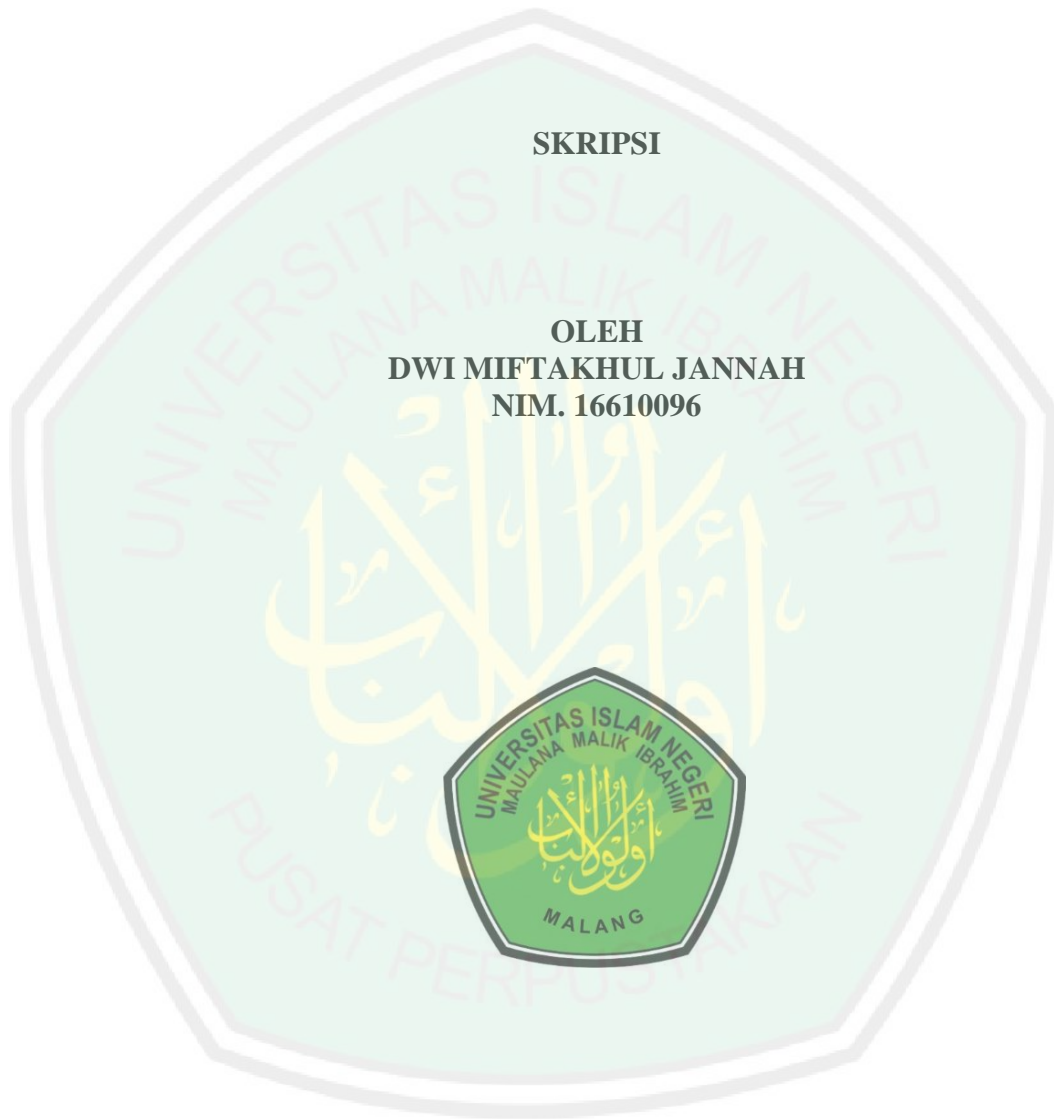


**KLASIFIKASI PENERIMAAN PESERTA DIDIK BARU DI MADRASAH  
BERTARAF INTERNATIONAL AMANATUL UMMAH PACET  
MENGUNAKAN *BOOTSTRAP AGGREGATING  
CLASSIFICATION AND REGRESSION TREE***

**SKRIPSI**

**OLEH  
DWI MIFTAKHUL JANNAH  
NIM. 16610096**



**JURUSAN MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2020**

**KLASIFIKASI PENERIMAAN PESERTA DIDIK BARU DI MADRASAH  
BERTARAF INTERNATIONAL AMANATUL UMMAH PACET  
MENGUNAKAN *BOOTSTRAP AGGREGATING  
CLASSIFICATION AND REGRESSION TREE***

**SKRIPSI**

**OLEH  
DWI MIFTAKHUL JANNAH  
NIM. 16610096**



**JURUSAN MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2020**

**KLASIFIKASI PENERIMAAN PESERTA DIDIK BARU DI MADRASAH  
BERTARAF INTERNATIONAL AMANATUL UMMAH PACET  
MENGUNAKAN *BOOTSTRAP AGGREGATING  
CLASSIFICATION AND REGRESSION TREE***

**SKRIPSI**

**Diajukan Kepada  
Fakultas Sains dan Teknologi  
Universitas Islam Negeri Maulana Malik Ibrahim Malang  
untuk Memenuhi Salah Satu Persyaratan dalam  
Memperoleh Gelar Sarjana Matematika (S.Mat)**

**Oleh:  
Dwi Miftakhul Jannah  
NIM 16610096**

**JURUSAN MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2020**

**KLASIFIKASI PENERIMAAN PESERTA DIDIK BARU DI MADRASAH  
BERTARAF INTERNATIONAL AMANATUL UMMAH PACET  
MENGUNAKAN *BOOTSTRAP AGGREGATING  
CLASSIFICATION AND REGRESSION TREE***

**SKRIPSI**

Oleh:  
**Dwi Miftakhul Jannah**  
NIM 16610096

Telah Diperiksa dan Disetujui untuk Diuji

Tanggal, 1 Juli 2020

Pembimbing I,



Heni Widayani, M.Si  
NIP. 19901006 20180201 2 229

Pembimbing II,



Angga Dwi Mulyanto, M.Si  
NIP. 19890813 201903 1 0 12

Mengetahui,  
Ketua Jurusan Matematika



Usman Pagalay, M.Si  
NIP. 19650414 200312 1 001

**KLASIFIKASI PENERIMAAN PESERTA DIDIK BARU DI MADRASAH  
BERTARAF INTERNATIONAL AMANATUL UMMAH PACET  
MENGUNAKAN BOOTSTRAP AGGREGATING  
CLASSIFICATION AND REGRESSION TREE**

**SKRIPSI**

**Oleh:  
Dwi Miftakhul Jannah  
NIM 16610096**

Telah Dipertahankan di Depan Dewan Penguji Skripsi  
dan Dinyatakan Diterima sebagai Salah Satu Persyaratan  
untuk Memperoleh Gelar Sarjana Matematika (S.Mat)

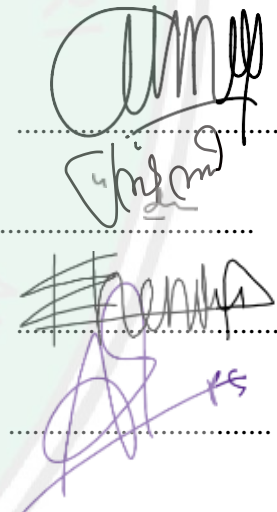
Tanggal, 1 Juli 2020

Penguji Utama : Mohammad Jamhuri, M.Si

Ketua Penguji : Ria Dhea Layla Nur Karisma, M.Si

Sekretaris Penguji : Heni Widayani, M.Si

Anggota penguji : Angga Dwi Mulyanto, M.Si



Mengetahui,

Ketua Jurusan Matematika



Usman Pagalay, M.Si  
NIP. 19650414 200312 1 001

## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Dwi Miftakhul Jannah

NIM : 16610096

Jurusan : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Klasifikasi penerimaan Peserta Didik Baru di Madrasah Bertaraf International Amanatul Ummah Pacet Menggunakan *Bootstrap Aggregating Classification and Regression Tree*.

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan pengambilan data, tulisan atau pikiran orang lain yang saya akui sebagai hasil tulisan atau pikiran saya sendiri, kecuali dengan mencantumkan sumber cuplikan pada daftar rujukan. Apabila di kemudian hari terbukti atau dapat dibuktikan skripsi ini hasil jiplakan, maka saya bersedia menerima sanksi atas perbuatan Tersebut.

Malang, 1 Juli 2020  
Yang membuat pernyataan



Dwi Miftakhul Jannah  
NIM. 16610096

## MOTO

“Hidup ini seperti sepeda, agar tetap seimbang, kau harus terus bergerak”  
(Albert Einstein)

“Jika kamu tidak membangun mimpimu, seseorang akan memperkerjakanmu  
untuk membangun mimpinya.”  
(Tony Ganskins)



## **PERSEMBAHAN**

**Bismillahirrohmanirrohim**

Skripsi ini dipersembahkan untuk:

Orang tua penulis Bapak Samsul dan Ibu Siti Romlah, kakak penulis

Amilusholichah yang selalu mendoakan penulis disetiap doa yang dipanjatkan,  
yang selalu memberikan semangat serta kasih sayang, selalu mendengarkan keluh  
kesah penulis dan menghibur penulis dalam pengerjaan skripsi ini.



## KATA PENGANTAR

*Assalamu'alaikum Warahmatullahi Wabarakatuh*

Segala puji bagi ALLAH SWT atas rahmat, taufiq, hidayah dan karunia-Nya yang selalu diberikan tanpa henti. Sehingga penulis dapat menyelesaikan penyusunan skripsi dengan judul “Klasifikasi Penerimaan Peserta Didik Baru di Madrasah Amanatul Ummah Pacet Menggunakan *Bootstrap Aggregating Classification and Regression Tree*” sebagai salah satu syarat untuk mendapatkan gelar sarjana dalam bidang matematika di Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Selama penyusunan skripsi ini, penulis banyak mendapatkan arahan dan semangat dari berbagai pihak. Oleh karena itu, penulis ucapkan terimakasih sebesar-besarnya kepada:

1. Prof. Dr. H. Abd. Haris, M.Ag, selaku rektor Universitas Islam Negeri Maulana Malik Ibrahim Malang.
2. Dr. Sri Harini, M.Si, selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
3. Dr. Usman Pagalay, M.Si, selaku ketua Jurusan Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang.
4. Heni Widayani, M.Si, selaku dosen pembimbing I yang telah memberikan arahan, motivasi, nasihat dan pengalaman yang berharga kepada penulis.
5. Ratnaning Palupi, M.Si dan Angga Dwi Mulyanto, M.Si, selaku dosen pembimbing II yang telah memberikan arahan, nasihat dan berbagai ilmu kepada penulis.

6. Segenap sivitas akademika Jurusan Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang terutama seluruh dosen, terima kasih atas ilmu dan bimbingannya.
7. Ayah dan ibu serta kakak tercinta yang telah memberikan doa, nasihat, semangat serta motivasi kepada penulis.
8. Seluruh teman-teman di Jurusan Matematika angkatan 2016, terutama Zahra Firdausa Khurin'in, Ramadhani Auliyana Rosa, Khilmi Hani, Ayunina Faidatul dan Emalia Nailun Makhfudho yang telah berjuang bersama dalam menyelesaikan skripsi ini.
9. Seluruh teman-teman NORTAVIUZ, terutama Conia Prajna Kathrine dan Sofie Nurmila atas semangat, motivasi dan hiburan dikala sedih.
10. Semua pihak yang secara langsung atau tidak langsung telah ikut membantu dalam penyelesaian skripsi ini.

Akhirnya penulis berharap dengan rahmat dan izin-Nya mudah-mudahan skripsi ini bermanfaat bagi penulis dan bagi pembaca. *Amin.*

*Wassalamu'alaikum Warahmatullohi Wabarakatuh*

Malang, 1 Juni 2020

Penulis

## DAFTAR ISI

<b>HALAMAN JUDUL</b>	
<b>HALAMAN PENGAJUAN</b>	
<b>HALAMAN PERSETUJUAN</b>	
<b>HALAMAN PENGESAHAN</b>	
<b>HALAMAN PERNYATAAN KEASLIAN TULISAN</b>	
<b>HALAMAN MOTO</b>	
<b>HALAMAN PERSEMBAHAN</b>	
<b>KATA PENGANTAR.....</b>	<b>viii</b>
<b>DAFTAR ISI.....</b>	<b>x</b>
<b>DAFTAR TABEL .....</b>	<b>xiii</b>
<b>DAFTAR GAMBAR.....</b>	<b>xiii</b>
<b>DAFTAR SIMBOL .....</b>	<b>xiv</b>
<b>ABSTRAK .....</b>	<b>xvii</b>
<b>ABSTRACT .....</b>	<b>xviii</b>
<b>ملخص.....</b>	<b>xviii</b>
<b>BAB 1 PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	6
1.3 Tujuan Penelitian .....	6
1.4 Manfaat Penelitian .....	7
1.5 Batasan Masalah.....	7
1.6 Sistematika Penulisan .....	8
<b>BAB II KAJIAN PUSTAKA .....</b>	<b>10</b>
2.1 Classification and Regression Tress (CART) .....	10
2.2 Bootstrap Aggregating (Bagging) .....	19
2.3 Cross Validation V–Fold Estimation .....	20
2.4 Ukuran Ketepatan Klasifikasi .....	22
2.5 Penerimaan Siswa Baru.....	24
2.6 Pendidikan Dalam Prespektif Islam .....	26
<b>BAB III METODOLOGI PENELITIAN .....</b>	<b>29</b>
3.1 Pendekatan Penelitian .....	29
3.2 Jenis dan Sumber Data .....	29
3.3 Identifikasi Variabel.....	29

3.4	Langkah-langkah Penelitian.....	31
3.5	Flow Chart.....	39
<b>BAB IV ANALISIS DATA DAN PEMBAHASAN.....</b>		<b>40</b>
4.1	Analisis Data .....	40
4.2	Analisis CART ( <i>Classification and Regression Trees</i> ) .....	47
4.3	Ketepatan Klasifikasi CART.....	56
4.4	<i>Bootstrap Aggregating (Bagging) Classification and Regression Tree (CART)</i> .....	58
4.5	Perbandingan Ketepatan Klasifikasi Analisis CART dan Analisis <i>Bagging</i> CART .....	61
4.6	Kajian Al-Quran Tentang <i>Bootstrap Aggregating Classification and Regression Tree</i> .....	61
<b>BAB V PENUTUP.....</b>		<b>63</b>
5.1	Kesimpulan .....	64
5.2	Saran.....	65

**DAFTAR RUJUKAN**

**LAMPIRAN**

**RIWAYAT HIDUP**

## DAFTAR TABEL

Tabel 2.1	Tabel Prediksi .....	22
Tabel 3. 1	Data <i>Learning</i> .....	33
Tabel 3. 2	Data <i>Testing</i> .....	33
Tabel 3. 3	Data Sampel <i>Learning</i> Pertama .....	34
Tabel 3. 4	Data Sampel <i>Learning</i> Kedua .....	34
Tabel 3. 5	Data Sampel <i>Learning</i> Ketiga.....	34
Tabel 3. 6	Hasil Klasifikasi dari Data <i>Testing</i> Pohon Pertama .....	36
Tabel 3. 7	Hasil Klasifikasi dari Data <i>Testing</i> Pohon Kedua.....	37
Tabel 3. 8	Hasil Klasifikasi dari Data <i>Testing</i> Pohon Ketiga .....	37
Tabel 3. 9	Kesimpulan <i>Majority Vote</i> Data <i>Learning</i> .....	38
Tabel 4. 1	Deskriptif Karakteristik Diterima dan Tidak Diterima Indikator Nilai Matematika .....	43
Tabel 4. 2	Deskriptif Karakteristik Diterima dan Tidak Diterima Indikator Nilai IPA.....	44
Tabel 4. 3	Deskriptif Karakteristik Diterima dan Tidak Diterima Indikator Nilai Bahasa Inggris .....	45
Tabel 4. 4	Deskriptif Karakteristik Diterima dan Tidak Diterima Indikator Nilai Tes IQ .....	46
Tabel 4. 5	Ketepatan Klasifikasi Presentase Data.....	48
Tabel 4. 6	Banyaknya Pemilihan Pemilah Variabel Prediktor.....	49
Tabel 4. 7	Skor Variabel Prediktor Pohon Maksimal .....	50
Tabel 4. 8	Skor Variabel Prediktor Pohon Otimal CART .....	53
Tabel 4. 9	Presentase Keakurasian Pada Data <i>Learning</i> .....	56
Tabel 4. 10	Klasifikasi Data <i>Testing</i> .....	58
Tabel 4. 11	Nilai Ketepatan Klasifikasi Data <i>Testing</i> .....	58
Tabel 4. 12	<i>Variable Important</i> .....	59
Tabel 4. 13	Ketepatan Klasifikasi Data <i>Testing Bagging</i> CART .....	60
Tabel 4. 14	Ketepatan Klasifikasi <i>Bagging</i> CART .....	60
Tabel 4. 15	Nilai Ketepatan Klasifikasi CART dan <i>Bagging</i> CART .....	61

## DAFTAR GAMBAR

Gambar 2. 1 Pohon Klasifikasi .....	12
Gambar 2. 2 Ilustrasi <i>Bagging</i> CART.....	20
Gambar 3. 1 Pohon Optimal Pertama .....	35
Gambar 3. 2 Pohon Optimal Kedua .....	35
Gambar 3. 3 Pohon Optimal Ketiga.....	36
Gambar 3. 4 <i>Flow Chart Bagging</i> CART .....	39
Gambar 4. 1 Jumlah Peserta Didik Baru Berdasarkan Hasil Seleksi .....	41
Gambar 4. 2 Perbandingan Diterima dan Tidak Diterima Berdasarkan Asal Sekolah.....	41
Gambar 4. 3 Perbandingan Diterima dan Tidak Diterima Indikator Jenis Kelamin .....	42
Gambar 4. 4 Pohon Klasifikasi Maksimal CART.....	51
Gambar 4. 5 Plot CP (Complexity Parameter).....	52
Gambar 4. 6 Pohon Klasifikasi Optimal CART .....	53

## DAFTAR SIMBOL

- $i(t)$  : *Impuriy* atau fungsi keheterogenan simpul  $r$   
 $p(i|t)$  : Proporsi kelas  $i$  pada simpul  $t$   
 $p(j|t)$  : Proporsi kelas  $j$  pada simpul  $t$   
 $\emptyset(s, t)$  : Nilai *goodness of split*  
 $p_L$  : Nilai proporsi pengamatan simpul kiri  
 $P_R$  : Nilai proporsi pengamatan simpul kanan  
 $i(t_L)$  : Nilai dari fungsi heterogenitas pada simpul anak kiri  
 $i(t_r)$  : Nilai dari fungsi heterogenitas pada simpul anak kanan.  
 $p(j_0|t)$  : Proporsi kelas  $j$  pada simpul  $t$   
 $N_j(t)$  : Banyaknya amatan kelas  $j$  pada *terminal node*  $t$   
 $N(t)$  : Jumlah total pengamatan dalam *terminal node*  
 $p(j_0|t)$  : Proporsi kelas  $j$  pada simpul  $t$   
 $N_j(t)$  : Banyaknya amatan kelas  $j$  pada *terminal node*  $t$   
 $N(t)$  : Jumlah total pengamatan dalam *terminal node*  
 $R^{ts}(T_t)$  : Total proporsi kesalahan *test sample estimate*  
 $N_2$  : Jumlah pengamatan dari data *learning* ( $L_1$ )  
 $X(\cdot)$  : Bernilai 0 ketika pernyataan salah dan bernilai 1 ketika pernyataan benar  
 $n_{11}$  : Banyaknya observasi  $Y$  kelas 1 yang diprediksi secara benar dengan  
 $n_{12}$  : Banyaknya observasi  $Y$  kelas 1 yang diprediksi secara tidak benar dengan observasi  $Y$  kelas 1

- $n_{21}$  : Banyaknya observasi Y kelas 2 yang diprediksi secara tidak benar dengan observasi Y kelas 2
- $n_{22}$  : Banyaknya observasi Y kelas 2 yang diprediksi secara benar dengan observasi Y kelas 2
- $N_{1a}$  : Banyaknya observasi Y kelas 1
- $N_{2a}$  : Banyaknya observasi Y kelas 2
- $N_{1b}$  : Banyaknya prediksi Y kelas 1
- $N_{2b}$  : Banyaknya prediksi Y kelas 2
- $N$  : Banyaknya total prediksi/observasi
- APER : Jumlah tingkat kesalahan klasifikasi

## ABSTRAK

Jannah, Dwi Miftakhul. 2020. **Klasifikasi Penerimaan Peserta Didik Baru di Madrasah Bertaraf International Amanatul Ummah Pacet Menggunakan *Bootstrap Aggregating Classification and Regression Tree***. Skripsi Jurusan Matematika Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing (1) Heni Widayani, M.Si. (II) Angga Dwi Mulyanto, M.Si.

**Kata kunci:** *10-fold cross validation*, Akurasi, *Bagging* CART, CART, Ketepatan Klasifikasi, Penerimaan

Pendaftaran peserta didik baru merupakan kegiatan setiap tahunnya yang dilakukan setiap sekolah di Indonesia. Madrasah Bertaraf International Amanatul Ummah merupakan sekolah berbasis islam yang banyak diminati oleh masyarakat. Tahun 2019 terdapat 1194 pendaftar, hanya 340 peserta didik yang diterima. Penerimaan peserta didik baru dikategorikan menjadi dua tipe yaitu diterima dan tidak diterima. Gelombang pertama tahun 2020 terdapat 390 pendaftar, 110 diantaranya diterima dan 280 pendaftar tidak diterima. Klasifikasi penerimaan peserta didik baru berguna untuk masyarakat umum yang tertarik mendaftar di sekolah tersebut. CART sesuai dalam pengelompokan pada kasus penerimaan peserta didik baru sedangkan *Bagging* CART untuk meningkatkan tingkat akurasi. Data dibagi menjadi 85% data *learning* dan 15% data *testing*. CART menggunakan *10-fold cross validation* untuk menghasilkan tingkat akurasi pada data *testing* sebesar 86,21% sedangkan *Bagging* CART sebesar 96,52%. *Bagging* CART mampu meningkatkan akurasi ketepatan klasifikasi sebesar 10,31% pada penelitian ini. Sedangkan faktor paling penting dalam klasifikasi penerimaan peserta didik baru yaitu nilai matematika dan bahasa inggris.

## ABSTRACT

Jannah, Dwi Miftakhul. 2020. **Classification of Receiving New Students in International Standart School of Amanatul Ummah Pacet Using Bootstrap Aggregating Classification and Regression Tree.** Thesis Department of Mathematics, Faculty of Science and Technology, Islamic State University of Maulana Malik Ibrahim Malang. Advisors (I) Heni Widayani, M.Si. (II)Angga Dwi Mulyanto, M.Si.

**Keyword:** 10-fold cross validation, Accuration, Admission, Bagging CART, CART, Classification Accuracy.

Every schools in Indonesia conduct new student registration activities every year. International Standart School of Amanatul Ummah is an Islamic based school that is much in demand by the comunity. In 2019 there were 1194 students, only 340 students were accepted. New student admissions are categorized into two types that is, accepted and not accepted. The first wave in 2020 there were 290 students, 110 of them were accepted and 280 students were not accepted. The new students admission classification is useful for the general public who are interested in registering at the school. CART is suitable for grouping in the case of acceptance of new students while *Bagging* CART is for increasing the level of accuracy. The data is divided into 85% learning data and 15% testing data. CART uses 10-fold cross validation to produce an accuracy level of testing data is 86,21% while *Bagging* CART is 96,52%. *Bagging* CART is able to improve the accuracy of classification accuracy classification error in this study equal to 10,31%. The most important factor in classification of new student admissions at the International Standart School of Amanatul Ummah is the value of mathematic and english.

## ملخص

جنة,دوي مفتاحل. 2020. تصنيف استلام الطلاب الجدد في المدرسة على مستوى الدولي أمانة الأمة بباچيت باستخدام شجرة بوتستراب للتصنيف و والانمدار التجميعة. البحث الجامعي. قسم الرياضيات كلية العلوم والتكنولوجيا الجامعة مولانا مالك إبراهيم الإسلامية الحكومية مالانج. المشرف: (1) هيني ويداياني، الماجستير (2) انجا دوى مولياتنو، الماجستير.

الكلمات الأساسية: *10-fold cross validation*، التنجيم، *CART*، *Bagging CART*، صحة التصنيف، استلام

كان جميع المدارس في اندونيسيا أنشطة تسجيل الطلاب الجدد كل عام. المدرسة على مستوى الدولي أمانة الأمة على الشريعة الإسلامية مطلوباً للمجتمع. في عام 2019 هناك 1194 مسجلاً، الاستلام منهم 340 طالباً فقط. تصنيف استلام الطلاب إلى نوعين وهما المقبولين وغير المقبولين. الموجة الأولى في عام 2020 و هناك 390 مسجلاً، 110 منهم قبولاً، و 280 مسجلاً غير مقبول. تصنيف استلام الطلاب الجدد مفيد للمجتمع راغب في تسجيل إلى تلك المدرسة. *CART* وفقاً للتجميع في حالة استلام الطلاب الجدد، بينما *Bagging CART* ليرفع دورالتنجيم. تنقسم البيانات إلى 85% لبيانات التعليم و 15% لبيانات الاختبار. يستخدم *CART* با *10-fold cross validation* لينتج مستوى على بيانات الانتبار إلى 86,21% بينما *Bagging CART* إلى 96,25%. *Bagging CART* قادرة على تحسين صحة التنجيم و لديها خطأ التصنيف هي 10,31% في هذا البحث. أهم عوامل في تصنيف استلام الطلاب الجدد في المدرسة على مستوى الدولي أمانة الأمة هي النتيجة ف الرياضيات واللغة الإنجليزية.

## BAB 1

### PENDAHULUAN

#### 1.1 Latar Belakang

Berdasarkan Undang-Undang Republik Indonesia Sistem Pendidikan Nasional Pasal 3, fungsi dari pendidikan nasional adalah untuk mengembangkan kemampuan dan membentuk watak serta peradapan bangsa yang bermartabat dalam rangka mencerdaskan kehidupan bangsa serta bertujuan untuk berkembangnya potensi peserta didik agar menjadi manusia yang beriman dan bertaqwa kepada Tuhan Yang Maha Esa, berakhlak mulia sehat, berilmu, cakap, kreatif, mandiri dan menjadi negara yang demokratis serta bertanggung jawab. (Indonesia, 2003)

Allah SWT memerintahkan kepada umatnya untuk selalu membaca yang dijelaskan pada salah satu firmannya seperti dibawah ini:

*“Bacalah dengan (menyebut) nama Tuhanmu yang menciptakan, Dia telah menciptakan manusia dari ‘Alaq, bacalah dan Tuhanmulah yang paling pemurah, yang mengajar manusia dengan pena dan dia mengajarkan kepada manusia apa yang belum diketahuinya” (QS. Al-‘Alaq /1-5)*

Pada Q.S Al-‘Alaq ayat 1-5 menyinggung sebuah pendidikan. Surat tersebut diawali dengan kata *iqra*’ yang memiliki arti “bacalah”. Berarti yang dimaksudkan dengan kata *iqra*’ memerintahkan kita untuk membaca, meneliti dan mendalami mengenai ciri-ciri, tanda-tanda serta sejarah baik itu dalam bentuk tertulis maupun tidak tertulis. Ayat ini jelas sekali perintah pertama yang ditunjukkan oleh Allah kepada manusia adalah perintah membaca, sebab membaca merupakan syarat utama dalam mengembangkan ilmu pengetahuan dan teknologi yang ada di muka muka bumi ini (Mustolehudin, 2011). Keterangan ayat tersebut dapat diketahui

pentingnya membaca, dimana membaca merupakan salah satu bagian dari pendidikan.

Tujuan penting dari pendidikan, yang pertama membuat peserta didik menentukan kualitas diri yaitu keinginan peserta didik dalam mengembangkan diri untuk mencapai masa depan yang diinginkan. Kedua untuk kegiatan sosial yaitu peserta didik dapat menerapkan ilmu sosial dari hasil menempuh pendidikannya. Fungsi pendidikan bukan hanya untuk melakukan kegiatan nasional yang telah ditentukan oleh pemerintah tetapi sebagai salah satu kekuatan sosial yang memberikan dampak positif kehidupan di masa depan (Rosadi, 2012). Oleh karena itu perlunya peserta didik mendapatkan kualitas pendidikan yang baik. Hal yang perlu dikembangkan dan diperhatikan pada manajemen sekolah adalah manajemen siswa.

Menurut Suling (2015), kegiatan perencanaan dalam manajemen kesiswaan sangat penting dilakukan untuk mengetahui banyaknya peserta didik baru yang akan diterima. Kegiatan penerimaan peserta didik baru merupakan salah satu kegiatan rutin yang diselenggarakan oleh sekolah setiap tahunnya. Kegiatan ini diketahui oleh masyarakat luas tanpa membedakan asal-usul, ras, agama, dan suku (Hermawan, 2006). Setiap sekolah memiliki kriteria-kriteria penerimaan peserta didik baru yang berbeda-beda.

Penerimaan peserta didik baru didasarkan oleh beberapa hal, yang pertama didasarkan oleh kebijakan penerimaan peserta didik baru yaitu setiap peserta didik yang diterima wajib memenuhi persyaratan yang telah ditentukan. Kedua, didasarkan oleh sistem seleksi penerimaan peserta didik baru. Terdapat dua

macam seleksi yang sering digunakan yaitu sistem seleksi dan sistem promosi. Ketiga, didasarkan oleh kriteria penerimaan peserta didik baru yaitu kriteria ini sebagai penentu peserta tersebut diterima atau tidak dengan kriteria yang telah ditentukan. Terakhir didasarkan oleh prosedur penerimaan (Suking, 2015).

MA Unggulan Amanatul Ummah memiliki beberapa program khusus, salah satunya Madrasah Bertaraf International (MBI) yang sudah terakreditasi A. Di lembaga ini setiap tahunnya selalu lulus 100% dalam Ujian Nasional (UN) dan sebagian besar (98%) peserta didik kelas XII melanjutkan ke perguruan tinggi. Sampai saat ini Madrasah Bertaraf Internasional Amanatul Ummah telah menunjukkan prestasi-prestasi yang gemilang di raih oleh siswa-siwanya di tingkat lokal, regional, nasional dan internasional. Sekolah ini menyajikan tiga kurikulum dalam pembelajarannya yaitu kurikulum nasional, kurikulum international dan kurikulum Al-Azhar (Kairo-Mesir).

Berdasarkan profil sekolah Madrasah Bertaraf Internasional Amanatul Ummah Pacet menunjukkan bahwa sekolah tersebut berprestasi di bidang pendidikan. Sehingga banyaknya peminat dari seluruh Indonesia untuk belajar di sekolah tersebut, pernyataan ini sesuai dengan data tahun 2019 dimana sebanyak 1194 calon peserta didik baru. Sebanyak 384 peserta didik baru berjenjang SMP yang dinyatakan diterima di sekolah tersebut. Setiap tahunnya Madrasah Bertaraf International Amanatul Ummah pacet membuka dua cara untuk bisa menempuh pendidikan di sekolah tersebut yaitu jalur prestasi dan jalur tes tulis.

Salah satu metode statistika nonparametrik yang dapat menggambarkan dan menjelaskan hubungan suatu variabel respon (dependen) dengan satu atau lebih

variabel prediktor (independen) yaitu metode CART (*Classification and Regression Trees*). Metode CART dapat digunakan untuk mendapatkan suatu kelompok data yang akurat sebagai pencari dari suatu pengklasifikasian (Leo Breiman, 1998). Pada metode ini dilakukan untuk pembentukan pohon klasifikasi sesuai dengan variabel yang digunakan. Perubahan kecil pada data *learning* akan mengakibatkan hasil pohon klasifikasi tidak stabil dan akan mempengaruhi tingkat keakurasian prediksi. Metode *Bootstrap Aggregating (Bagging) Classification and Regression Trees* (CART) dapat digunakan untuk memperbaiki stabilitas dan kekuatan prediksi pohon klasifikasi.

Penelitian penggunaan metode *Classification and Regression Trees* (CART) pernah dilakukan oleh, Sumartini (2015) klasifikasi rekruesi pasien kanker serviks di RSUD DR. Soetomo Surabaya yang menghasilkan rata-rata nilai akurasi klasifikasi CART untuk data *learning* sebesar 77,75% dan 69,14% untuk data prediksi yang terbukti bahwa penggunaan metode CART ini mempunyai tingkat keakurasian yang tinggi. Pembagian data *learning* dan *testing* menggunakan metode *5-fold cross validation*. Perhitungan ketepatan klasifikasi pohon dengan akurasi, *sensitivity*, dan *specificity* didapatkan nilai-nilai sebagai berikut 69,14%, 64,91% dan 72,22%.

Analisis perbandingan kinerja metode klasifikasi dalam data mining dilakukan oleh Wibowo (2015). Ada dua metode yang digunakan yaitu metode CART dan *Bagging* CART. Menurut perhitungan keakurasian dari masing-masing metode didapatkan 75% untuk metode CART dan untuk metode *Bagging* CART didapatkan 87,5%. Sehingga dapat disimpulkan bahwa metode *Bagging*

CART lebih akurasi untuk pengklasifikasian. Hal ini disebabkan karena metode *Bagging* CART ada pembangkitan data *learning* yang akan mereduksi variansi atribut prediktor, sehingga ketika dikombinasikan hasilnya lebih baik bila dibandingkan dengan prediktor tunggal yang akan dibangun untuk menyelesaikan masalah yang sama.

Pada penelitian ini, membahas tentang pengelompokkan nasabah kredit lancar atau tidak lancar dilakukan oleh (Desy Ratnaningrum, 2016) dengan menggunakan metode *Bagging Classification and Regression Trees* menghasilkan nilai ketetapan klasifikasi sebesar 81,44%. Nilai tersebut menunjukkan bahwa metode *Bagging* CART dapat mengklasifikasikan dengan tepat sebesar 81,44%. Bahwa metode *Bagging Classification and Regression trees* dapat diterapkan untuk mengklasifikasikan nasabah kredit lancar atau tidak lancar.

Oleh karena itu peneliti akan membahas tentang metode *Bagging Classification and Regression Trees (CART)* untuk mengklasifikasikan penerimaan peserta didik baru. Data yang digunakan pada penelitian ini yaitu data sekunder penerimaan peserta didik baru di Madrasah Bertaraf International Amanatul Ummah. Pada penelitian ini peneliti menyajikan dengan judul “Klasifikasi Penerimaan Peserta Didik Baru di Madrasah Bertaraf International (MBI) Amanatul Ummah dengan menggunakan *Bagging Classification and Regression Trees (CART)*).

## 1.2 Rumusan Masalah

Berdasarkan latar belakang, rumusan masalah pada penelitian ini adalah:

1. Bagaimana variabel yang mempengaruhi penerimaan peserta didik baru di Madrasah Bertaraf International (MBI) Amanatul Ummah Pacet menggunakan *Bootstrap Aggregating (Bagging) Classification and Regression Tree*?
2. Bagaimana tingkat akurasi yang didapatkan dari klasifikasi penerimaan peserta didik baru di Madrasah Bertaraf International (MBI) Amanatul Ummah Pacet menggunakan *Bootstrap Aggregating (Bagging) Classification and Regression Tree*?

## 1.3 Tujuan Penelitian

Berdasarkan rumusan masalah yang telah dijelaskan, maka tujuan penelitian ini sebagai berikut:

1. Mengetahui variabel yang mempengaruhi penerimaan peserta didik baru di Madrasah Bertaraf International (MBI) Amanatul Ummah Pacet menggunakan *Bootstrap Aggregating (Bagging) Classification and Regression Tree*.
2. Mengetahui tingkat akurasi yang didapatkan dari klasifikasi penerimaan peserta didik baru di Madrasah Bertaraf International (MBI) Amanatul Ummah Pacet menggunakan *Bootstrap Aggregating (Bagging) Classification and Regression Tree*.

#### 1.4 Manfaat Penelitian

Berdasarkan tujuan penelitian, maka manfaat penelitian pada penelitian ini adalah :

1. Penelitian ini berguna untuk memberikan informasi kepada masyarakat yang berminat untuk mendaftar di sekolah Madrasah Bertaraf International Amanatul Ummah terkait klasifikasi penerimaan peserta didik baru. Selain itu, dengan menggunakan metode *Bagging* CART diharapkan dapat memberikan kemudahan dalam proses klasifikasi penerimaan peserta didik baru di sekolah Madrasah Bertaraf International Amanatul Ummah.
2. Penelitian ini diharapkan dapat menambah wawasan dan pengetahuan bagi pembaca tentang klasifikasi menggunakan *Bagging* CART.

#### 1.5 Batasan Masalah

Batasan masalah pada penelitian ini sebagai berikut:

1. Data yang digunakan merupakan data sekunder dari hasil tes ujian masuk Madrasah Bertaraf Internasional (MBI) Amanatul Ummah Pacet pada gelombang satu tahun 2020 yang dilaksanakan pada tanggal 22 februari 2020.
2. Klasifikasi kelulusan masuk melalui tes ujian masuk Madrasah Bertaraf Internasional (MBI) Amanatul Ummah Pacet pada gelombang satu tahun 2020 menggunakan metode *Bootsrap Aggregating (Bagging) Classification and Regression Tress (CART)*.

## 1.6 Sistematika Penulisan

Sistematika penulisan digunakan untuk mempermudah memahami dalam penyusunan atau penulisan penelitian ini. Adapun beberapa bab yang terdapat di penulisan penelitian sebagai berikut:

### BAB I PENDAHULUAN

Pada bab pendahuluan ini dijelaskan tentang latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah dan sistematika penulisan.

### BAB II KAJIAN PUSTAKA

Pada bab II dijelaskan tentang kajian teori-teori yang digunakan pada penelitian ini, diantaranya adalah *classification and regression trees* (CART), *cross validation V-fold estimation*, *bootstrap aggregating* (*Bagging*), ukuran ketepatan klasifikasi, penerimaan siswa baru dan pendidikan dalam prespektif islam.

### BAB III METODOLOGI PENELITIAN

Pada bab ini berisi penjelasan metode penelitian yang digunakan dalam penelitian, yaitu sebagai berikut: pendekatan penelitian, jenis dan sumber data, variabel penelitian, langkah-langkah penelitian dan *flow chart*.

### BAB IV PEMBAHASAN

Pada bab ini akan dipaparkan dan dijelaskan mengenai hasil penelitian yang telah dilakukan pada studi kasus.

## BAB V KESIMPULAN

Pada bab ini akan dipaparkan tentang kesimpulan dari penelitian yang sudah dilakukan dan saran untuk peneliti selanjutnya.



## BAB II

### KAJIAN PUSTAKA

#### 2.1 *Classification and Regression Trees (CART)*

*Classification and Regression Trees (CART)* merupakan salah satu metode statistika nonparametrik dengan menggunakan teknik pohon klasifikasi. Tujuan dari metode ini adalah untuk menghasilkan suatu kelompok klasifikasi yang akurat dari suatu penelitian. CART digunakan untuk menggambarkan hubungan dari variabel respon (independen atau tak bebas) dengan variabel prediktor (dependen atau bebas) dalam bentuk pohon klasifikasi. CART dapat dikatakan metode yang sederhana tapi metode ini memiliki tingkat akurasi yang tinggi. Pada metode ini dapat menghasilkan dua macam pohon, diantaranya adalah pohon klasifikasi dan pohon regresi. Jika variabel respon bertipe kontinu atau numerik dengan skala rasio atau interval maka yang dihasilkan adalah pohon regresi, sedangkan jika variabel respon bertipe kategorik dengan rasio nominal atau ordinal maka yang dihasilkan adalah pohon klasifikasi. Pohon CART dapat dikatakan *binary decision tree* karena pada metode ini memisahkan *node* menjadi dua *child node* secara berulang-ulang dengan diawali *parent node* (Leo Breiman, 1998).

Lewis (2000) menyatakan bahwa proses pembuatan pohon keputusan menggunakan metode CART dikenal dengan *binary recursive partitioning*. *Binary* yang artinya memisahkan *parent node* menjadi dua *child node* yang diterapkan pada sekelompok data yang ada di ruang yang sama. Istilah *recursive* dapat diartikan proses biner yang dilakukan secara berulang-ulang. *Recursive*

didapatkan dengan pemisahan *parent node* menjadi dua *child node*, dan kedua *child node* tersebut akan menjadi *parent node* kemudian dilakukan pemisahan menjadi dua *child node* dimasing-masing *parent node*, proses ini dilakukan sampai tidak dapat lagi untuk dilakukan pemisahan lagi. Sedangkan istilah *partitioning* diartikan proses klasifikasi dilakukan dengan pemilahan data *learning* menjadi bagian-bagian yang lebih kecil.

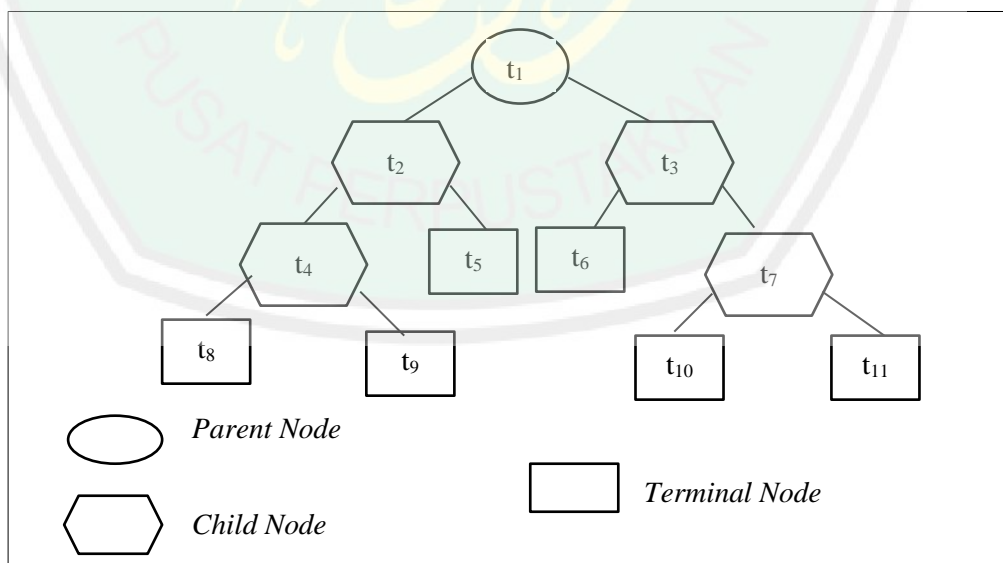
Secara rinci bentuk struktural CART dimulai dengan simpul awal sebagai pemilah terbaik dan terpenting untuk pengklasifikasian dalam kelas-kelas pada data *learning* yang akan diteliti. Simpul awal dapat disebut juga dengan *parent node*. *Parent node* akan dipilah menjadi dua bagian yaitu simpul anak kiri dan kanan, simpul anak dapat disebut dengan *child node*. Setiap simpul hanya bisa dipilah menjadi dua simpul anak baru, dimana simpul awal sebagai simpul induk. Pada proses biner kedua *child node* akan menjadi simpul induk dan dipilah menjadi dua simpul anak baru. Proses ini dilakukan secara berulang-ulang hingga terbentuk *terminal node*. *Terminal node* adalah sebuah simpul yang sudah tidak bisa dipilah lagi dan data yang ada pada simpul ini bersifat tidak homogen atau bisa disebut heterogen (Lewis, 2000).

Menurut Lewis (2000) adapun beberapa kelebihan dari metode *Classification and Regression Tress* (CART) sebagai berikut:

1. CART masuk dalam kategori metode statistika nonparametrik dikarenakan metode ini tidak mendasarkan variabel respon dan variabel prediktor pada sebuah distribusi probabilitas.

2. Variabel responnya dapat berupa kategorik atau kontinu dan variabel prediktornya dapat berupa rasio, nominal, kategorik atau ordinal.
3. Metode CART tidak berlaku sistem transformasi. Jika data telah di transformasi tidak mengubah hasil pohon keputusan CART.
4. Hasil pohon keputusan berbentuk sederhana sehingga dapat mudah diinterpretasikan dan mudah dipahami oleh kalangan umum.

Pohon klasifikasi yang terbentuk dari metode CART diawali dengan simpul awal atau *parent node* dengan berbentuk oval yaitu  $t_1$ . *Child node* berbentuk segi enam yaitu  $t_2, t_3, t_4$  dan  $t_7$  sedangkan untuk *terminal node* berbentuk persegi yaitu  $t_5, t_6, t_8, t_9, t_{10}$  dan  $t_{11}$ . Pada pohon klasifikasi terdapat level-level dalam setiap node, level pertama terletak pada  $t_1$ , sedangkan  $t_2$  dan  $t_3$  terletak pada level kedua dan begitu seterusnya sampai pada *terminal node*  $t_8, t_9, t_{10}$  dan  $t_{11}$  terletak pada level empat. Ilustrasi struktur pohon klasifikasi terdapat pada Gambar 2.1:



Gambar 2. 1 Pohon Klasifikasi

(Sumber : Leo Breiman, 1998)

Terdapat 3 langkah untuk pembentukan pohon klasifikasi menggunakan algoritma CART diantaranya adalah pembentukan pohon klasifikasi, pemangkasan pohon klasifikasi dan penentuan pohon klasifikasi optimal.

### 2. 1.1 Pembentukan Pohon Klasifikasi

Hal pertama yang dilakukan untuk membentuk pohon klasifikasi adalah menyiapkan banyaknya data variabel-variabel yang layak digunakan menjadi pemilah untuk tiap simpul. Pembentukan pohon ini dibutuhkan data *learning* sebanyak  $L$  yang terdiri dari atas  $N$  amatan. Tingkat keakurasian pohon dapat dilihat dari ukuran pohon. Ada tiga tahapan untuk pembentukan pohon klasifikasi sebagai berikut:

#### a. Pemilihan Pemilah

Tujuan pada proses pemilihan pemilah adalah untuk mengurangi keheterogenan simpul induk atau *parent node* dan meningkatkan tingkat kehomogenan data pada simpul anak atau *child node*. Data yang digunakan adalah data *learning*  $L$  yang kemudian akan dipilah menggunakan aturan *Goodness of split*. Menurut Leo Breiman (1998) indeks gini merupakan fungsi keheterogenitas yang paling umum. Kelebihan metode indeks gini adalah perhitungannya yang sederhana, mudah dipahami dan layak digunakan untuk semua kasus. Pemilah terbaik dipilih dari semua kemungkinan pemilah variabel prediktor yang memiliki nilai penurunan keheterogenitas yang tinggi. Pemilah terbaik digunakan untuk menjadi induk pemilah pada berbagi kasus di satu kelompok. Persamaan indeks gini dapat dituliskan seperti persamaan 2.1:

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

dengan:

$i(t)$  = *impurity* atau fungsi keterogenan simpul  $t$

$p(i|t)$  = proporsi kelas  $i$  pada simpul  $t$

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

Menurut Leo Breiman (1998), aturan pemilahan *parent node* menjadi 2 *child node* yaitu masing-masing pemilah hanya bergantung pada nilai satu variabel prediktor yang sudah ditentukan. Dua *child node* harus lebih homogen dari *parent node*nya. Hal ini dapat dilakukan dengan menganalisis fungsi keherogenitas pada setiap *node* (*impurity/  $i(t)$* ). Adapun beberapa ketentuan-ketentuan untuk menentukan jumlah pemilah pada setiap variabel prediktor berdasarkan masing-masing jenis skala sebagai berikut:

1. Jika  $X_j$  merupakan variabel berskala kategorik, maka pemilahan diperoleh dari semua kemungkinan pemilahan berdasarkan simpul yang telah terbentuk yang saling lepas atau *disjoint*. Ketika  $L$  merupakan banyaknya pemilah untuk skala ordinal maka dapat diperoleh banyaknya pemilah sebesar  $L - 1$ . Sedangkan ketika  $L$  merupakan banyaknya pemilah untuk skala kategorik maka diperoleh banyaknya pemilah sebesar  $L^{L-1} - 1$ .
2. Jika  $X_j$  merupakan variabel berskala kontinu, maka pemilahan yang diperbolehkan  $X_j \leq M_i$  sebagai simpul kiri dan  $X_j > M_i$  simpul kanan dengan  $i = 1, 2, 3 \dots, n - 1$ , dimana  $M_i$  adalah nilai tengah atau *median* pada setiap dua data amatan berurutan yang berbeda. Misalkan banyaknya

amatan berjumlah  $N$ , dan banyaknya data yang berbeda sama dengan  $N$  maka banyaknya pemilah berskala kontinu sebesar  $N - 1$  kemungkinan banyaknya pemilah yang berbeda.

Kemudian dilakukan proses *goodness of split* yang bertujuan untuk menentukan pemilah terbaik ditunjukkan dengan nilai penurunan keheterogenitasnya yang paling tinggi dari semua variabel prediktor. Berikut ini merupakan persamaan untuk mencari nilai *goodness of split* seperti persamaan 2.2:

$$\phi(s, t) = \Delta i(s, t) = i(t) - P_L i(t_L) - P_R i(t_R) \quad (2.2)$$

dengan:

$\phi(s, t)$  = nilai *goodness of split*

$i(t)$  = nilai dari fungsi heterogenitas pada simpul  $t$

$P_L$  = nilai proporsi pengamatan simpul kiri

$P_R$  = nilai proporsi pengamatan simpul kanan

$i(t_L)$  = nilai dari fungsi heterogenitas pada simpul anak kiri

$i(t_R)$  = nilai dari fungsi heterogenitas pada simpul anak kanan.

#### b. Penentuan Simpul/*Terminal node*

Tahap kedua dari pembentukan pohon klasifikasi adalah penentuan *terminal node*. Dimana pada tahap ini penentuan untuk simpul  $t$  dapat di lanjutkan untuk menjadi pemilah atau menjadi *terminal node*. Simpul  $t$  dapat dijadikan *terminal node* jika sudah tidak terjadi penurunan heterogenitas yang berarti. Menurut Leo Breiman (1998), adapun kriteria untuk menentukan simpul  $t$  menjadi *terminal node* yaitu jika pada simpul tersebut banyaknya  $n < 5$  dan pembentukan pohon

akan berhenti jika sudah sampai pada level yang ditentukan dalam pohon maksimal.

### c. Penandaan Label Kelas

Tujuan dari penandaan label kelas pada setiap *terminal node* adalah untuk mengetahui karakteristik dan klasifikasi pada hasil pengamatan disetiap kelas variabel respon yang terbentuk. Penentuan penandaan label kelas disetiap *terminal node* dapat menggunakan aturan jumlah terbanyak dengan persamaan 2.3 (Leo Breiman, 1998):

$$p(j_0|t) = \max p(j|t) = \max \frac{N_j(t)}{N(t)} \quad (2.3)$$

dengan:

$p(j_0|t)$  = proporsi kelas  $j$  pada simpul  $t$

$N_j(t)$  = banyaknya amatan kelas  $j$  pada *terminal node*  $t$

$N(t)$  = jumlah total pengamatan dalam *terminal node*

$j_0$  merupakan label kelas untuk *terminal node*  $t$ . Asumsi yang digunakan adalah kesalahan klasifikasi untuk setiap kelas besarnya sama, pemberian label kelas  $j_0$  tersebut memberikan nilai dugaan kesalahan pengklasifikasian pada simpul  $t$  paling kecil yaitu sebesar  $r(t) = 1 - p(j_0|t)$ .

### 2.1.2 Pemangkasan Pohon Klasifikasi

Setelah melakukan proses pemilihan pemilih menggunakan kriteria pemilihan dan *goodness of split* sehingga didapatkan pohon klasifikasi dengan ukuran yang sangat besar karena pemberhentian pohon berdasarkan banyaknya simpul dan tingkat kehomogenan suatu simpul. Jika dalam proses pembentukan

pohon diberikan batasan, padahal simpul tersebut masih layak untuk dilakukan pemangkasan maka terjadi *underfitting*. Semakin banyak pemilah maka tingkat keakuratan semakin tinggi atau tingkat kesalahan pada nilai prediksi semakin kecil bisa mengakibatkan *overfitting* yaitu nilai prediksi lebih tinggi dari pada nilai sebenarnya. Pohon yang layak akan didapatkan dengan dilakukan pemangkasan pohon (*pruning*) yaitu proses pengurangan simpul pohon dengan menilai ukuran pohon tanpa mengurangi nilai keakuratan sehingga dicapai pohon yang layak. Proses pemangkasan pohon menggunakan metode *cost complexity* (Leo Breiman, 1998). Ukuran *cost complexity* dengan  $\alpha > 0$  menggunakan persamaan 2.4:

$$R_{\alpha}(T) = R(T) + \alpha|\tilde{T}| \quad (2.4)$$

dengan:

$R_{\alpha}(T)$  = *resubstitution* suatu pohon T pada kompleksitas  $\alpha$

$\alpha$  = parameter *cost complexity* bagi penambahan satu simpul akhir pohon T

$R(T)$  = *resubstitution estimate* (pendugaan pengganti)

$\tilde{T}$  = banyaknya simpul terminal pohon T

Ada beberapa tahap dalam proses pemangkasan pohon yaitu, mengambil  $t_R$  dan  $t_L$  dari  $T_{max}$  yang telah dihasilkan oleh *parent node* t. Jika diperoleh simpul anak dan simpul induk yang memenuhi persamaan  $R(t) = R(t_L) + R(t_R)$  maka dilakukan pemangkasan. Sehingga diperoleh pohon  $T_1$  yang memenuhi kriteria  $T_1$  yang diperoleh memenuhi kriteria dengan persamaan  $R(T_1) = R(T_{max})$ . Proses ini dilakukan berulang-ulang hingga tidak mungkin lagi dilakukan pemangkasan.

### 2. 1.3 Penentuan Pohon Optimal

Penentuan pohon optimal dapat dilihat dari besar kecilnya pohon, besarnya pohon yang *overfitting* mengakibatkan nilai *cost complexity* besar. Hal itu diakibatkan oleh struktur data yang berbentuk kompleks. Sehingga diperlukan pemilihan pohon optimal dengan pengganti terkecil. Jika  $R(t)$  merupakan pengganti terkecil atau disebut dengan penduga terbaik, maka cenderung dipilih pohon yang besar, karena pohon yang besar memiliki nilai  $R(t)$  kecil.

*Test Sample Estimate* merupakan metode penentuan pohon optimal dengan data yang besar. Langkah pertama pada proses ini adalah membagi data menjadi dua yaitu data *learning* (L1) dan data *testing* (L2). Data *learning* digunakan untuk menentukan pohon klasifikasi sedangkan data *testing* digunakan untuk penentuan nilai  $R^{ts}(T_t)$ . Persamaan *test sample estimate* sebagai berikut:

$$R^{ts}(T_t) = \frac{1}{N_2} \sum_{(x_n, j_n) \in L_2} X(d(x_n) \neq j_n) \quad (2.5)$$

dengan:

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

$N_2$  = jumlah pengamatan dari data *learning* (L1)

$X(\cdot)$  = bernilai 0 ketika pernyataan salah dan bernilai 1 ketika pernyataan benar

Pada proses ini bertujuan untuk menduga proporsi kesalahan dari pohon klasifikasi sehingga dipilih pohon optimal adalah  $T_t$  dengan  $R^{ts}(T_t) = \min_t R^{ts}(T_t)$ .

## 2.2 *Bootstrap Aggregating (Bagging)*

Breiman (1996) mengusulkan suatu teknik yang digunakan untuk mereduksi variansi pada sebuah variabel prediktor pada beberapa metode regresi dan klasifikasi yang bertujuan untuk memperbaiki kualitas prediksi yaitu teknik *bootstrap aggregating*. Istilah *bootstrap* yaitu teknik resampling atau pengambilan suatu data sample dari sebuah variabel prediktor yang secara bebas dan dilakukan secara berulang-ulang untuk mengurangi tingkat kesalahan (*error*). *Bagging* itu sendiri dapat diartikan dengan penggabungan hasil prediksi atau dugaan klasifikasi dari setiap pohon yang sudah dibentuk oleh data yang diambil secara acak, banyaknya pohon berdasarkan banyaknya resampling yang dilakukan. Hasil klasifikasi ditentukan dari dugaan atau prediksi data *testing* voting terbanyak untuk klasifikasi dan rata-rata error terkecil untuk regresi. Pengulangan dilakukan sebanyak 100 pengulangan (Breiman, 1996).

Pada suatu penelitian dilakukan perbandingan dalam menghitung keakurasian hasil *Bagging* yang dibandingkan dengan berapa banyak kali pengulangan. Salah satunya yaitu 100 kali dan menunjukkan bahwa dengan pengulangan sebanyak 100 kali menghasilkan nilai keakurasian yang tinggi dari pada banyaknya percobaan yang lain (Ying Huang, 2004). Salah satu kegunaan dari *Bagging* untuk mengatasi ketidakstabilan dalam pohon klasifikasi dan regresi.

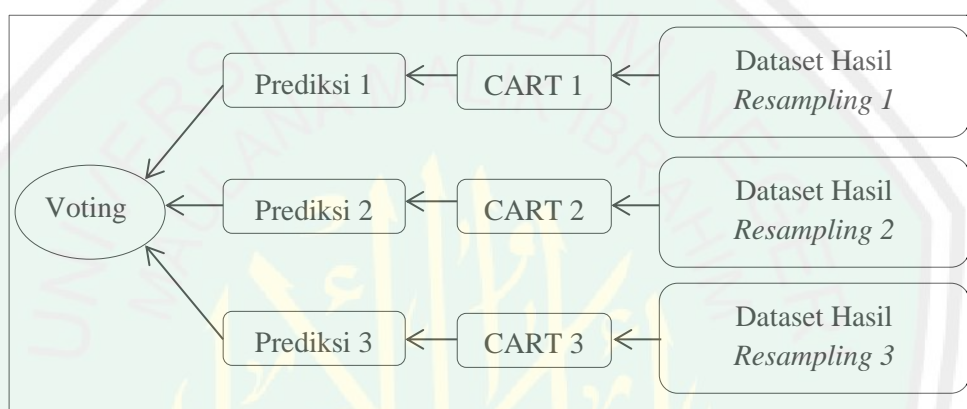
Menurut Sutton (2005), ada beberapa tahapan untuk pembentukan pohon secara *Bagging* sebagai berikut:

1. *Bootstrap* (resampling)
  - a. Ambil sampel acak berukuran  $n$  dari data *learning*.

- b. Bentuk pohon klasifikasi terbaik dari data tersebut
- c. Ulangi tahap a sampai dengan b sebanyak  $n$  kali.

## 2. *Aggregating*

Pada proses ini digunakan untuk menggabungkan hasil dugaan klasifikasi dari setiap pohon klasifikasi dengan banyaknya  $N$  pohon dengan menggunakan aturan *majority vote* (pemilihan suara terbanyak).



Gambar 2. 2 Ilustrasi *Bagging* CART

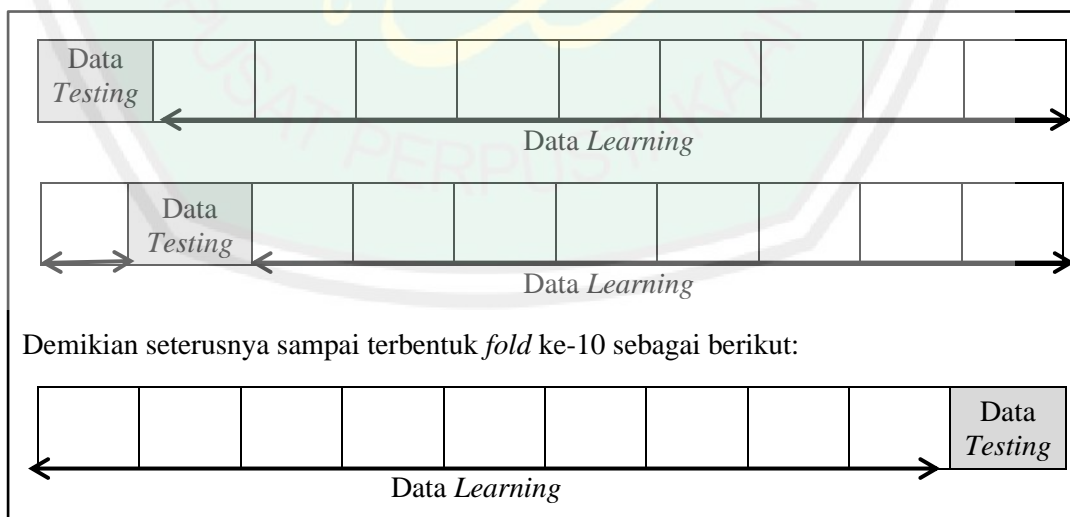
Sumber : (Muhammad Jamal Muttaqin, 2013)

### 2.3 *Cross Validation V-Fold Estimation*

Penduga pengganti ini sering digunakan pada data pengamatan yang relatif kecil atau sedikit. Data pengamatan dibagi secara random sebanyak *V-Fold* dengan ukuran yang sama disetiap fold (bagian) dengan saling lepas. Pembagian data penelitian ini dibagi menjadi 2 (dua) yaitu data learning dan data testing. Setiap bagian (*fold*) yang sudah terbentuk salah satu akan menjadi data *learning* dan yang lainnya akan menjadi data *testing* dalam prosedur pembuatan pohon klasifikasi. Maka akan terbentuk pohon sebanyak *V-Fold* pengulangan data.

Kemudian dihitung rata-rata ketepatan klasifikasi masing-masing *fold* (bagian) (Lewis, 2000).

Permasalahan *overlapping* pada data *testing* sudah banyak terjadi pada penelitian-penelitian sebelumnya. Permasalah itu dapat diatasi dengan menggunakan metode *cross validation*. Pada metode ini nilai  $V$  yang sering digunakan yaitu 10 sehingga menjadi 10 *fold cross validation*. Nilai tersebut merupakan nilai terbaik karena menghasilkan estimasi errornya rendah diperoleh dari proses yang sudah dilakukan. *V-fold cross validation* menyebabkan pembagian data  $L$  menjadi seimbang dan *balance*. Sehingga nilai akurasi yang didapatkan dari data yang sudah *balance* lebih akurat (Ian H. Witten, 2011). Sehingga nilai 10 *fold cross validation* dijadikan sebagai standar klasifikasi pembagian data *learning* dan data *testing*. Jika menggunakan nilai 10 *fold cross validation* akan mendapatkan nilai akurasi sebanyak 10 dari data *learning* dan data *testing*. Berikut ini merupakan ilustrasi dalam pembagian data penelitian:



Gambar 2.3 Ilustrasi *V fold cross validation*

## 2.4 Ukuran Ketepatan Klasifikasi

Pada pengukuran ketepatan klasifikasi terdapat empat hal utama masalah klasifikasi. Pertama variabel respon yang bersifat kategorik. Kedua variabel prediktor memiliki pengaruh terhadap variabel respon. Selanjutnya, adanya data *learning* yang digunakan untuk pembuatan pohon klasifikasi dan yang terakhir adalah penggunaan data *testing* untuk pengukuran ketepatan klasifikasi (Lewis, 2000). Menurut Richard A. Johnson (2002), pengukuran ketepatan klasifikasi dengan menghitung akurasi klasifikasi. Pada penelitian ini menggunakan perhitungan nilai 1-APER (*apperent error rate*) atau total akurasi. Total tingkat kesalahan adalah banyaknya proporsi yang diprediksi secara tidak benar sedangkan total tingkat akurasi total merupakan banyaknya proporsi yang diprediksi secara benar oleh fungsi klasifikasi. Berikut ini menggunakan tabel prediksi untuk menentukan ketepatan klasifikasi (Mock. Abdul Mukid, 2015):

Tabel 2. 1 Tabel Prediksi

Kelas Observasi Y	Kelas Prediksi		Total
	1	2	
1	$n_{11}$	$n_{12}$	$N_{1a}$
2	$n_{21}$	$n_{22}$	$N_{2a}$
<b>Total</b>	$N_{1b}$	$N_{2b}$	N

keterangan :

$n_{11}$  = banyaknya observasi Y kelas 1 yang diprediksi secara benar dengan observasi Y kelas 1

$n_{12}$  = banyaknya observasi Y kelas 1 yang diprediksi secara tidak benar dengan observasi Y kelas 1

$n_{21}$  = banyaknya observasi Y kelas 2 yang diprediksi secara tidak benar dengan observasi Y kelas 2

$n_{22}$  = banyaknya observasi Y kelas 2 yang diprediksi secara benar dengan observasi Y kelas 2

$N_{1a}$  = banyaknya observasi Y kelas 1

$N_{2a}$  = banyaknya observasi Y kelas 2

$N_{1b}$  = banyaknya prediksi Y kelas 1

$N_{2b}$  = banyaknya prediksi Y kelas 2

$N$  = banyaknya total prediksi/observasi

Pengukuran ketepatan klasifikasi menggunakan ukuran-ukuran sebagai berikut:

1. *Apparent Error Rate* (APER) merupakan jumlah tingkat kesalahan klasifikasi, dengan formula di persamaan 2.6:

$$APER = \frac{\text{Jumlah prediksi salah}}{\text{Jumlah total prediksi}} \quad (2.6)$$

2. *Total Accuracy Rate* (1-APER) merupakan jumlah tingkat ketepatan klasifikasi, dengan formula di persamaan 2.7:

$$1 - APER = \frac{n_{11} + n_{22}}{N} \quad (2.7)$$

3. *Sensitivity* menggambarkan tingkat keakurasian pada sampel kelas  $i$ . dengan menggunakan rumus di persamaan 2.8:

$$\text{Sensitivity} = \frac{n_{11}}{N_1} \quad (2.8)$$

4. *Specificity* menggambarkan tingkat keakurasian pada sampel kelas  $j$ . dengan menggunakan rumus di persamaan 2.9:

$$Specificity = \frac{n_{22}}{N_2} \quad (2.9)$$

## 2.5 Penerimaan Siswa Baru

Hal pertama yang dilakukan di manajemen kesiswaan yaitu penerimaan peserta didik. Proses ini dimulai dengan membentuk panitia yang dimana panitia tersebut bukan pengajar dari kelas XII. Ada beberapa panitia yang perlu dibentuk yaitu, panitia lokal, panitia pusat dan panitia lembaga. Tugas dari panitia ini yaitu untuk melakukan sosialisasi ke sekolah-sekolah SMP dengan memberikan informasi-informasi kriteria peserta didik yang dibutuhkan di sekolah tersebut. Berikut sistematika yang penerimaan peserta didik baru (Suking, 2015), 1.) Membentuk panitia penerimaan peserta didik baru, 2.) Membuat persyaratan bagi pendaftar, 3.) Menyediakan formulir pendaftaran, 4.)Pengumuman terkait pendaftar, 5.) Menyediakan buku tentang pendaftaran peserta didik baru, 6.) Menginformasikan waktu pendaftaran dan banyaknya peserta didik baru yang diterima.

Sedangkan menurut Badruddin (2014), terdapat beberapa langkah-langkah dalam pelaksanaan penerimaan peserta didik baru, diantaranya yaitu:

- a. Membentuk Panitia. Panitia penerimaan peserta didik baru beranggotakan kepala sekolah dan beberapa guru yang diperlukan. Tugas dari panitia tersebut adalah untuk membentuk persyaratan-persyaratan peserta didik

baru, pengumuman terkait penerimaan, buku pendaftaran, formulir pendaftaran dan jumlah peserta didik yang dibutuhkan.

- b. Menentukan syarat pendaftaran peserta didik baru. Secara umum persyaratan sudah ditentukan oleh pemerintahan dari masing-masing wilayah yang berpedoman dengan kementerian pendidikan Indonesia.
- c. Menyediakan formulir pendaftaran.
- d. Pengumuman pendaftaran. Informasi yang harus ada di pengumuman tersebut adalah tentang informasi singkat lembaga tersebut, tanggal pendaftaran, biaya pendaftaran, waktu dan biaya seleksi.
- e. Menyediakan buku pendaftaran
- f. Menentukan waktu pendaftaran.

Sistem penerimaan peserta didik baru menjelaskan tentang cara penerimaan peserta didik baru. Ada dua macam dalam penerimaan peserta didik baru yaitu dengan cara seleksi atau promosi. Menurut Badruddin (2014), terdapat 3 (tiga) pembagian cara seleksi diantaranya yaitu berdasarkan nilai ujian nasional (UN), melalui prestasi bakat dan yang terakhir berdasarkan nilai ujian atau tes seleksi (tes psikolog, tes kesehatan, tes akademis dan tes jasmani). Kebijakan penerimaan peserta didik baru harus sesuai dengan persyaratan yang telah ditentukan. Kebijakan operasional dari penerimaan peserta didik baru berisikan tentang banyaknya peserta didik baru yang dibutuhkan. Penentuan banyaknya peserta yang dibutuhkan dapat dilihat dari kondisi sekolah. Kondisi sekolah dapat dilihat dengan, daya tampung kelas, sarana dan prasarana yang ada, kriteria peserta didik yang dibutuhkan dan tenaga pengajar yang cukup (Badruddin, 2014).

Kriteria penerimaan peserta didik baru digunakan untuk tolak ukur yang menentukan peserta didik bisa diterima atau tidak. Ada 3 (tiga) macam kriteria dalam menentukan peserta didik baru, diantaranya (Badruddin, 2014):

- a. Kriteria acuan norma (*norm criterion referenced*). pada kriteria ini penerimaan peserta didik baru dilakukan dengan cara keseluruhan prestasi pendaftar.
- b. Kriteria acuan patokan (*standart criterion referenced*). Penerimaan dengan kriteria ini yaitu, peserta didik diterima dengan patokan-patokan sebelumnya.
- c. Kriteria berdasarkan daya tampung sekolah. Kriteria ini dilakukan dengan cara meranking semua prestasi akademis peserta yang mengikuti seleksi dari yang tertinggi hingga terendah sampai memenuhi daya tampung sekolah.

## 2. 6 Pendidikan Dalam Prespektif Islam

Pendidikan islam adalah proses belajar mengajar dalam pengembangan diri meliputi akal, jasmani dan hati untuk membentuk kepribadian islami dengan menerapkan nilai-nilai agama. Hal tersebut mengakibatkan agar manusia dapat hidup sesuai kaidah islam dan sebagai khalifah. Tujuan dari pendidikan islam yaitu untuk mengembangkan pikiran setiap orang dan dapat mengatur tingkah laku dan perasaanya berdasarkan islam, dimana hal tersebut sejalan atau searah dengan tujuan diciptakan manusia. Pendidikan islam tidak hanya memberi ajaran-ajaran sebagai pengetahuan melainkan harus bisa membimbing peserta didik agar dapat menerapkan ajaran-ajaran tersebut dikehidupan sehari-hari (Tatang, 2011).

Menurut Miftahur Rohman (2018), adapun dua poin penting yang dijadikan sebagai tujuan pendidikan islam yaitu dapat mewujudkan kebahagiaan hidup di dunia dan akhirat. Hal ini menjelaskan bahwa pendidikan islam merupakan pendidikan yang ideal atau sempurna, karena pada pendidikan ini menjelaskan bahwa nantinya semua manusia akan binasa serta diberi akal ketika didunia. Di era jaman sekarang dapat diketahui bahwa perkembangan perilaku sosial di masyarakat yang tidak stabil, yang sulit ditebak akan perilakunya, sehingga perlu dilakukan pemahaman ulang bahwa tujuan pendidikan islam yang bersifat khusus dan aplikatif. Perlu diketahui bahwa Al-Quran dan hadist dijadikan pedoman di pendidikan islam, maka perlu dilakukan interpretasi ulang dengan memadukan nilai-nilai sosial kultural.

Fa taqabbalahaa rabbuhaa biqabulin hasaniw wa ambatahaa nabaatan hasanaw wa kaffalahaa zakariyyaa, kullama dakhala ‘alaiha zakariyyalmihraaba wajada ‘indahaa rizqaa, qaala yaa maryamu annaa laki haazaa, qaalat huwa min ‘indillaah, innallaaha yarzuqu may yasya’u bigairi hisaab.

Artinya:

Maka Dia (Allah) menerimanya dengan penerimaan yang baik, membesarkannya dengan pertumbuhan yang baik dan menyerahkan pemeliharannya kepada Zakariah. Setiap kali Zakaria masuk menemuinya di mihrab (kamar khusus ibadah), dia dapati makanan di sisinya. Dia berkata, “Wahai Maryam! dari mana ini engkau peroleh?” Dia (Maryam) menjawab, “itu dari Allah.”. Sesungguhnya Allah memberi rezeki kepada siapa yang Dia kehendaki tanpa perhitungan (QS. Al-Imran, 37).

Allah menerima Maryam sebagai nazar disebabkan permohonan ibunya. Allah meridhoinya untuk menjadi orang yang yang semata-mata beribadah dan

berkhidmat di Baitulmakdis walaupun Maryam masih kec0il dan hanya seorang perempuan. Padahal orang yang dikhususkan untuk berkhidmat di Baitulmakdis biasanya laki-laki yang akil baliq dan sanggup melaksanakan pengkhidmatan. Allah juga memelihara dan mendidiknya serta membesarkannya dengan sebaik-baiknya. Pendidikan yang diberikan Allah kepada Maryam, meliputi pendidikan rohani dan jasmani. Maka dia menjadi orang yang berbadan sehat dan kuat serta berbudi baik, bersih rohani dan jasmaninya. Allah telah pula menjadikan Nabi Zakariah sebagai pengasuh dan pelindungnya.



## BAB III

### METODOLOGI PENELITIAN

#### 3.1 Pendekatan Penelitian

Pendekatan yang digunakan pada penelitian ini adalah kuantitatif dan studi literatur. Pendekatan kuantitatif adalah dimana data yang digunakan dalam penelitian berupa angka. Pendekatan studi literatur adalah pendekatan dengan mengkaji dari buku, jurnal dan artikel yang berkaitan dengan penelitian yang digunakan sebagai sumber.

#### 3.2 Jenis dan Sumber Data

Pada penelitian ini data yang digunakan adalah data sekunder. Pengambilan data dilakukan pada hari Senin tanggal 25 Februari 2020 di Madrasah Bertaraf International Amanatul Ummah. Data yang diperoleh dari hasil tes tulis ujian masuk gelombang satu tahun 2020 dengan jumlah 390 data di MBI (Madrasah Bertaraf International) Amanatul Ummah Pacet Mojokerto.

#### 3.3 Identifikasi Variabel

Terdapat dua variabel yang digunakan pada penelitian ini yaitu variabel respon dan variabel prediktor. Variabel respon dan variabel prediktor yang digunakan pada penelitian ini sebagai berikut:

##### 3.3.1 Variabel Respon

Penentuan variabel respon (Y) yang terdiri dari diterima atau tidak diterima seleksi ujian masuk MBI Amanatul Ummah Pacet. Pada variabel respon ini menggunakan skala nominal yaitu:

Y=1, jika responden diterima seleksi ujian masuk

Y=2, jika responden tidak diterima seleksi ujian masuk

### 3.3.2 Variabel Prediktor

Variabel prediktor yang digunakan pada penelitian ini adalah variabel yang dianggap berpengaruh terhadap penentuan hasil seleksi ujian masuk MBI Amanatul Ummah. Variabel-variabel tersebut sebagai berikut:

1. Jenis kelamin

Variabel prediktor jenis kelamin sebagai variabel prediktor pertama dengan dilambangkan X1. Variabel ini dibedakan menjadi dua kategori, diantaranya:

X1=1, jika responden berjenis kelamin laki-laki

X1=2, jika responden berjenis kelamin perempuan

2. Asal sekolah

Variabel prediktor asal sekolah sebagai variabel prediktor kedua dilambangkan X2. Variabel ini dibedakan menjadi dua kategori yaitu:

X2=1, jika responden berasal dari SMP berbasis non islam

X2=2, jika responden berasal dari SMP berbasis islam

3. Nilai tes IPA

Variabel nilai tes IPA sebagai variabel prediktor ketiga dilambangkan dengan X3. Variabel ini menggunakan skala rasio 0-100.

4. Nilai tes Matematika

Variabel nilai tes Matematika sebagai variabel prediktor keempat dilambangkan dengan X4. Variabel ini menggunakan skala rasio 0-100.

5. Nilai tes B.Inggris.

Variabel nilai tes Bahasa Inggris sebagai variabel kelima variabel prediktor kelima dilambangkan dengan X5. Variabel ini menggunakan skala rasio 0-100.

6. Nilai tes IQ

Variabel nilai tes IQ sebagai variabel keenam variabel prediktor keenam dilambangkan dengan X6. Variabel ini menggunakan skala kontinu.

### 3.4 Langkah-langkah Penelitian

Pada penelitian ini menggunakan Bootsrap Aggregating (*Bagging*) Classification and Regression Trees (CART) dengan bantuan menggunakan *Software R* dan *minitab 19* dengan tahap-tahap sebagai berikut ini:

1. Mengumpulkan data sekunder hasil tes seleksi ujian masuk MBI Amanatul Ummah Pacet gelombang 1 (satu) tahun 2020.
2. Menentukan variabel respon dan variabel prediktor yang dibutuhkan pada penelitian.
3. Melakukan analisis deskriptif data sekunder hasil tes seleksi ujian masuk MBI Amanatul Ummah Pacet gelombang 1 (satu) tahun 2020.
4. Membagi data menjadi dua yaitu, data learning dan data testing dengan perbandingan 90%:10%, 85%:15%, 80% : 20% dan 75%:25% secara acak.
5. Pembentukan pohon klasifikasi algoritma CART dengan beberapa tahapan dibawah ini:

- a. Memilih pemilah dari variabel prediktor dengan memilih nilai *goodness of split* yang paling besar akan menjadi *parent node*.

Dengan persamaan indeks gini :

$$i(t) = \sum_{j, i \neq j} p(i|t)p(j|t), i \neq j$$

Persamaan *goodness of split*:

$$\phi(s, t) = \Delta i(s, t) = i(t) - P_L i(t_L) - P_R i(t_R)$$

- b. Penentuan *terminal node*, dikatakan *terminal node* jika suatu simpul memiliki  $n < 5$ .
- c. Penandaan label kelas dapat dilihat dari semua simpul terminal dengan melihat nilai  $N(J|t)$ . Untuk menentukan kelasnya dipilih  $P(J|t)$  yang terbesar.

$$p(j_0|t) = \max p(j|t) = \max \frac{N_j(t)}{N(t)}$$

- d. Pemangkasan pohon
6. Prediksi data menggunakan data *testing* dengan algoritma pohon CART yang telah terbentuk dengan menggunakan data *testing*.
7. Proses *Bootstrap Aggregating (Bagging)*
- a. Mengulangi proses 5 sampai proses 7 dengan replikasi sebanyak 100 kali, sehingga didapatkan 100 pohon CART. Kemudian dilakukan prediksi pada masing-masing pohon CART.
- b. Proses *aggregating* menggunakan *majority vote* untuk mendapatkan pendugaan gabungan.
- c. Ilustrasi *Bootstrap Aggregating*

Ilustrasi *Bagging* akan dilakukan dengan banyaknya  $b=3$  dan data *learning* yang digunakan sebanyak 8 data yang tertera di tabel 3.1. Pada ilustrasi ini melibatkan variabel respon dengan skala kategorik yaitu diterima dan tidak di terima dan variabel prediktor sebanyak dua variabel yaitu jenis kelamin ( $x_1$ ) dan asal sekolah ( $x_2$ ) berskala kategorik dan nilai matematika ( $x_3$ ), nilai ipa ( $x_4$ ), nilai B.Inggris ( $x_5$ ) serta nilai IQ ( $x_6$ ) berskala rasio.

Tabel 3. 1 Data *Learning*

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	3	60.0	48.6	62.0	125	DITERIMA
1	3	80.0	62.9	70.0	106	DITERIMA
1	3	80.0	54.3	42.0	120	DITERIMA
2	3	50.0	42.9	60.0	112	TIDAK DITERIMA
2	3	45.0	37.1	70.0	125	TIDAK DITERIMA
2	3	75.0	28.6	58.0	112	DITERIMA
1	4	55.0	31.4	48.0	106	TIDAK DITERIMA
2	3	80.0	40.0	60.0	98	DITERIMA

Tabel 3. 2 Data *Testing*

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	3	20.0	37.1	22.0	73	TIDAK DITERIMA
1	3	30.0	28.6	58.0	96	TIDAK DITERIMA
2	4	40.0	48.6	34.0	98	TIDAK DITERIMA
2	3	55.0	28.6	46.0	112	TIDAK DITERIMA
1	4	75.0	54.3	66.0	121	DITERIMA

Data *learning* pada Tabel 3.1 akan dilakukan pengulangan sebanyak 3 kali sehingga diperoleh 3 kelompok data *learning*. Setelah didapatkan data sampelnya,

selanjutnya dilakukan pembentukan pohon CART. Sehingga didapatkan pohon optimal dari ketiga data sampel, sebagai berikut:

Tabel 3. 3 Data Sampel *Learning* Pertama

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	3	60.0	48.6	62.0	125	DITERIMA
2	3	75.0	28.6	58.0	112	DITERIMA
1	3	60.0	48.6	62.0	125	DITERIMA
2	3	80.0	40.0	60.0	98	DITERIMA
2	3	45.0	37.1	70.0	125	TIDAK DITERIMA
2	3	45.0	37.1	70.0	125	TIDAK DITERIMA
1	4	55.0	31.4	48.0	106	TIDAK DITERIMA
1	3	80.0	62.9	70.0	106	DITERIMA

Tabel 3. 4 Data Sampel *Learning* Kedua

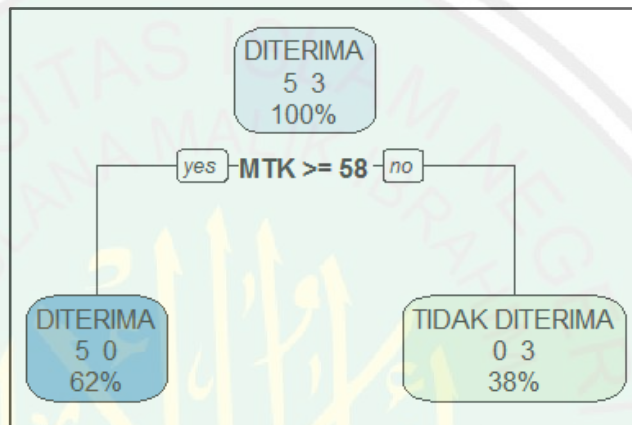
JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
2	3	45.0	37.1	70.0	125	TIDAK DITERIMA
1	3	60.0	48.6	62.0	125	DITERIMA
1	3	80.0	62.9	70.0	106	DITERIMA
1	3	80.0	54.3	42.0	120	DITERIMA
2	3	75.0	28.6	58.0	112	DITERIMA
1	3	60.0	48.6	62.0	125	DITERIMA
1	3	80.0	62.9	70.0	106	DITERIMA
1	3	80.0	54.3	42.0	120	DITERIMA

Tabel 3. 5 Data Sampel *Learning* Ketiga

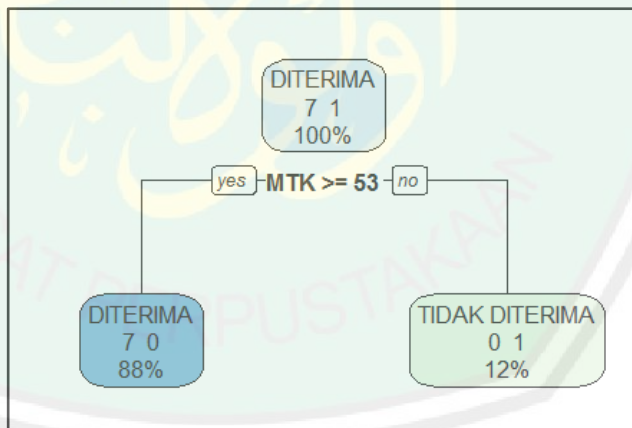
JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	3	80.0	62.9	70.0	106	DITERIMA
1	4	55.0	31.4	48.0	106	TIDAK DITERIMA
2	3	45.0	37.1	70.0	125	TIDAK DITERIMA
2	3	45.0	37.1	70.0	125	TIDAK DITERIMA
2	3	75.0	28.6	58.0	112	DITERIMA

Lanjutan Tabel 3.5 Data Sampel *Learning* Ketiga

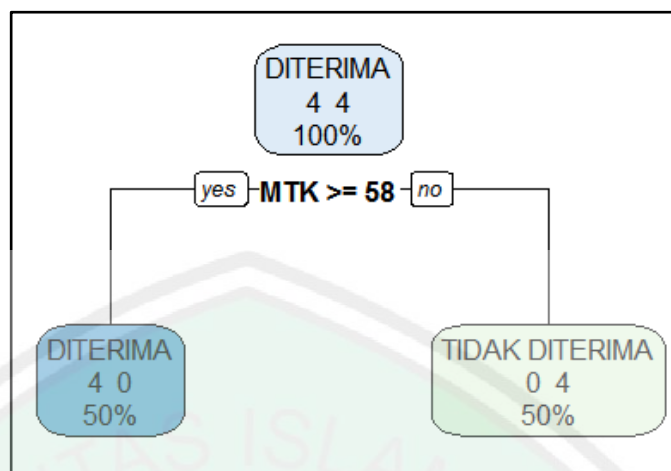
JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
2	3	45.0	37.1	70.0	125	TIDAK DITERIMA
2	3	75.0	28.6	58.0	112	DITERIMA
1	3	60.0	48.6	62.0	125	DITERIMA



Gambar 3. 1 Pohon Optimal Pertama



Gambar 3. 2 Pohon Optimal Kedua



Gambar 3. 3 Pohon Optimal Ketiga

Ketiga pohon klasifikasi yang telah didapatkan memiliki bentuk pohon yang berbeda-beda, hal ini menunjukkan bahwa dari ketiga data sampel berbeda dan dilakukan pengambilan secara acak dengan pengembalian. Setelah itu akan dilakukan prediksi dengan data testing dari setiap pohon klasifikasi untuk didapatkan ketepatan klasifikasi menggunakan *majority vote* (suara terbanyak), tahap ini dinamakan *aggregating*. Data testing yang digunakan yaitu pada tabel 3.2.

Tabel 3. 6 Hasil Klasifikasi dari Data *Testing* Pohon Pertama

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	3	20.0	37.1	22.0	73	TIDAK DITERIMA
1	3	30.0	28.6	58.0	96	TIDAK DITERIMA
2	4	40.0	48.6	34.0	98	TIDAK DITERIMA
2	3	55.0	28.6	46.0	112	TIDAK DITERIMA
1	4	75.0	54.3	66.0	121	DITERIMA

Tabel 3. 7 Hasil Klasifikasi dari Data *Testing* Pohon Kedua

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	3	20.0	37.1	22.0	73	TIDAK DITERIMA
1	3	30.0	28.6	58.0	96	TIDAK DITERIMA
2	4	40.0	48.6	34.0	98	TIDAK DITERIMA
2	3	55.0	28.6	46.0	112	DITERIMA
1	4	75.0	54.3	66.0	121	DITERIMA

Tabel 3. 8 Hasil Klasifikasi dari Data *Testing* Pohon Ketiga

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	3	20.0	37.1	22.0	73	TIDAK DITERIMA
1	3	30.0	28.6	58.0	96	TIDAK DITERIMA
2	4	40.0	48.6	34.0	98	TIDAK DITERIMA
2	3	55.0	28.6	46.0	112	TIDAK DITERIMA
1	4	75.0	54.3	66.0	121	DITERIMA

Pada tahap *aggregating* dilakukan dengan aturan *majority vote* yaitu pengambilan suara terbanyak dari hasil klasifikasi, misalkan pada data pertama pohon pertama hasil klasifikasinya tidak diterima, pohon kedua hasil klasifikasinya tidak diterima dan pada pohon ketiga hasil klasifikasinya tidak diterima sehingga *majority vote* dari data pertama yaitu tidak diterima, karena hasil klasifikasi tidak diterima lebih banyak dari pada hasil klasifikasi. Berikut tabel kesimpulan *majority vote* dari semua data *testing*:

Tabel 3. 9 Kesimpulan *Majority Vote Data Learning*

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	3	20.0	37.1	22.0	73	TIDAK DITERIMA
1	3	30.0	28.6	58.0	96	TIDAK DITERIMA
2	4	40.0	48.6	34.0	98	TIDAK DITERIMA
2	3	55.0	28.6	46.0	112	TIDAK DITERIMA
1	4	75.0	54.3	66.0	121	DITERIMA

Berdasarkan keempat calon peserta didik baru dapat diketahui hasil klasifikasi diterima sebanyak 1 dan tidak diterima sebanyak 4 calon peserta didik baru. Setelah didapatkan hasil klasifikasi dari data *testing*, selanjutnya akan dilakukan perhitungan ketepatan klasifikasi.

- Mengukur ketepatan klasifikasi pohon *Bagging* CART dengan menggunakan *10-fold cross validation* untuk menghitung nilai APER *Accuracy rate*, *sensitivity* dan *specificity*.

$$APER = \frac{\text{Jumlah prediksi salah}}{\text{Jumlah total prediksi}}$$

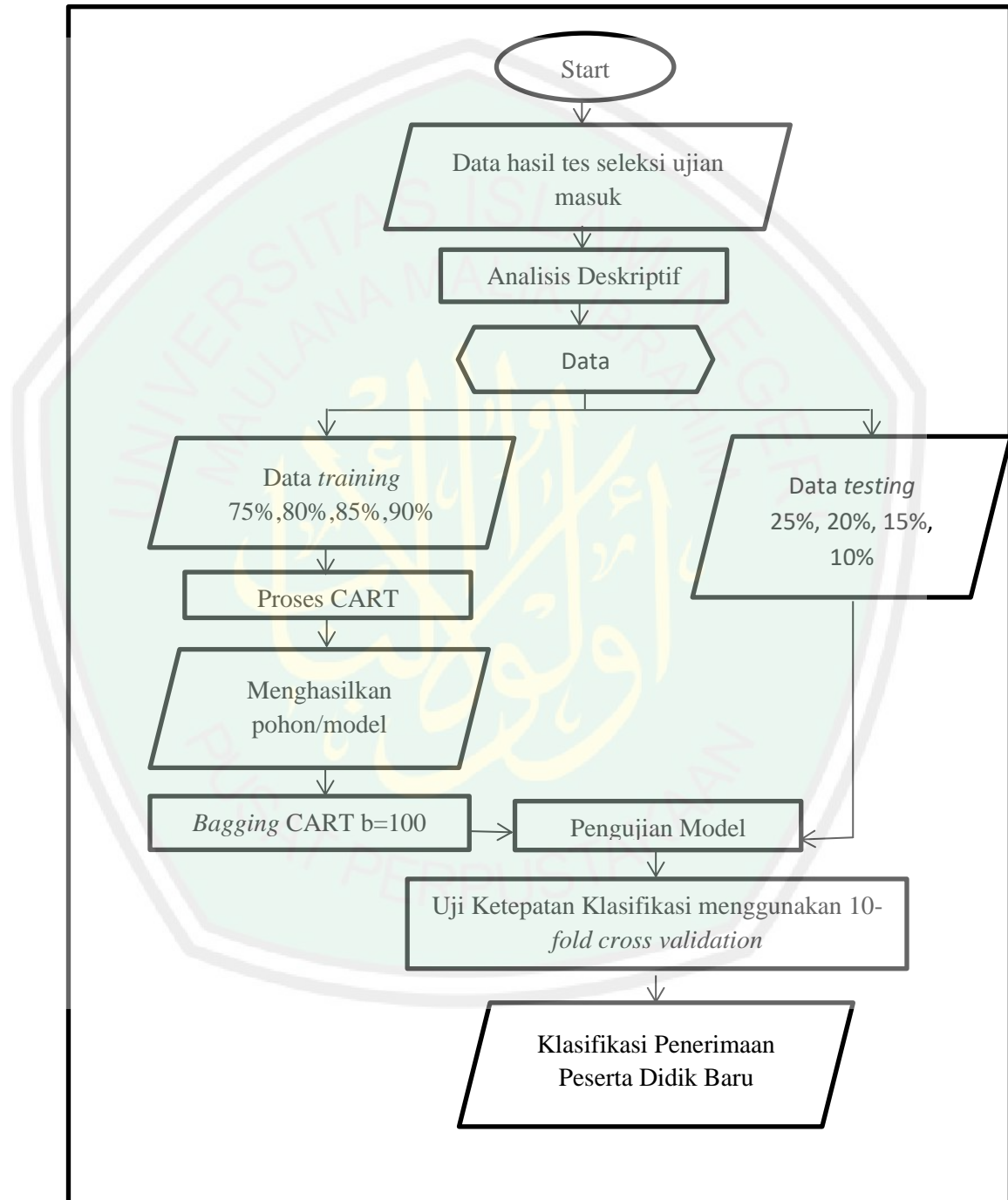
$$Accuracy = \frac{n_{11} + n_{22}}{N}$$

$$Sensitivity = \frac{n_{11}}{N_1}$$

$$Specificity = \frac{n_{22}}{N_2}$$

### 3.5 Flow Chart

Langkah-langkah penelitian yang dilakukan dapat digambarkan dengan diagram alir sebagai berikut:



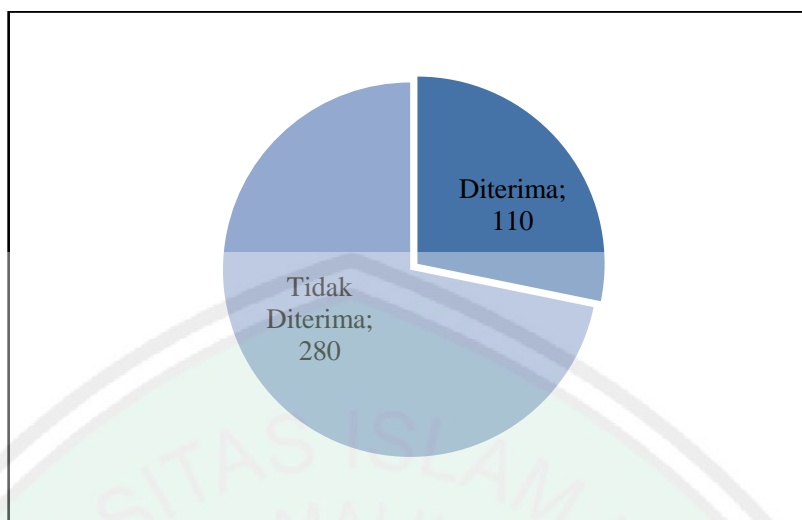
Gambar 3.4 Flow Chart Bagging CART

## BAB IV

### ANALISIS DATA DAN PEMBAHASAN

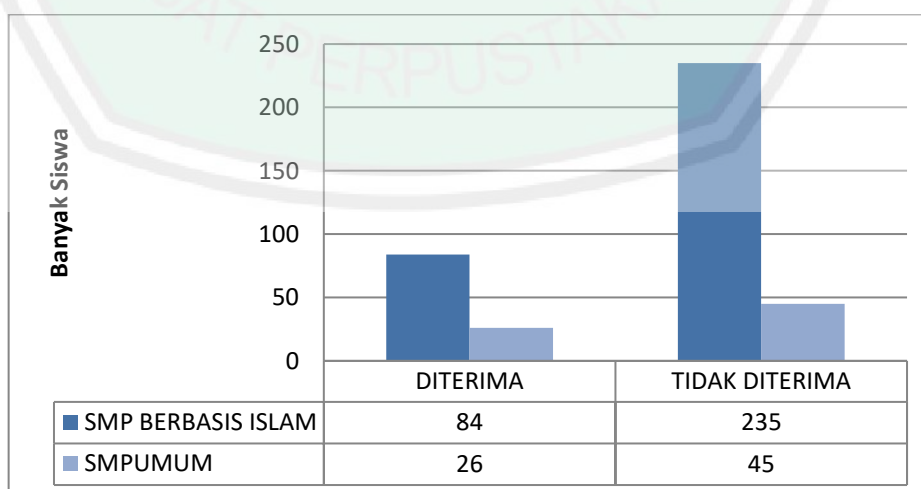
#### 4.1 Analisis Data

Jumlah banyaknya calon peserta didik baru gelombang satu tahun ajaran 2020-2021 di Madrasah Bertaraf International Amanatul Ummah pacet sebanyak 390 peserta. Berdasarkan jumlah tersebut sebanyak 110 calon peserta didik baru dinyatakan telah diterima seleksi penerimaan peserta didik baru dan sebanyak 280 dinyatakan tidak diterima seleksi penerimaan peserta didik baru. Terima dan tidak diterimanya calon peserta didik baru menjadi variabel respon pada penelitian ini. Terdapat indikator-indikator yang mempengaruhi dalam penentuan diterima atau tidak diterimanya dalam seleksi penerimaan peserta didik baru gelombang 1 di Madrasah Bertaraf International Amanatul Ummah Pacet, indikator tersebut diantaranya yaitu jenis kelamin (perempuan atau laki-laki), asal sekolah (SMP berbasis islam atau SMP umum), nilai tes IPA, nilai tes matematika, nilai tes bahasa Inggris dan nilai tes IQ. Indikator-indikator tersebut dapat membantu untuk menentukan klasifikasi diterima dan tidak diterima pada penerimaan peserta didik baru. Perbandingan banyaknya calon peserta didik baru yang diterima dan tidak diterima akan digambarkan dengan diagram lingkaran seperti gambar 4.1:



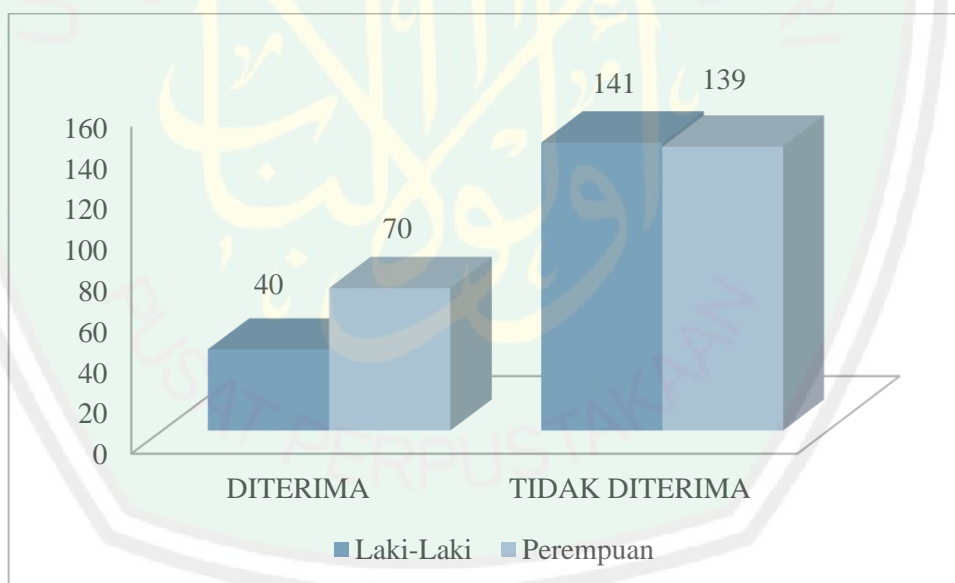
Gambar 4. 1 Jumlah Peserta Didik Baru Berdasarkan Hasil Seleksi

Berdasarkan diagram 4.1, didapatkan informasi dari 390 calon peserta didik baru yang mendaftar 110 diterima dan 280 tidak diterima. Indikator pertama yang mempengaruhi seleksi penerimaan peserta didik baru di Madrasah Bertaraf International Amanatul Ummah Pacet yaitu asal sekolah dari setiap calon peserta didik baru. Didapatkan informasi tentang klasifikasi penerimaan santri baru berdasarkan asal sekolah yang digambarkan dengan diagram batang seperti gambar 4.2:



Gambar 4. 2 Perbandingan Diterima dan Tidak Diterima Berdasarkan Asal Sekolah

Gambar 4.2 memberikan informasi bahwa banyaknya calon peserta santri baru yang diterima berdasarkan SMP berbasis Islam sebanyak 84 peserta sedangkan yang tidak diterima sebanyak 235 peserta. Setiap calon peserta didik baru yang berasal dari SMP berbasis Islam memiliki peluang untuk bisa diterima sebesar 0,263. Peserta santri baru yang berasal dari sekolah SMP umum yang diterima sebanyak 26 peserta dan yang tidak diterima berjumlah 45 peserta, setiap peserta calon peserta didik baru yang berasal dari SMP umum memiliki peluang untuk diterima sebesar 0,266. Jadi dapat disimpulkan dari diagram pada gambar 4.2 bahwa peluang diterimanya peserta didik baru yang berasal dari SMP umum lebih besar dari pada dari SMP berbasis Islam.



Gambar 4. 3 Perbandingan Diterima dan Tidak Diterima Indikator Jenis Kelamin

Gambar 4.3 menjelaskan tentang perbandingan hasil seleksi penerimaan peserta didik baru dengan indikator jenis kelamin. Didapatkan informasi bahwa jenis kelamin perempuan lebih banyak dari laki-laki yang diterima menjadi peserta didik baru di Madrasah Bertaraf Internasional Amanatul Ummah Pacet

dengan selisih 30 lebih banyak jenis kelamin perempuan dari pada laki-laki. Sedangkan untuk perbandingan hasil seleksi tidak diterima berbanding terbalik dengan perbandingan hasil seleksi diterima, yaitu sebanyak 141 peserta didik baru berjenis kelamin laki-laki dan 139 berjenis kelamin perempuan. Sehingga dapat diidentifikasi bahwa jenis kelamin perempuan mendominasi dalam penerimaan peserta didik baru.

Indikator yang ketiga yaitu nilai matematika, dimana data nilai matematika berskala rasio dari 1 hingga 100. Berikut ini deskriptif karakteristiknya berdasarkan masing-masing hasil seleksi diantaranya diterima dan tidak diterima seperti tabel 4.1:

Tabel 4. 1 Deskriptif Karakteristik Diterima dan Tidak Diterima Indikator Nilai Matematika

Variabel	Rata-rata	Standart Deviasi	Minimum	Maksimum
<b>Diterima</b>	69,73	12,56	35,00	95,00
<b>Tidak Diterima</b>	36,76	14,70	5,00	75,00

Tabel 4.1 didapatkan informasi mengenai indikator nilai matematika terhadap hasil seleksi penerimaan peserta didik baru yaitu diterima atau tidak diterima. Untuk hasil seleksi diterima rata-rata nilai matematika sebesar 69,73 dengan standar deviasi 12,56. Dapat diketahui juga nilai matematika dengan rentan antara 35,00 sampai dengan 95,00 mempunyai peluang untuk bisa diterima. Tetapi dengan memiliki nilai minimum belum tentu diterima kecuali didukung dengan indikator-indikator lainnya.

Indikator hasil seleksi tidak diterima ditunjukkan dengan rata-rata nilai matematika sebesar 36,76 dari keseluruhan peserta santri baru yang tidak diterima. Sedangkan nilai standar deviasinya sebesar 14,70 yang menunjukkan bahwa nilai matematika sangat bervariasi. Nilai matematika dengan rentan 5,00 hingga 75,00 dinyatakan tidak diterima ada seleksi. Ketika terdapat peserta yang memiliki nilai maksimum di indikator ini dinyatakan tidak diterima padahal nilai 75,00 bisa masuk kategori peserta diterima, hal itu diakibatkan karena indikator-indikator lainnya tidak sesuai dengan kriteria.

Indikator selanjutnya yaitu nilai mata pelajaran IPA dengan skala rasio dari 1 hingga 100. Tabel 4.2 merupakan deskriptif karakteristik dari indikator nilai IPA, sebagai berikut:

Tabel 4. 2 Deskriptif Karakteristik Diterima dan Tidak Diterima Indikator Nilai IPA

Variabel	Rata-rata	Standart Deviasi	Minimum	Maksimum
Diterima	44,934	9,625	25,7	74,3
Tidak Diterima	32,110	7,704	5,7	51,4

Tabel 4.2 didapatkan beberapa informasi tentang deskriptif karakteristik hasil seleksi penerimaan peserta didik baru yaitu diterima atau tidak diterima. Nilai rata-rata IPA sebesar 44,934 untuk hasil yang diterima dengan standar deviasi 9,625 dari keseluruhan peserta yang mengikuti seleksi penerimaan santri baru. Minimum nilai IPA yang harus didapatkan agar dapat diterima seleksi adalah 25,7 dan nilai maksimum dari keseluruhan peserta adalah 74,3.

Tidak diterimanya peserta santri baru dapat dilihat dari nilai IPA, jika nilainya masuk dalam rentan antara 5,7 hingga 51,4 maka ada peluang tidak

diterima. Karena nilai rata-rata nilai IPA dari keseluruhan peserta yang tidak diterima di gelombang satu sebesar 32,11 dan standar deviasinya sebesar 7,7. Oleh karena itu peserta harus mendapatkan nilai dalam rentan minimum dan maksimal pada kategori diterima di indikator nilai IPA.

Indikator selanjutnya adalah nilai bahasa Inggris dengan skala rasio antara 1 hingga 100. Berikut ini deskriptif karakteristik dari nilai bahasa Inggris terhadap hasil seleksi, diantaranya:

Tabel 4. 3 Deskriptif Karakteristik Diterima dan Tidak Diterima Indikator Nilai Bahasa Inggris

Variabel	Rata-rata	Standart Deviasi	Minimum	Maksimum
Diterima	65,964	10,053	40	90
Tidak Diterima	46,693	12, 877	14	78

Tabel 4.3 menjelaskan tentang deskriptif karakteristik indikator nilai bahasa Inggris terhadap hasil seleksi yaitu diterima atau tidak diterima. Peserta dapat dikategorikan diterima seleksi jika nilai tes bahasa Inggris dalam rentan nilai minimum sebesar 40 dan nilai maksimum sebesar 90. Dikarenakan nilai rata-rata dari semua peserta yang diterima adalah 65,964 dan standar deviasinya sebesar 10,053. Maka dianjurkan untuk peserta yang ingin diterima untuk mendapatkan nilai bahasa inggris di atas rata-rata nilai diterima atau di antara rentan nilai minimum dan maksimal. Peserta bisa dikatakan tidak diterima seleksi jika nilai peserta mendapatkan nilai bahasa inggris dibawah rata-rata dari semua siswa yang tidak diterima sebesar 46,693 dan standar deviasinya 12, 877.

Indikator selanjutnya adalah nilai IQ dengan skala rasio. Berikut ini deskriptif karakteristik dari nilai bahasa Inggris terhadap hasil seleksi, diantaranya:

Tabel 4. 4 Deskriptif Karakteristik Diterima dan Tidak Diterima Indikator Nilai Tes IQ

Variabel	Rata-rata	Standart Deviasi	Minimum	Maksimum
Diterima	116,63	15,54	91	142
Tidak Diterima	103,75	14,9	68	136

Tabel 4.4 didapatkan informasi mengenai indikator nilai tes IQ terhadap hasil seleksi penerimaan peserta didik baru yaitu diterima atau tidak diterima. Untuk hasil seleksi diterima rata-rata nilai tes IQ sebesar 116,63 dengan standar deviasi 15,54. Dapat diketahui juga nilai IQ dengan rentan antara 91 sampai dengan 142 mempunyai peluang untuk bisa diterima. Tetapi dengan memiliki nilai minimum belum tentu diterima kecuali didukung dengan indikator-indikator lainnya.

Indikator hasil seleksi tidak diterima ditunjukkan dengan rata-rata nilai tes IQ sebesar 103,75 dari keseluruhan peserta didik baru yang tidak diterima. Sedangkan nilai standar deviasinya sebesar 14,9 yang menunjukkan bahwa nilai IQ sangat bervariasi. Nilai tes IQ dengan rentan 68 hingga 136 dinyatakan tidak diterima pada seleksi. Ketika terdapat peserta yang memiliki nilai maksimum di indikator ini dinyatakan tidak diterima padahal nilai tes IQ sebesar 136 bisa masuk kategori peserta diterima, hal itu diakibatkan karena indikator-indikator lainnya tidak sesuai dengan kriteria.

## 4.2 Analisis CART (*Classification and Regression Trees*)

Analisis CART yang dilakukan untuk mengetahui klasifikasi penerimaan peserta didik baru dengan studi kasus di Madrasah Bertaraf International Amanatul Ummah Pacet dengan 2 kategori yaitu diterima seleksi dan tidak diterima seleksi. Langkah awal pada penelitian ini adalah membagi data menjadi 2 bagian yaitu data *learning* dan data *testing* dengan beberapa macam presentase dan untuk menguji validasi menggunakan metode *V-fold cross validation*. Nilai *V* yang digunakan di penelitian ini adalah 10 sehingga menjadi 10 *fold cross validation* terhadap sebanyak 390 data penerimaan peserta didik baru tahun 2021-2022 gelombang 1 yang dilaksanakan pada tanggal 22 februari 2020 di MBI Amanatul Ummah Pacet.

### 4.2.1 Penentuan Presentase Pembagian Data Learning dan Data Testing

Presentase pembagian data *learning* dan data *testing* pada penelitian ini menggunakan 4 kombinasi presentase, yaitu: (75%:25%), (80%,20%), (85%,15%), dan (90%, 10%). Dilakukan pembentukan pohon optimal disetiap kombinasi presentase pembagian data di lampiran 3, setelah itu dihitung ketepatan klasifikasi di tiap-tiap klasifikasi presentase pembagian data. Hasil ketepatan klasifikasi tersebut dilakukan perbandingan, kombinasi akan dipilih jika nilai ketepatan klasifikasinya seimbang antara data *learning* dan data *testing*. Data *learning* digunakan untuk pembentukan pohon klasifikasi sedangkan data *testing* digunakan untuk mengukur kebaikan/kelayakan model.

Tabel 4. 5 Ketepatan Klasifikasi Presentase Data

No	Presentasi Data(%)		Ketepatan Klasifikasi	
	Learning	Testing	Learning	Testing
1	75	25	0.9488055	0.8762887
2	80	20	0.9391026	0.9230769
3*	85	15	0.9277108	0.8965517
4	90	10	0.5726496	0.6153846

\*Pembagian kombinasi data terbaik

Tabel 4.5 dapat menunjukkan bahwa nilai ketepatan klasifikasi yang seimbang antara data *learning* dan data *testing* diperoleh pada presentase data sebesar 85% data *learning* dan 15 % data *testing*. Nilai yang diperoleh dari ketepatan klasifikasi sebesar 89,66% untuk data *testing* dan untuk data *learning* sebesar 92,77%. Hasil kombinasi dari kelima kombinasi presentase pembagian data yang dibandingkan pertama adalah nilai dari ketepatan klasifikasi data *testing*, karena data *testing* berguna untuk melihat kelayakan pohon klasifikasi yang telah dibentuk oleh data *learning*. Proses selanjutnya yaitu analisis CART (*classification and regression trees*) dengan menggunakan data 85% data *learning* dan 15% data *testing*.

#### 4.2.2 Pembentukan Pohon Klasifikasi

Pembentukan pohon klasifikasi diawali dengan pemilihan variabel terpenting yang berperan sebagai pemilah terbaik dengan dihitung nilai setiap variabel prediktor (nilai *threshold*) dan *goodness of split*. Nilai *threshold* dan variabel pemilah diperoleh dari kemungkinan pemilah setiap variabel prediktor dengan skala data yang berbeda. Berikut ini perhitungan untuk menentukan kemungkinan pemilah setiap variabel prediktor ditunjukkan pada tabel 4.6:

Tabel 4. 6 Banyaknya Pemilihan Pemilah Variabel Prediktor

Variabel	Nama Variabel	Skala Data	Banyaknya Kategori	Kemungkinan Pemilah
X1	Jenis Kelamin	Nominal	2	$2^{2-1} - 1 = 1$ pemilah
X2	Asal Sekolah	Nominal	2	$2^{2-1} - 1 = 1$ pemilah
X3	Nilai Tes IPA	Rasio	1-100	$100-1=99$ pemilah
X4	Nilai Tes MTK	Rasio	1-100	$100-1=99$ pemilah
X5	Nilai Tes B.Ingggris	Rasio	1-100	$100-1=99$ pemilah
X6	Nilai Tes IQ	Rasio	1-130	$130-1=129$ pemilah

Tabel 4.6 menunjukkan banyaknya kemungkinan pemilah setiap variabel prediktor. Variabel X1 (Jenis Kelamin) terdapat satu kemungkinan pemilah. Variabel X2 (Asal Sekolah) terdapat satu kemungkinan pemilah. Jika dilihat dari X3 (Nilai Tes IPA), X4 (Nilai Tes MTK) dan X5 (Nilai Tes B.Ingggris) memiliki data berskala rasio 1 hingga 100 sehingga terdapat kemungkinan 99 pemilah di masing-masing variabel prediktor tersebut. Sedangkan untuk nilai tes IQ (X6) merupakan data berskala rasio 1 hingga 130 sehingga didapatkan banyaknya kemungkinan pemilah yaitu 129 pemilah.

Kemungkinan pemilah setiap variabel prediktor, selanjutnya menentukan pemilah terbaik menggunakan kriteria *goodness of split*. Pada penelitian ini perhitungan *goodness of split* menggunakan Indeks Gini, dimana Indeks Gini merupakan perhitungan yang sederhana dan efisien. Perhitungan dengan Indeks Gini dilakukan dengan pemilah simpul yang berfokus dengan simpul kanan dan simpul kiri. Pemilah yang terpilih yaitu pemilah yang memiliki nilai *threshold* dan nilai *goodness of split* yang terbesar. Pemilah yang terpilih biasanya disebut dengan *parent node* merupakan variabel yang terpenting pada pembuatan pohon

klasifikasi. Banyaknya kontribusi suatu variabel pemilah dapat dilihat dengan nilai skor masing-masing variabel.

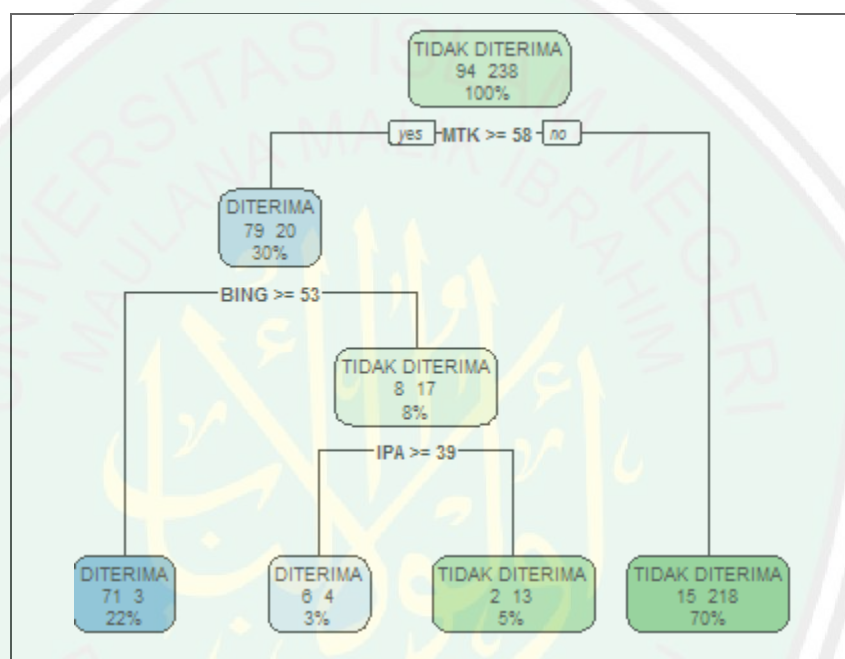
Tabel 4. 7 Skor Variabel Prediktor Pohon Maksimal

Variabel	Skor(%)
MTK (X4)	56
BING (X5)	18
IPA (X3)	15
Nilai.IQ (X6)	11
AS (Asal Sekolah) (X2)	0
JK (Jenis Kelamin) (X1)	0

Tabel 4.7 dapat diketahui bahwa variabel prediktor yang paling banyak berkontribusi yaitu nilai MTK dengan nilai skor 56%. Skor ini menunjukkan bahwa sebanyak 186 data yang berkontribusi untuk pembentukan pohon pemilah nilai matematika. Nilai MTK dapat dikatakan menjadi *parent node* pada pembentukan pohon klasifikasi maksimal. Semua variabel prediktor mempunyai kontribusi pohon klasifikasi berdasarkan hasil perhitungan skor diatas. Variabel prediktor yang memberikan kontribusi paling sedikit yaitu AS (Asal Sekolah) dan JK (Jenis Kelamin) dengan nilai skor yang sama yaitu sebesar 0%. Skor ini menunjukkan bahwa sebanyak 0 data yang berkontribusi pada pembentukan pohon pemilah asal sekolah dan jenis kelamin. Variabel prediktor nilai MTK dapat menurunkan keheterogenan terbesar sehingga simpul yang dihasilkan lebih homogen.

Proses pemilihan pemilah di pembentukan pohon klasifikasi selanjutnya yaitu pembentukan model pohon klasifikasi maksimal. Sifat dari pohon klasifikasi maksimal adalah pohon yang terbentuk dengan memiliki simpul terminal

terbanyak dengan kedalaman tertinggi. Pembentukan model pohon klasifikasi maksimal dengan menggunakan indeks gini diperoleh simpul terminal sebanyak 4 simpul dan tingkat kedalam pohon yang diperoleh sebesar 4. Ilustrasi pohon maksimal yang diperoleh dengan menggunakan pemilah indeks gini seperti gambar 4.4:

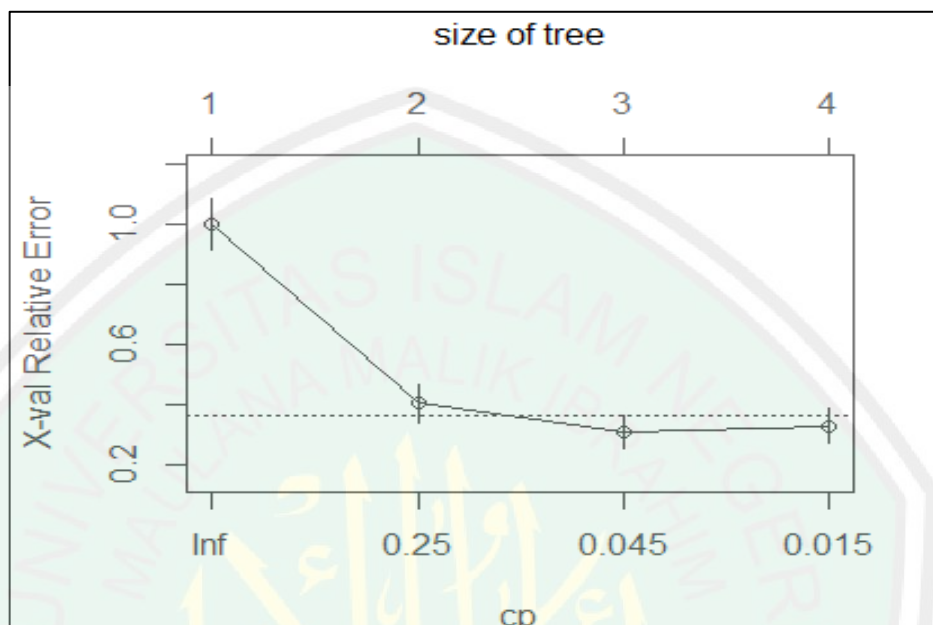


Gambar 4. 4 Pohon Klasifikasi Maksimal CART

#### 4.2.3 Pemangkasan Pohon Klasifikasi Maksimal

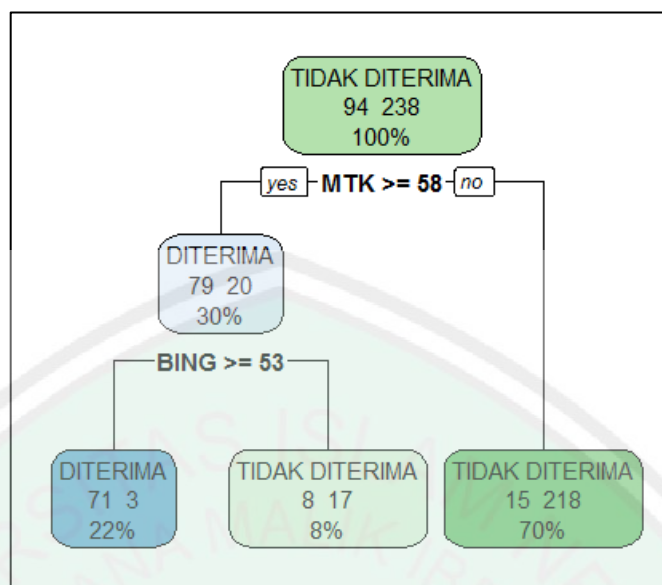
Setelah di dapatkan pohon maksimal klasifikasi, selanjutnya akan dilakukan pemangkasan pohon sehingga di dapatkan pohon yang optimal. Pemangkasan pohon dilakukan agar terhindar dari *overfitting* (nilai akurasi yang didapatkan tidak sesuai atau melebihi nilai kenyataannya) dan *underfitting* (nilai akurasi yang didapatkan kurang dari nilai kenyataannya). *Overfitting* dapat terjadi karena pohon yang terlalu besar menyebabkan nilai kompleksitas terlalu tinggi dikarenakan struktur data yang terlalu kompleks sehingga dilakukan

pemangkasan. Pemangkasan pohon dapat dilihat dari nilai *X-val relative error* yang paling kecil, dan diilustrasikan pada gambar plot di gambar 4.5:



Gambar 4. 5 Plot CP (Complexity Parameter)

Gambar 4.5 menjelaskan tentang plot nilai *X-val relative error* dan nilai *CP(Complexity Parameter)*. Sumbu Y menjelaskan nilai *X-val relative error* dan sumbu x menjelaskan nilai *CP(Complexity Parameter)*. Dilihat dari plot pada gambar 4.5 nilai *X-val relative error* yang paling kecil menghasilkan nilai CP sebesar 0,045 dan banyaknya *terminal node* sebanyak 3. Hal ini menunjukkan bahwa pohon maksimal perlu dilakukan pemangkasan, jika tidak dilakukan pemangkasan maka terjadi *overfitting*. Pohon optimal yang didapatkan ditunjukkan pada gambar 4.6:



Gambar 4. 6 Pohon Klasifikasi Optimal CART

Pada gambar 4.6 ditunjukkan ilustrasi dari pohon klasifikasi optimal yang telah dilakukan pemangkasan pada pohon maksimal. Diketahui dari pohon optimal tersebut bahwa terdapat *terminal node* sebanyak tiga *node*. Variabel yang mempunyai kontribusi terbesar yaitu nilai MTK dan ada 3 variabel lagi yang memberikan kontribusi pada pohon optimal yaitu nilai BING, IPA dan nilai IQ. Banyaknya skor setiap variabel akan ditunjukkan pada tabel 4.8.

Tabel 4. 8 Skor Variabel Prediktor Pohon Otimal CART

Variabel	Skor (%)
MTK (X4)	58
BING (X5)	18
IPA (X3)	14
Nilai IQ (X6)	11

Tabel 4.8 dapat diketahui bahwa variabel prediktor yang paling banyak berkontribusi yaitu nilai MTK dengan nilai skor 58%. Skor ini menunjukkan bahwa sebanyak 193 data yang berkontribusi pembentukan pohon setelah

pemangkasan pohon pada pemilah nilai MTK. Nilai MTK dapat dikatakan menjadi *parent node* pada pembentukan pohon klasifikasi optimal. Tiga variabel lainnya yang memberikan kontribusi juga pada pohon optimal yaitu nilai BING dengan skor 18% (60 data yang berkontribusi dalam pembentukan pohon setelah pemangkasan), nilai IPA dengan nilai skor 14% (46 data yang berkontribusi dalam pembentukan pohon setelah pemangkasan) dan nilai IQ dengan skor 11% (37 data yang berkontribusi dalam pembentukan pohon setelah pemangkasan). Sedangkan untuk dua variabel prediktor yaitu jenis kelamin dan asal sekolah calon peserta didik baru tidak memberikan kontribusi dalam pembentukan model klasifikasi penerimaan siswa baru.

Langkah awal dalam pembentukan pohon klasifikasi dengan data *learning* yaitu dengan melakukan pemilah data menggunakan variabel yang memiliki skor tertinggi adalah nilai MTK. Diawali dengan banyaknya data *learning* sebesar 332 dipilah menjadi dua bagian yaitu bagian simpul kanan dan simpul kiri. Simpul kanan didapatkan sebanyak 99 data dengan kriteria nilai MTK lebih dari sama dengan 58 sedangkan untuk simpul kiri didapatkan sebanyak 368 data dengan kriteria nilai MTK kurang dari 58. Simpul kanan atau disebut dengan *node 2* diberikan label diterima karena dari 99 data, 79 data menunjukkan bahwa siswa baru diterima dan sisanya menunjukkan siswa tidak diterima. Simpul kiri bisa disebut dengan *terminal node 1* karena pada simpul tersebut sudah homogen sehingga tidak perlu dilakukan pemilahan lagi. *Terminal node* tersebut diberi label tidak diterima calon peserta didik baru dikarenakan banyaknya data siswa tidak diterima lebih dominan dari pada siswa yang diterima.

Kemudian dari *node 2* dipilah lagi dengan variabel yang memiliki skor tertinggi kedua setelah dari nilai MTK yaitu nilai BING. 99 data dipilah menjadi dua diperoleh 74 data untuk simpul kanan dan 25 simpul kiri. Simpul kanan yang didapatkan dari *node 2* dapat disebut dengan *terminal node 2* dan diberi label diterima. Pelabelan tersebut dikarenakan presentase data siswa diterima lebih dominan dari pada siswa tidak diterima. Sedangkan simpul kiri bisa disebut dengan *terminal node 3* dan diberi label siswa tidak diterima. Pelabelan tersebut didapatkan karena banyaknya data tidak diterima sebanyak 7 lebih banyak dari data siswa diterima.

Secara keseluruhan dari pohon klasifikasi optimal dapat disimpulkan bahwa terdapat *terminal node* sebanyak 3 *node*. Sebanyak 2 *terminal node* yang diklasifikasikan sebagai calon peserta didik baru tidak diterima. Sebanyak 1 *terminal node* diklasifikasikan sebagai calon peserta didik baru diterima. Karakteristik *terminal node* yang memiliki status diterima atau tidak diterimanya calon peserta didik baru dijelaskan sebagai berikut:

1. *Terminal node 1* terdiri dari 15 calon peserta diterima dan 218 calon peserta didik tidak diterima sehingga diberikan label untuk *node* ini yaitu tidak diterima. Karakteristik dari *node* ini yaitu calon peserta yang memiliki nilai MTK (matematika) kurang dari 58.
2. *Terminal node 2* terdiri dari 71 calon peserta yang diterima dan 3 calon peserta didik yang tidak diterima sehingga diberikan label untuk *node* ini calon peserta didik baru diterima. Karakteristik dari *node* ini yaitu calon

peserta didik yang mempunyai nilai MTK (matematika) lebih dari 58 dan memiliki nilai BING (bahasa inggris) lebih dari 53.

3. *Terminal node 3* terdiri dari 8 calon peserta didik yang diterima dan 17 calon peserta didik yang tidak diterima sehingga diberikan label untuk *node* ini calon peserta didik diterima. Karakteristik dari *node* ini yaitu calon peserta didik yang memiliki nilai MTK (matematika) lebih dari 58 dan nilai BING (bahasa inggris) nya kurang dari 53.

#### 4.3 Ketepatan Klasifikasi CART

Ketepatan suatu model pohon klasifikasi yang terbentuk pada penerimaan siswa baru untuk menentukan diterima atau tidak diterimanya siswa tersebut tidak hanya dapat dilihat dari nilai *Apparent Error Rate* (APER) dan *Total Accuracy Rate* (1-APER). Melainkan dapat juga dilihat dari tingkat *sensitivity* dan *specificity*. *Sensitivity* digunakan untuk melihat keakurasian model untuk menentukan diterimanya siswa baru tersebut. Sedangkan *specificity* digunakan untuk melihat keakurasian model untuk menentukan tidak diterimanya siswa baru tersebut.

Hasil ketepatan klasifikasi untuk data *learning* dengan menggunakan 10-*fold validation* ditampilkan pada tabel 4.9 dengan masing masing *fold*:

Tabel 4. 9 Presentase Keakurasian Pada Data *Learning*

<i>Fold</i>	APER	Akurasi	<i>Sensitivity</i>	<i>Specificity</i>
1	8,83%	91,17%	73,74%	98,02%
2	8,55%	91,45%	75,76%	97,62%
3	8,55%	91,45%	74,75%	98,02%

Lanjutan Tabel 4. 10 Presentase Keakurasian Pada Data *Learning*

<i>Fold</i>	APER	Akurasi	<i>Sensitivity</i>	<i>Specificity</i>
4	8,26%	91,74%	75,76%	98,02%
5*	7,98%	92,02%	75,76%	98,41%
6	8,83%	91,17%	74,75%	97,62%
7	9,12%	90,88%	73,74%	97,62%
8	7,69%	92,31%	74,75%	97,62%
9	8,83%	91,17%	73,74%	98,41%
10	9,12%	90,88%	73,74%	97,22%
<b>Rata-rata</b>	8,58%	91,42%	74,65%	97,868%

\*: Model pohon terbaik

Tabel 4.9 menunjukkan hasil perhitungan tingkat akurasi tiap data *learning* dengan menggunakan *10-fold cross validation*. Tabel 4.9 menunjukkan pembentukan model pohon terbaik adalah pada *fold* ke 5, karena *fold* 5 memiliki nilai akurasi tertinggi dari pada kesembilan *fold* lainnya. Nilai akurasinya sebesar 92,02%, APER sebesar 7,98%, *sensitivity* sebesar 75,76% dan nilai *specificity* sebesar 98,41%. Kesepuluh sampel pada tabel 4.9 diperoleh akurasi dengan nilai akurasi rata-rata dari kesepuluh *fold* tersebut adalah 91,42%. Nilai rata-rata tersebut dapat dikatakan nilai akurasi yang keakuratannya tinggi karena mendekati 100%, sehingga dapat dinyatakan bahwa setiap model pohon *fold* sudah stabil.

Setelah mengetahui tingkat keakurasian data *learning*, selanjutnya akan menghitung nilai akurasi data *testing* dari kombinasi pembagian data 85%:25%. Data *testing* berguna untuk melihat kelayakan suatu model pohon yang sudah terbentuk. Berikut ini merupakan perhitungan untuk keakuratan sebuah model menggunakan data *testing* dengan indeks gini:

Tabel 4. 11 Klasifikasi Data *Testing*

		Prediksi	
		Diterima	Tidak Diterima
Observasi	Diterima	11	5
	Tidak Diterima	3	39

Perhitungan yang dilakukan nilai akurasi untuk data *testing* sebesar 86,21%, maka dapat dikatakan bahwa pohon yang terbentuk dari data *learning* sudah baik dan sesuai jika digunakan untuk mengklasifikasikan data baru. Sedangkan nilai *sensitivity* sebesar 68,75% dan nilai *specificity* sebesar 92,86%.

Tabel 4. 12 Nilai Ketepatan Klasifikasi Data *Testing*

APER	Akurasi	<i>Sensitivity</i>	<i>Specificity</i>
13,79%	86,21%	68,75%	92,86%

#### 4.4 *Bootstrap Aggregating (Bagging) Classification and Regression Tree (CART)*

Proses *Bagging* CART dibagi menjadi 2 tahap yaitu yang pertama *bootstrap* dan tahap kedua yaitu *aggregating* dengan menggunakan metode *majority vote* (suara terbanyak) untuk prediksi dari keseluruhan pohon. Pada tahap *bootstrap* dilakukan pembentukan pohon dengan sebanyak 100 pengulangan pada data *learning*. Data *learning* yang digunakan yaitu 85% dari 390 data. Setelah dilakukan 100 pengulangan dengan pengembalian akan didapatkan 100 pohon klasifikasi. Pembentukan pohon klasifikasi di masing-masing data *learning* yang didapatkan memiliki tahapan-tahapan yang sama dengan pembentukan pohon CART sebelumnya hingga didapatkan 100 pohon klasifikasi optimal penerimaan peserta didik baru dengan 2 kategori yaitu diterima dan tidak diterima. Didapatkan

*variable important* dari 100 pohon yang terbentuk. Tabel 4.12 memberikan informasi bahwa skor tertinggi diperoleh variabel MTK (matematika) dengan nilai skor 98,161% yang artinya, nilai matematika memiliki kontribusi yang paling besar dalam pembentukan pohon klasifikasi. Variabel yang memiliki kontribusi paling kecil yaitu AS (asal sekolah) dengan nilai skor sebesar 5,848.

Tabel 4. 13 *Variable Important*

Variabel	Skor (%)
MTK ( $X_4$ )	98,161
BING ( $X_5$ )	75,806
IPA ( $X_3$ )	69,424
lai IQ ( $X_6$ )	40,974
JK ( $X_1$ )	6,611
AS ( $X_2$ )	5,848

Selanjutnya akan dilakukan perhitungan ketepatan klasifikasi terhadap data *testing* sehingga diketahui kebaikan dan kesalahan klasifikasi. Perhitungan ketepatan klasifikasi dilakukan pada data *testing* di setiap pohon optimal yang telah terbentuk. Data *testing* yang digunakan data *testing* kombinasi pembagian data 85%:15% di lampiran 8. Hasil prediksi data testing dari setiap pohon optimal bagging dapat dilihat di lampiran 9.

Dilihat dari lampiran 9, pada tahap *aggregating* ini menggunakan aturan *majority vote* (suara terbanyak). Pada data pertama di pohon pertama hasilnya tidak diterima, di pohon kedua hasilnya tidak diterima, dan seterusnya hingga pohon ke-100 maka hasil dari *majority vote* tidak diterima. Sehingga dapat disimpulkan pada data pertama yaitu tidak diterima karena pohon yang menyatakan tidak diterima lebih banyak dari pada pohon yang menyatakan

diterima. Pengambilan suara terbanyak ini dilakukan terus-menerus hingga data ke-58. Hasil dari *majority vote* data *testing* di lampiran 10.

Berdasarkan dari 58 calon peserta didik baru dapat diketahui hasil klasifikasi diterima sebanyak 16 dan yang tidak diterima sebanyak 42 calon peserta didik baru. Setelah didapatkan hasil klasifikasi dari data *testing*, selanjutnya akan dilakukan perhitungan ketepatan klasifikasi. Proses menentukan ketepatan klasifikasi, HASIL (diterima, tidak diterima) dari prediksi klasifikasi dicek dengan hasil data *testing*. Data pertama pada prediksi klasifikasi dinyatakan tidak diterima sedangkan pada data *testing* dinyatakan tidak diterima, maka prediksi data pertama tepat.

Tabel 4. 14 Ketepatan Klasifikasi Data *Testing Bagging* CART

	<b>DITERIMA</b>	<b>TIDAK DITERIMA</b>
<b>DITERIMA</b>	15	1
<b>TIDAK DITERIMA</b>	1	41

Dari perhitungan yang dilakukan nilai akurasi untuk data *testing* sebesar 96,52%, maka dapat dikatakan bahwa pohon yang terbentuk dari data *learning* sudah baik dan sesuai jika digunakan untuk pendugaan klasifikasi data baru dan memiliki kesalahan klasifikasi sebesar 3,48%. Sedangkan nilai *sensitivity* sebesar 93,75% dan nilai *specificity* sebesar 97,62%.

Tabel 4. 15 Ketepatan Klasifikasi *Bagging* CART

<b>APER</b>	<b>Akurasi</b>	<b>Sensitivity</b>	<b>Specificity</b>
3,48%	96,52%	93,75%	97,62%

#### 4.5 Perbandingan Ketepatan Klasifikasi Analisis CART dan Analisis

##### *Bagging* CART

Setelah dilakukan perhitungan ketepatan klasifikasi pada data *testing* menggunakan analisis CART dan *Bagging* CART dengan replikasi, berikut diperoleh nilai ketepatan klasifikasi ada tabel 4.17:

Tabel 4. 16 Nilai Ketepatan Klasifikasi CART dan *Bagging* CART

	APER	AKURASI	SENSITIVITY	SPECITIVITY
<b>CART</b>	13,79%	86,21%	68,75%	92,86%
<b><i>Bagging</i> CART</b>	3,48%	96,52%	93,75%	97,62%

Tabel 4.17 dapat diketahui bahwa ketepatan klasifikasi data *testing* CART penerimaan peserta didik baru di MBI Amnataul Ummah sebesar 86,21% dan nilai kesalahan klasifikasi sebesar 13,79%. Selanjutnya penggunaan replikasi *Bagging* CART pada data penerimaan peserta didik baru di MBI Amanatul Ummah ketepatan klasifikasinya sebesar 96,52% dan kesalahan klasifikasi 3,48%. Metode *Bagging* CART dapat meningkatkan ketepatan akurasi sebesar 10,31%, tetapi pada *bagging* CART visualisasi hasil dan interpretasi hasil tidak semudah CART.

#### 4.6 Kajian Al-Quran Tentang Bootstrap Aggregating *Classification and Regression Tree*

*Classification and regression tree* biasa dikenal dengan sebutan CART merupakan metode klasifikasi dengan algoritma *recursive biner* yang dilakukan pemilahan pada suatu kelompok menjadi dua bagian. Tujuan dari CART yaitu untuk menemukan penciri atau karakteristik dari suatu kelompok dengan akurat.

*Bootstrap aggregating* yaitu salah satu dari metode esamble yang bertujuan untuk meningkatkan tingkat akurasi pada prediksi, yang dilakukan pengulangan berkali-kali untuk mendapatkan kesalahan yang tidak terlalu besar. Firman Allah di surat Ar-Rahman sedikit menjelaskan tentang klasifikasi dan pengulangan sebagai berikut:

*“ Kedua surga itu (kelihatan) hijau tua warnanya, Maka nikmat Tuhanmu yang manakah yang kamu dustakan?, Di dalam keduanya (surga itu) ada dua buah mata air yang memanca, Maka nikmat Tuhanmu yang manakah yang kamu dustakan, Di dalam kedua surga itu ada buah-buahan, kurma dan delima, , Maka nikmat Tuhanmu yang manakah yang kamu dustakan?, Di dalam surga-surga itu ada bidadari-bidadari yang baik-baik dan jelita, , Maka nikmat Tuhanmu yang manakah yang kamu dustakan?. (Q.S. Ar-Rahman 62-73)”*

Bagi siapa yang takutkan saat menghadap Tuhannya ada dua surga. Surga untuk manusia dan surga untuk jin. Ada juga musafir yang berpendapat surga dunia dan surga akhirat. Selain dua surga yang tersebut di atas ada lagi dua surga untuk orang-orang mukmin yang kurang derajatnya dari orang-orang mukmin yang dimasukkan ke dalam surga yang pertama (Arif Fakhrudin, 2012). Dari ayat Ar- Rahman di atas dapat diketahui bahwa Allah mengklasifikasikan setiap hambanya di akhir zaman nanti untuk menempatkannya di antara surga-surganya, sesuai dengan amal ibadah hambanya. Pada surat Al-Rahman sangat banyaknya yang diulang beberapa kali, karena pada surat ini mengajak kepada seluruh manusia untuk memperbanyak bersyukur kepada Allah karena nikmat-nikmat yang telah Allah berikan kepada hambanya.

## BAB V

### PENUTUP

#### 5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan klasifikasi penerimaan peserta didik baru di Madrasah Amanatul Ummah Pacet diperoleh kesimpulan sebagai berikut:

1. Berdasarkan analisis *Bagging* CART, pengulangan yang dilakukan sebanyak 100 replikasi sehingga didapatkan pohon optimal sebanyak 100. Terdapat *variable important* yang mempengaruhi dalam pembentukan 100 pohon tersebut diantaranya yaitu nilai matematika merupakan variabel yang memberikan kontribusi terbesar dengan nilai 98,161%, yang kedua yaitu nilai bahasa inggris dengan skor 75,866%, nilai IPA dengan skor 69,424%, nilai IQ dengan skor 40,974%, JK (Jenis Kelamin) dengan skor 6,611 dan yang memiliki kontribusi terksecil dalam pembentukan pohon klasifikasi yaitu AS (Asal Sekolah) dengan skor 5,848%.
2. Tingkat akurasi ketepatan klasifikasi pada klasifikasi penerimaan peserta didik baru di Madrasah Bertaraf International Amanatul Ummah Pacet menggunakan metode *Bootstrap Aggregating (Bagging) Classification and Regression Tree (CART)* sebesar 96,52% serta nilai *specitifty* dan *sencitiviy* sebesar 93,75% dan 97,625 sedangkan nilai APERnya 3,48%. *Bagging* CART terbukti dapat meningkatkan tingkat akurasi CART pada data penerimaan peserta didik baru sebesar 10,31%. Nilai tingkat akurasi total untuk data *testing* tergolong tinggi, sehingga dapat dikatakan bahwa pohon

optmal yang terbentuk sudah baik dan sesuai atau layak bila digunakan untuk mengklasifikasikan data baru.

## 5.2 Saran

Bagi peneliti selanjutnya dapat menerapkan pembagian data *learning* dan data *testing* dengan kombinasi yang lain, pada metode *Bagging* CART dapat ditambahkan beberapa replikasi yang lebih banyak variasinya agar didapatkan hasil yang berbeda dari penelitian sebelumnya. Selain itu, peneliti dapat membentuk pohon klasifikasi dari beberapa tahun yang lalu data penerimaan peserta didik baru agar didapatkan karakteristik klasifikasi penerimaan peserta didik baru yang lebih akurat, dikarenakan untuk kasus penerimaan peserta didik baru selalu berubah-ubah setiap tahunnya dan kemampuan calon peserta didik baru yang mendaftar berbeda yang mungkin akan mengakibatkan menghasilkan pohon klasifikasi yang berbeda.

### DAFTAR RUJUKAN

- "Al-Quran dan Terjemahannya". 2002. Jakarta:Departemen Agama RI.
- Arif Fakhruddin, M. S. 2012. *Al Hidayah Al-Qur'an Tafsir Per Kata Tajwid Kode Angka*. Tangerang Selatan: PT. Kalim .
- Badruddin. 2014. *Managemen Pesrta Didik*. Jakarta: PT Indeks .
- Breiman, L. 1996. Bagging Predictors. *Machine Learning*, 24, 123-140.
- Desy Ratnaningrum, M. A.2016. Analisis Klasifikasi Nasabah Kredit Menggunakan Bootstrap Aggregating Classification and Regression Trees (Bagging CART). *Jurnal Gaussian*, Vol.5, No.1, 81-98.
- Hermawan, A. 2006. *Jaringan Saraf Tiruan*. Yogyakarta: Pustaka .
- Ian H. Witten, E. F. 2011. *Data Mining Pratical Machine Learning Tools and Techniques 3rd Edition*. USA: Elsevier .
- Pemerintahan Indonesia. 2003. *Undang-Undang Republik Indonesia No.20 Tahun 2003 Tentang Pendidikan Nasional .Lembaran RI Tahuun 2003 No. 20*. Jakarta: Sekretariat Negara.
- Leo Breiman, J. H. 1998. *Classification and Regression Trees*. New York Washington, D.C: CRC Press Print.
- Lewis, R. J. 2000. *An Introduction to Classification and Regression Tree (CART)Analysis*. Fransisco, California: UCLA Medical Center.
- Miftahur Rohman, H. 2018. Konsep Tujuan Pendiidkan Islam Perspektif Nilai-nilai Sosial Kultural. *Jurnal Pendiidkan Islam*, Vol. 9, No. 1 , 21-35.
- Mock. Abdul Mukid, T. W. 2015. BAGGING Classification and Regression Trees untuk Prediksi Risiko Preeklampsia. *Media Statistika*, Vol. No. 2, 111-120.

- Muhammad Jamal Muttaqin, B. W. 2013. Metode Esamble pada CART untuk Perbaikan Klasifikasi Kemiskinan di Kabupaten Jombang. *Institute Teknologi Sepuluh November*, 1-6.
- Mustolehudin. 2011. Tradisi Baca Tulis dalam Islam Kajian Terhadap Teksi AL-Quran Surah Al 'Alaq Ayat 1-5. *Jurnal Analisa Vol. XVIII*, 145-154.
- Richard A. Johnson, D. W. 2002. *Alied Multivariabel Statistical Analysis Fifth Edition* . USA: Prentice-Hall, inc.
- Rosadi, K. I. 2012. Efektifitas Kinerja Lembaga Penjaminan Mutu endidikan (LPMP) dalam Meningkatkan Mutu Pendiidkan. *Jurnal Administrasi Pendidikan Vol. XIV No.1*, 1-19.
- Suking, A. 2015. Sistem Penerimaan Siswa Baru di Sekolah Efektif. *Prosiding TEMILNAS* (hal. 503-512). Surabaya: Dwiputra Pustaka Jaya.
- Sumartini, Siti Holis dan Purnami, Santi Wulan. 2015. Penggunaan Metode *CLassification and Regreassion Trees* (CART) untuk Klasifikasi Rekurensi Pasien Kanker Serviks di RSUD Dr. Soetomo. *Jurnal SAINS dan Seni ITS, Vol. 4, No. 2*, 2301-928X Print
- Sutton, C. D. 2005. Classification and Regression Trees, Bagging and Boosting. *Handbok of Statistic, Vol. 24*, 304-327.
- Tatang Hidayatt, A. S. 2011. Pendidikan dalam Prespektif Islam dan Peranannya dalam Membeni Kepribadian. *Jurnal Mudarrisuna, Vol. 8, No. 2*, 218-244.
- Wibowo, Ari. 2015. Analisis Perbandingan Kinerja Metode Klasifikasi Dalam Data Mining. *Jurnal Integrasi, Vol. 7, No. 1*, 23-30

Ying Huang, J. C. 2004. Classifying G-protein Couled Receptors with Bagging Classification Tree. *Computational Biology and Cemistry*, 275-280.



## LAMPIRAN

Lampiran 1 : Data Penerimaan Siswa Baru di Madrasah Amanatul Ummah Pacet

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	1	25.0	31.4	58.0	98	TIDAK DITERIMA
1	1	40.0	22.9	68.0	121	TIDAK DITERIMA
1	2	25.0	37.1	50.0	91	TIDAK DITERIMA
1	2	45.0	34.3	56.0	104	TIDAK DITERIMA
1	1	50.0	51.4	50.0	113	TIDAK DITERIMA
1	1	60.0	42.9	84.0	113	DITERIMA
1	1	40.0	34.3	56.0	88	TIDAK DITERIMA
2	1	25.0	20.0	66.0	96	TIDAK DITERIMA
2	2	75.0	57.1	70.0	127	DITERIMA
2	1	20.0	17.1	56.0	116	TIDAK DITERIMA
1	1	20.0	20.0	44.0	107	TIDAK DITERIMA
1	1	60.0	48.6	62.0	125	DITERIMA
2	1	45.0	45.7	64.0	112	TIDAK DITERIMA
2	1	25.0	34.3	38.0	112	TIDAK DITERIMA
1	1	10.0	20.0	58.0	113	TIDAK DITERIMA
1	1	30.0	25.7	22.0	81	TIDAK DITERIMA
1	1	20.0	37.1	44.0	104	TIDAK DITERIMA
1	1	80.0	62.9	70.0	106	DITERIMA
1	1	65.0	42.9	36.0	120	TIDAK DITERIMA
1	1	55.0	28.6	40.0	125	TIDAK DITERIMA
1	1	25.0	34.3	30.0	104	TIDAK DITERIMA
1	1	20.0	31.4	46.0	81	TIDAK DITERIMA
1	1	20.0	20.0	40.0	68	TIDAK DITERIMA
1	1	80.0	54.3	42.0	120	DITERIMA
1	1	30.0	20.0	52.0	98	TIDAK DITERIMA
1	1	40.0	34.3	52.0	88	TIDAK DITERIMA
1	1	40.0	28.6	28.0	104	TIDAK DITERIMA
1	2	25.0	34.3	62.0	104	TIDAK DITERIMA
1	2	40.0	31.4	66.0	120	TIDAK DITERIMA
1	1	40.0	37.1	36.0	113	TIDAK DITERIMA
2	1	40.0	40.0	56.0	113	TIDAK DITERIMA

## Lampiran 1 : Lanjutan

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	1	25.0	22.9	34.0	91	TIDAK DITERIMA
2	2	30.0	34.3	44.0	98	TIDAK DITERIMA
1	1	50.0	57.1	64.0	106	DITERIMA
1	1	35.0	28.6	36.0	98	TIDAK DITERIMA
1	1	40.0	25.7	44.0	96	TIDAK DITERIMA
2	1	60.0	34.3	60.0	113	TIDAK DITERIMA
1	2	35.0	45.7	60.0	125	TIDAK DITERIMA
2	1	20.0	42.9	32.0	91	TIDAK DITERIMA
2	1	35.0	25.7	38.0	76	TIDAK DITERIMA
2	1	50.0	22.9	70.0	120	TIDAK DITERIMA
1	2	45.0	14.3	28.0	84	TIDAK DITERIMA
2	1	75.0	48.6	80.0	128	DITERIMA
2	1	55.0	34.3	62.0	125	TIDAK DITERIMA
2	1	50.0	40.0	64.0	120	TIDAK DITERIMA
2	1	35.0	28.6	54.0	98	TIDAK DITERIMA
1	1	5.0	37.1	64.0	84	TIDAK DITERIMA
2	1	40.0	25.7	62.0	113	TIDAK DITERIMA
2	1	20.0	37.1	44.0	91	TIDAK DITERIMA
1	2	25.0	31.4	28.0	88	TIDAK DITERIMA
1	1	70.0	40.0	54.0	98	DITERIMA
1	1	40.0	31.4	44.0	128	TIDAK DITERIMA
2	1	25.0	25.7	32.0	98	TIDAK DITERIMA
2	1	15.0	25.7	44.0	96	TIDAK DITERIMA
2	1	60.0	40.0	60.0	112	DITERIMA
1	1	50.0	31.4	44.0	98	TIDAK DITERIMA
2	1	30.0	34.3	40.0	104	TIDAK DITERIMA
1	1	10.0	28.6	54.0	91	TIDAK DITERIMA
1	1	30.0	22.9	50.0	120	TIDAK DITERIMA
2	2	60.0	37.1	68.0	121	DITERIMA
1	1	20.0	37.1	30.0	84	TIDAK DITERIMA
1	1	50.0	40.0	68.0	112	DITERIMA
2	1	45.0	31.4	42.0	120	TIDAK DITERIMA

## Lampiran 1 : Lanjutan

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	1	20.0	20.0	32.0	91	TIDAK DITERIMA
2	1	85.0	42.9	72.0	112	DITERIMA
1	1	20.0	20.0	28.0	73	TIDAK DITERIMA
2	1	55.0	51.4	66.0	112	DITERIMA
2	1	55.0	40.0	36.0	112	TIDAK DITERIMA
2	1	60.0	37.1	44.0	127	TIDAK DITERIMA
1	1	60.0	34.3	24.0	112	TIDAK DITERIMA
1	1	65.0	31.4	66.0	127	DITERIMA
2	1	45.0	45.7	40.0	116	TIDAK DITERIMA
2	1	25.0	34.3	40.0	125	TIDAK DITERIMA
2	1	75.0	31.4	72.0	125	DITERIMA
1	2	15.0	45.7	42.0	120	TIDAK DITERIMA
2	1	25.0	31.4	54.0	121	TIDAK DITERIMA
2	2	70.0	37.1	78.0	136	DITERIMA
2	2	70.0	57.1	62.0	104	DITERIMA
1	1	55.0	45.7	62.0	106	DITERIMA
2	1	25.0	31.4	24.0	104	TIDAK DITERIMA
2	1	45.0	37.1	48.0	104	TIDAK DITERIMA
2	1	85.0	42.9	76.0	113	DITERIMA
2	1	30.0	37.1	50.0	98	TIDAK DITERIMA
1	1	20.0	31.4	28.0	88	TIDAK DITERIMA
2	1	60.0	34.3	72.0	104	DITERIMA
2	1	50.0	42.9	60.0	112	TIDAK DITERIMA
1	1	35.0	25.7	32.0	98	TIDAK DITERIMA
2	1	60.0	51.4	82.0	113	DITERIMA
2	1	50.0	37.1	54.0	94	TIDAK DITERIMA
1	2	35.0	34.3	50.0	109	TIDAK DITERIMA
1	1	35.0	40.0	44.0	113	TIDAK DITERIMA
1	1	65.0	48.6	56.0	106	DITERIMA
2	1	55.0	25.7	68.0	98	TIDAK DITERIMA
1	1	50.0	37.1	44.0	112	TIDAK DITERIMA
2	1	50.0	34.3	42.0	98	TIDAK DITERIMA

## Lampiran 1 : Lanjutan

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
2	2	35.0	34.3	30.0	91	TIDAK DITERIMA
2	1	50.0	42.9	32.0	112	TIDAK DITERIMA
1	1	60.0	22.9	52.0	125	TIDAK DITERIMA
2	1	20.0	22.9	38.0	106	TIDAK DITERIMA
2	2	90.0	40.0	80.0	136	DITERIMA
2	2	70.0	57.1	60.0	127	DITERIMA
2	2	75.0	37.1	54.0	127	DITERIMA
2	2	20.0	25.7	54.0	100	TIDAK DITERIMA
2	1	30.0	22.9	48.0	104	TIDAK DITERIMA
1	1	50.0	28.6	58.0	104	TIDAK DITERIMA
2	1	50.0	22.9	50.0	124	TIDAK DITERIMA
2	1	45.0	25.7	46.0	133	TIDAK DITERIMA
2	1	20.0	17.1	18.0	68	TIDAK DITERIMA
1	1	30.0	34.3	52.0	120	TIDAK DITERIMA
1	2	50.0	45.7	64.0	127	DITERIMA
1	1	35.0	42.9	46.0	120	TIDAK DITERIMA
2	1	70.0	51.4	72.0	127	DITERIMA
1	1	20.0	28.6	32.0	84	TIDAK DITERIMA
1	2	30.0	37.1	32.0	76	TIDAK DITERIMA
2	1	30.0	28.6	48.0	112	TIDAK DITERIMA
1	1	5.0	22.9	40.0	88	TIDAK DITERIMA
1	1	60.0	28.6	40.0	91	TIDAK DITERIMA
2	1	60.0	28.6	66.0	98	TIDAK DITERIMA
2	1	85.0	60.0	72.0	133	DITERIMA
1	1	45.0	45.7	78.0	98	DITERIMA
1	1	30.0	11.4	18.0	98	TIDAK DITERIMA
1	1	40.0	34.3	38.0	106	TIDAK DITERIMA
2	1	45.0	40.0	68.0	96	TIDAK DITERIMA
2	1	50.0	37.1	50.0	106	TIDAK DITERIMA
2	1	10.0	34.3	38.0	88	TIDAK DITERIMA
1	1	40.0	34.3	50.0	106	TIDAK DITERIMA
1	2	55.0	31.4	48.0	106	TIDAK DITERIMA

**Lampran 1 : Lanjutan**

<b>JK</b>	<b>AS</b>	<b>MTK</b>	<b>IPA</b>	<b>BING</b>	<b>NILAI IQ</b>	<b>HASIL</b>
2	1	80.0	40.0	60.0	98	DITERIMA
2	1	45.0	37.1	70.0	125	TIDAK DITERIMA
2	1	75.0	28.6	58.0	112	DITERIMA
2	1	90.0	51.4	60.0	136	DITERIMA
2	1	30.0	40.0	38.0	106	TIDAK DITERIMA
2	1	40.0	37.1	76.0	104	TIDAK DITERIMA
2	2	45.0	40.0	44.0	96	TIDAK DITERIMA
2	2	30.0	31.4	38.0	127	TIDAK DITERIMA
1	1	50.0	25.7	46.0	84	TIDAK DITERIMA
2	1	50.0	31.4	64.0	133	TIDAK DITERIMA
1	2	50.0	34.3	54.0	127	TIDAK DITERIMA
2	1	55.0	40.0	56.0	120	TIDAK DITERIMA
2	2	35.0	45.7	64.0	91	TIDAK DITERIMA
2	2	35.0	48.6	60.0	104	TIDAK DITERIMA
2	1	70.0	37.1	38.0	121	TIDAK DITERIMA
2	1	65.0	40.0	76.0	121	DITERIMA
2	1	40.0	37.1	58.0	104	TIDAK DITERIMA
2	2	45.0	42.9	60.0	106	TIDAK DITERIMA
1	1	35.0	28.6	62.0	96	TIDAK DITERIMA
1	1	80.0	40.0	46.0	133	DITERIMA
2	1	90.0	51.4	68.0	113	DITERIMA
1	1	60.0	31.4	80.0	112	DITERIMA
1	1	90.0	34.3	66.0	121	DITERIMA
1	1	50.0	28.6	68.0	113	TIDAK DITERIMA
2	1	50.0	28.6	34.0	104	TIDAK DITERIMA
2	1	65.0	45.7	66.0	112	DITERIMA
1	1	20.0	34.3	24.0	98	TIDAK DITERIMA
2	1	35.0	42.9	64.0	96	TIDAK DITERIMA
2	1	70.0	45.7	66.0	96	DITERIMA
2	1	55.0	31.4	30.0	104	TIDAK DITERIMA
1	2	60.0	37.1	48.0	106	TIDAK DITERIMA
2	1	35.0	42.9	44.0	104	TIDAK DITERIMA

## Lampiran 1 : Lanjutan

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
2	1	30.0	34.3	56.0	104	TIDAK DITERIMA
2	1	55.0	28.6	48.0	104	TIDAK DITERIMA
1	1	85.0	54.3	62.0	127	DITERIMA
2	1	20.0	42.9	54.0	98	TIDAK DITERIMA
2	1	50.0	34.3	36.0	112	TIDAK DITERIMA
2	1	55.0	51.4	64.0	121	DITERIMA
2	1	20.0	34.3	60.0	96	TIDAK DITERIMA
2	1	70.0	54.3	60.0	112	DITERIMA
2	2	50.0	28.6	54.0	106	TIDAK DITERIMA
2	2	75.0	54.3	66.0	127	DITERIMA
2	1	30.0	48.6	68.0	113	TIDAK DITERIMA
2	1	45.0	25.7	36.0	81	TIDAK DITERIMA
2	1	30.0	31.4	36.0	98	TIDAK DITERIMA
2	1	60.0	45.7	64.0	133	DITERIMA
2	1	60.0	62.9	78.0	120	DITERIMA
1	1	45.0	25.7	50.0	96	TIDAK DITERIMA
1	1	10.0	28.6	40.0	104	TIDAK DITERIMA
1	1	45.0	25.7	76.0	113	TIDAK DITERIMA
1	1	35.0	28.6	24.0	84	TIDAK DITERIMA
1	1	35.0	40.0	48.0	81	TIDAK DITERIMA
1	1	75.0	48.6	72.0	113	DITERIMA
1	1	30.0	34.3	60.0	84	TIDAK DITERIMA
1	1	55.0	31.4	58.0	106	TIDAK DITERIMA
1	1	70.0	22.9	44.0	120	TIDAK DITERIMA
1	1	75.0	57.1	74.0	142	DITERIMA
1	1	25.0	22.9	42.0	91	TIDAK DITERIMA
1	2	95.0	45.7	64.0	136	DITERIMA
1	1	15.0	28.6	38.0	109	TIDAK DITERIMA
1	2	75.0	28.6	62.0	136	DITERIMA
1	1	35.0	40.0	74.0	125	TIDAK DITERIMA
1	1	20.0	37.1	22.0	73	TIDAK DITERIMA
1	1	30.0	28.6	58.0	96	TIDAK DITERIMA

## Lampiran 1 : Lanjutan

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
2	2	40.0	48.6	34.0	98	TIDAK DITERIMA
2	1	55.0	28.6	46.0	112	TIDAK DITERIMA
1	2	75.0	54.3	66.0	121	DITERIMA
1	1	35.0	34.3	56.0	98	TIDAK DITERIMA
1	1	25.0	25.7	74.0	112	TIDAK DITERIMA
1	1	65.0	31.4	62.0	127	DITERIMA
2	1	10.0	25.7	40.0	104	TIDAK DITERIMA
2	1	65.0	48.6	50.0	98	DITERIMA
2	2	50.0	37.1	36.0	98	TIDAK DITERIMA
2	1	30.0	34.3	42.0	112	TIDAK DITERIMA
1	1	25.0	28.6	36.0	96	TIDAK DITERIMA
1	1	55.0	34.3	50.0	112	TIDAK DITERIMA
1	1	10.0	31.4	24.0	88	TIDAK DITERIMA
1	1	25.0	25.7	66.0	112	TIDAK DITERIMA
1	1	40.0	25.7	44.0	109	TIDAK DITERIMA
1	1	50.0	25.7	54.0	104	TIDAK DITERIMA
1	1	35.0	25.7	24.0	112	TIDAK DITERIMA
1	1	15.0	25.7	64.0	98	TIDAK DITERIMA
1	1	40.0	37.1	36.0	76	TIDAK DITERIMA
1	1	35.0	5.7	28.0	84	TIDAK DITERIMA
1	1	60.0	28.6	52.0	98	TIDAK DITERIMA
1	1	50.0	48.6	38.0	91	TIDAK DITERIMA
1	1	15.0	45.7	64.0	96	TIDAK DITERIMA
1	1	15.0	22.9	40.0	88	TIDAK DITERIMA
1	1	90.0	54.3	66.0	120	DITERIMA
1	2	65.0	34.3	46.0	121	TIDAK DITERIMA
1	1	60.0	37.1	54.0	120	TIDAK DITERIMA
1	1	55.0	28.6	24.0	120	TIDAK DITERIMA
1	2	30.0	28.6	56.0	113	TIDAK DITERIMA
1	1	10.0	25.7	78.0	96	TIDAK DITERIMA
1	1	55.0	45.7	56.0	106	DITERIMA
1	1	30.0	22.9	62.0	120	TIDAK DITERIMA

**Lampiran 1: Lanjutan**

<b>JK</b>	<b>AS</b>	<b>MTK</b>	<b>IPA</b>	<b>BING</b>	<b>NILAI IQ</b>	<b>HASIL</b>
1	1	15.0	31.4	52.0	96	TIDAK DITERIMA
1	1	85.0	45.7	46.0	120	DITERIMA
1	1	45.0	34.3	52.0	91	TIDAK DITERIMA
1	1	60.0	42.9	52.0	112	DITERIMA
1	2	40.0	28.6	50.0	113	TIDAK DITERIMA
1	1	50.0	42.9	48.0	104	TIDAK DITERIMA
1	2	50.0	28.6	48.0	104	TIDAK DITERIMA
1	1	20.0	37.1	28.0	113	TIDAK DITERIMA
1	1	40.0	31.4	28.0	112	TIDAK DITERIMA
1	1	15.0	28.6	46.0	84	TIDAK DITERIMA
1	1	45.0	37.1	44.0	98	TIDAK DITERIMA
1	1	35.0	40.0	38.0	104	TIDAK DITERIMA
1	1	70.0	45.7	84.0	128	DITERIMA
1	2	50.0	40.0	72.0	113	DITERIMA
1	1	65.0	42.9	78.0	91	DITERIMA
1	2	20.0	42.9	24.0	91	TIDAK DITERIMA
1	1	20.0	40.0	42.0	98	TIDAK DITERIMA
1	2	30.0	37.1	46.0	112	TIDAK DITERIMA
1	1	60.0	48.6	64.0	127	DITERIMA
1	1	40.0	28.6	26.0	120	TIDAK DITERIMA
1	1	55.0	37.1	74.0	1	DITERIMA
1	1	65.0	40.0	50.0	121	DITERIMA
1	1	50.0	25.7	78.0	120	TIDAK DITERIMA
1	1	45.0	37.1	62.0	128	TIDAK DITERIMA
1	1	65.0	14.3	54.0	91	TIDAK DITERIMA
1	1	35.0	34.3	38.0	104	TIDAK DITERIMA
1	1	65.0	40.0	54.0	106	DITERIMA
1	1	20.0	37.1	34.0	112	TIDAK DITERIMA
1	1	50.0	17.1	48.0	106	TIDAK DITERIMA
1	2	65.0	34.3	64.0	120	DITERIMA
1	1	25.0	34.3	22.0	88	TIDAK DITERIMA
2	1	75.0	42.9	34.0	127	TIDAK DITERIMA

## Lampiran 1 : Lanjutan

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
2	2	65.0	40.0	56.0	112	DITERIMA
2	1	70.0	25.7	46.0	98	TIDAK DITERIMA
2	1	55.0	40.0	74.0	125	DITERIMA
2	1	40.0	37.1	46.0	98	TIDAK DITERIMA
2	1	75.0	37.1	44.0	112	DITERIMA
2	1	50.0	45.7	48.0	96	TIDAK DITERIMA
2	1	60.0	25.7	64.0	98	TIDAK DITERIMA
2	1	45.0	40.0	62.0	104	TIDAK DITERIMA
2	1	60.0	31.4	38.0	136	TIDAK DITERIMA
2	1	30.0	31.4	40.0	104	TIDAK DITERIMA
2	1	95.0	74.3	78.0	121	DITERIMA
2	1	60.0	34.3	48.0	127	TIDAK DITERIMA
2	1	65.0	37.1	66.0	120	DITERIMA
2	1	85.0	37.1	60.0	112	DITERIMA
2	1	40.0	28.6	56.0	84	TIDAK DITERIMA
2	2	50.0	48.6	76.0	120	DITERIMA
1	2	85.0	62.9	90.0	121	DITERIMA
2	1	30.0	20.0	42.0	96	TIDAK DITERIMA
2	1	35.0	25.7	58.0	120	TIDAK DITERIMA
2	1	45.0	40.0	50.0	104	TIDAK DITERIMA
2	2	45.0	37.1	54.0	120	TIDAK DITERIMA
2	1	10.0	28.6	58.0	104	TIDAK DITERIMA
2	1	50.0	20.0	36.0	104	TIDAK DITERIMA
2	2	40.0	48.6	48.0	109	TIDAK DITERIMA
2	1	25.0	51.4	74.0	113	TIDAK DITERIMA
1	1	65.0	28.6	46.0	113	TIDAK DITERIMA
2	1	55.0	34.3	50.0	127	TIDAK DITERIMA
1	2	45.0	34.3	64.0	127	TIDAK DITERIMA
2	1	80.0	60.0	64.0	127	DITERIMA
2	1	30.0	25.7	36.0	98	TIDAK DITERIMA
2	1	20.0	37.1	68.0	120	TIDAK DITERIMA
2	1	60.0	37.1	64.0	120	DITERIMA

**Lampiran 1 : Lanjutan**

<b>JK</b>	<b>AS</b>	<b>MTK</b>	<b>IPA</b>	<b>BING</b>	<b>NILAI IQ</b>	<b>HASIL</b>
2	1	40.0	31.4	56.0	104	TIDAK DITERIMA
2	1	55.0	31.4	48.0	104	TIDAK DITERIMA
2	1	25.0	20.0	48.0	9	TIDAK DITERIMA
2	1	45.0	31.4	66.0	98	TIDAK DITERIMA
2	1	85.0	31.4	46.0	120	DITERIMA
2	2	30.0	31.4	46.0	116	TIDAK DITERIMA
2	1	45.0	34.3	76.0	104	DITERIMA
2	1	65.0	48.6	56.0	106	DITERIMA
2	1	40.0	37.1	54.0	133	TIDAK DITERIMA
2	1	60.0	42.9	52.0	125	TIDAK DITERIMA
2	1	15.0	31.4	48.0	84	TIDAK DITERIMA
2	2	65.0	28.6	68.0	109	DITERIMA
2	1	25.0	14.3	38.0	88	TIDAK DITERIMA
2	1	55.0	37.1	58.0	106	TIDAK DITERIMA
2	1	35.0	20.0	48.0	125	TIDAK DITERIMA
2	1	25.0	31.4	42.0	136	TIDAK DITERIMA
2	1	95.0	57.1	64.0	128	DITERIMA
2	1	30.0	22.9	74.0	104	TIDAK DITERIMA
2	1	35.0	28.6	48.0	112	TIDAK DITERIMA
2	1	40.0	28.6	58.0	120	TIDAK DITERIMA
2	1	70.0	40.0	76.0	96	DITERIMA
2	1	55.0	40.0	42.0	133	TIDAK DITERIMA
2	1	80.0	40.0	60.0	120	DITERIMA
2	1	30.0	34.3	32.0	104	TIDAK DITERIMA
1	2	15.0	31.4	38.0	104	TIDAK DITERIMA
1	1	40.0	25.7	42.0	96	TIDAK DITERIMA
2	1	45.0	34.3	50.0	91	TIDAK DITERIMA
2	1	85.0	42.9	58.0	113	DITERIMA
2	1	60.0	42.9	66.0	112	DITERIMA
2	1	25.0	31.4	50.0	84	TIDAK DITERIMA
2	1	70.0	54.3	86.0	136	DITERIMA
2	2	70.0	40.0	50.0	112	DITERIMA

## Lampiran 1 : Lanjutan

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
1	1	65.0	25.7	60.0	112	TIDAK DITERIMA
2	1	70.0	51.4	70.0	106	DITERIMA
1	1	40.0	34.3	44.0	112	TIDAK DITERIMA
1	1	40.0	40.0	54.0	112	TIDAK DITERIMA
1	2	55.0	34.3	56.0	106	TIDAK DITERIMA
1	1	60.0	40.0	48.0	112	TIDAK DITERIMA
1	1	5.0	31.4	60.0	104	TIDAK DITERIMA
2	1	50.0	31.4	40.0	104	TIDAK DITERIMA
1	1	20.0	22.9	48.0	91	TIDAK DITERIMA
1	2	35.0	20.0	62.0	96	TIDAK DITERIMA
2	1	40.0	34.3	28.0	112	TIDAK DITERIMA
1	1	35.0	17.1	26.0	112	TIDAK DITERIMA
2	1	25.0	42.9	50.0	91	TIDAK DITERIMA
1	1	30.0	40.0	46.0	100	TIDAK DITERIMA
2	2	30.0	37.1	46.0	106	TIDAK DITERIMA
2	2	30.0	34.3	68.0	127	TIDAK DITERIMA
1	1	15.0	34.3	28.0	88	TIDAK DITERIMA
1	1	90.0	42.9	74.0	113	DITERIMA
2	1	60.0	31.4	46.0	121	TIDAK DITERIMA
2	2	50.0	31.4	54.0	120	TIDAK DITERIMA
1	1	70.0	51.4	58.0	120	DITERIMA
2	1	75.0	48.6	66.0	113	DITERIMA
2	2	15.0	34.3	50.0	106	TIDAK DITERIMA
2	2	55.0	40.0	60.0	112	DITERIMA
2	1	70.0	48.6	58.0	142	DITERIMA
2	1	35.0	34.3	28.0	91	TIDAK DITERIMA
1	1	40.0	42.9	42.0	104	TIDAK DITERIMA
2	2	75.0	31.4	62.0	127	DITERIMA
1	2	80.0	42.9	60.0	121	DITERIMA
2	1	30.0	28.6	38.0	91	TIDAK DITERIMA
2	1	55.0	40.0	44.0	112	TIDAK DITERIMA
2	1	75.0	57.1	64.0	112	DITERIMA

## Lampiran 1 : Lanjutan

JK	AS	MTK	IPA	BING	NILAI IQ	HASIL
2	1	35.0	40.0	52.0	120	TIDAK DITERIMA
2	2	25.0	25.7	36.0	96	TIDAK DITERIMA
2	1	35.0	37.1	68.0	98	TIDAK DITERIMA
2	1	75.0	28.6	66.0	125	DITERIMA
2	1	95.0	62.9	62.0	127	DITERIMA
2	1	65.0	25.7	70.0	112	DITERIMA
2	1	25.0	28.6	22.0	96	TIDAK DITERIMA
2	1	30.0	34.3	36.0	98	TIDAK DITERIMA
1	2	25.0	20.0	42.0	81	TIDAK DITERIMA
2	2	80.0	37.1	74.0	120	DITERIMA
1	1	30.0	40.0	32.0	104	TIDAK DITERIMA
1	1	25.0	22.9	14.0	91	TIDAK DITERIMA
1	1	20.0	37.1	32.0	96	TIDAK DITERIMA
2	1	50.0	42.9	50.0	112	TIDAK DITERIMA
2	1	70.0	31.4	58.0	127	DITERIMA
2	1	45.0	22.9	52.0	88	TIDAK DITERIMA
2	1	55.0	51.4	76.0	136	DITERIMA
2	1	25.0	25.7	38.0	91	TIDAK DITERIMA
2	1	20.0	20.0	42.0	76	TIDAK DITERIMA
2	1	75.0	51.4	40.0	104	DITERIMA
2	1	20.0	45.7	40.0	120	TIDAK DITERIMA
1	2	20.0	25.7	48.0	84	TIDAK DITERIMA
2	1	55.0	28.6	40.0	113	TIDAK DITERIMA
1	2	45.0	22.9	48.0	96	TIDAK DITERIMA
1	2	80.0	54.3	80.0	128	DITERIMA
2	1	65.0	48.6	70.0	120	DITERIMA
1	1	30.0	25.7	32.0	127	TIDAK DITERIMA
2	1	20.0	28.6	32.0	84	TIDAK DITERIMA
2	1	55.0	48.6	60.0	127	DITERIMA
2	2	85.0	62.9	68.0	112	DITERIMA
2	2	65.0	25.7	76.0	113	DITERIMA
1	1	60.0	45.7	54.0	112	DITERIMA

**Lampiran 1 : Lanjutan**

<b>JK</b>	<b>AS</b>	<b>MTK</b>	<b>IPA</b>	<b>BING</b>	<b>NILAI IQ</b>	<b>HASIL</b>
2	1	65.0	25.7	38.0	91	TIDAK DITERIMA
1	1	20.0	34.3	40.0	84	TIDAK DITERIMA
2	1	50.0	28.6	84.0	96	DITERIMA
2	1	35.0	51.4	74.0	112	DITERIMA
2	1	65.0	37.1	78.0	113	DITERIMA
1	1	45.0	40.0	68.0	125	TIDAK DITERIMA
2	1	20.0	46.0	42.0	76	TIDAK DITERIMA



**Lampiran 2 : Statistik deskriptif variabel prediktor**

## Statistik deskriptif variabel nilai MTK

<b>Statistics</b>				
<b>Variable</b>	<b>Mean</b>	<b>StDev</b>	<b>Minimum</b>	<b>Maximum</b>
Diterima	69,73	12,56	35,00	95,00
Tidak Diterima	36,786	14,708	5,000	75,000

## Statistik deskriptif variabel nilai BING

<b>Statistics</b>				
<b>Variable</b>	<b>Mean</b>	<b>StDev</b>	<b>Minimum</b>	<b>Maximum</b>
Diterima	44,934	9,625	25,700	74,300
Tidak Diterima	32,110	7,704	5,700	51,400

## Statistik deskriptif variabel nilai IPA

<b>Statistics</b>				
<b>Variable</b>	<b>Mean</b>	<b>StDev</b>	<b>Minimum</b>	<b>Maximum</b>
Diterima	65,964	10,053	40,000	90,000
Tidak Diterima	46,693	12,877	14,000	78,000

## Statistik Deskriptif Variabel Nilai IQ

<b>Statistics</b>					
<b>Variable</b>	<b>HASIL</b>	<b>Mean</b>	<b>StDev</b>	<b>Minimum</b>	<b>Maximum</b>
NILAI IQ	DITERIMA	117,77	10,88	91,00	142,00
	TIDAK DITERIMA	104,06	13,87	68,00	136,00

**Lampiran 3 : Aturan pohon dan pohon optimal setiap kombinasi pembagian data**

Aturan pohon maksimal CART presentase data 75%:25%

```

> p175
n= 293

(node), split, n, loss, yval, (yprob)
* denotes terminal node

1) root 293 83 TIDAK DITERIMA (0.283276451 0.716723549)
2) MTK<=62.5 69 8 DITERIMA (0.884057971 0.115942029)
4) BING<=48 56 0 DITERIMA (1.000000000 0.000000000) *
5) BING< 48 13 5 TIDAK DITERIMA (0.384615385 0.615384615) *
3) MTK< 62.5 224 22 TIDAK DITERIMA (0.098214286 0.901785714)
6) BING<=71 15 4 DITERIMA (0.733333333 0.266666667) *
7) BING< 71 209 11 TIDAK DITERIMA (0.052631579 0.947368421)
14) BING<=59 41 10 TIDAK DITERIMA (0.243902439 0.756097561)
28) MTK<=52.5 11 3 DITERIMA (0.727272727 0.272727273) *
29) MTK< 52.5 30 2 TIDAK DITERIMA (0.066666667 0.933333333) *
15) BING< 59 168 1 TIDAK DITERIMA (0.005952381 0.994047619) *

> |
    
```

Aturan pohon optimal CART presentase data 75%:25%

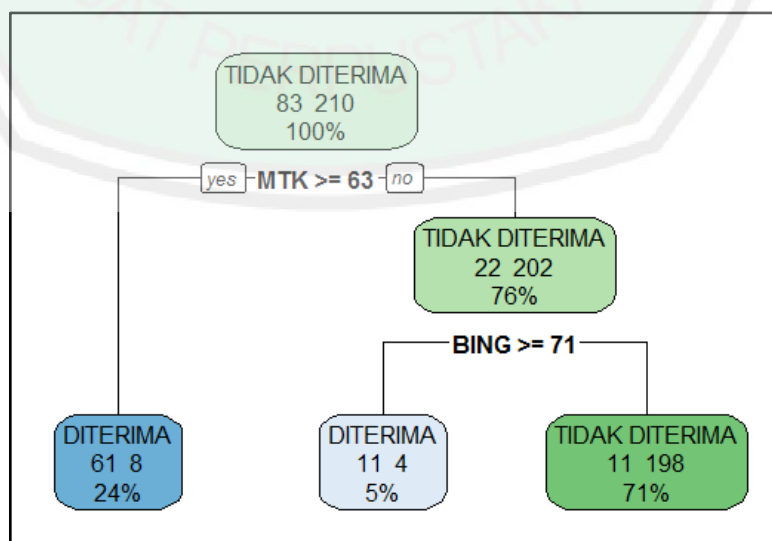
```

n= 293

(node), split, n, loss, yval, (yprob)
* denotes terminal node

1) root 293 83 TIDAK DITERIMA (0.28327645 0.71672355)
2) MTK<=62.5 69 8 DITERIMA (0.88405797 0.11594203) *
3) MTK< 62.5 224 22 TIDAK DITERIMA (0.09821429 0.90178571)
6) BING<=71 15 4 DITERIMA (0.73333333 0.26666667) *
7) BING< 71 209 11 TIDAK DITERIMA (0.05263158 0.94736842) *
    
```

Pohon optimal CART presentase data 75%:25%



Aturan pohon maksimal CART presentase data 80%:20%

```
n= 312
(node), split, n, loss, yval, (yprob)
* denotes terminal node

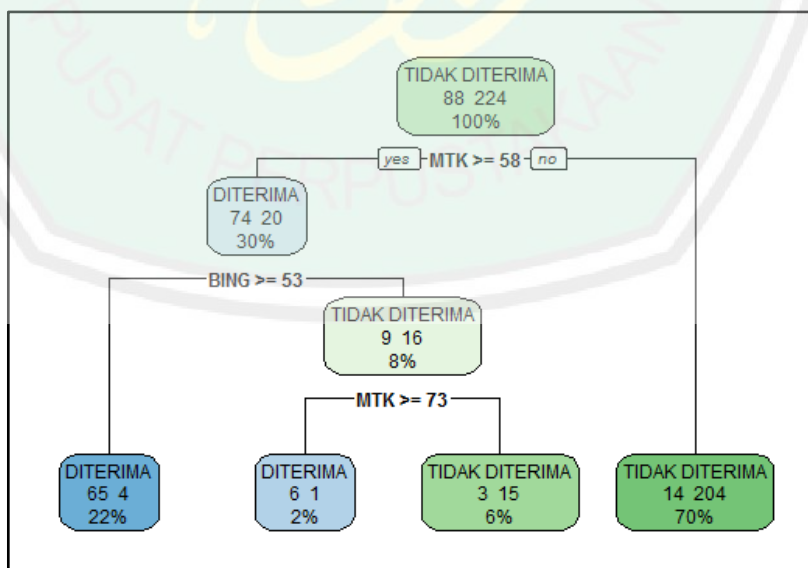
1) root 312 88 TIDAK DITERIMA (0.282051282 0.717948718)
2) MTK>=57.5 94 20 DITERIMA (0.787234043 0.212765957)
4) BING>=53 69 4 DITERIMA (0.942028986 0.057971014) *
5) BING< 53 25 9 TIDAK DITERIMA (0.360000000 0.640000000)
10) MTK>=72.5 7 1 DITERIMA (0.857142857 0.142857143) *
11) MTK< 72.5 18 3 TIDAK DITERIMA (0.166666667 0.833333333) *
3) MTK< 57.5 218 14 TIDAK DITERIMA (0.064220183 0.935779817)
6) BING>=71 15 7 TIDAK DITERIMA (0.466666667 0.533333333) *
7) BING< 71 203 7 TIDAK DITERIMA (0.034482759 0.965517241)
14) IPA>=44.3 20 6 TIDAK DITERIMA (0.300000000 0.700000000)
28) MTK>=47.5 9 3 DITERIMA (0.666666667 0.333333333) *
29) MTK< 47.5 11 0 TIDAK DITERIMA (0.000000000 1.000000000) *
15) IPA< 44.3 183 1 TIDAK DITERIMA (0.005464481 0.994535519) *
```

Aturan pohon optimal CART presentase data 80%:20%

```
> prunetree
n= 312
(node), split, n, loss, yval, (yprob)
* denotes terminal node

1) root 312 88 TIDAK DITERIMA (0.28205128 0.71794872)
2) MTK>=57.5 94 20 DITERIMA (0.78723404 0.21276596)
4) BING>=53 69 4 DITERIMA (0.94202899 0.05797101) *
5) BING< 53 25 9 TIDAK DITERIMA (0.36000000 0.64000000)
10) MTK>=72.5 7 1 DITERIMA (0.85714286 0.14285714) *
11) MTK< 72.5 18 3 TIDAK DITERIMA (0.16666667 0.83333333) *
3) MTK< 57.5 218 14 TIDAK DITERIMA (0.06422018 0.93577982) *
```

Pohon optimal CART presentase data 80%:20%



Aturan pohon maksimal CART presentase data 85%:15%

```

> p185
n= 332

node), split, n, loss, yval, (yprob)
  * denotes terminal node

1) root 332 94 TIDAK DITERIMA (0.28313253 0.71686747)
2) MTK>=57.5 99 20 DITERIMA (0.79797980 0.20202020)
4) BING>=53 74 3 DITERIMA (0.95945946 0.04054054) *
5) BING< 53 25 8 TIDAK DITERIMA (0.32000000 0.68000000)
10) IPA>=38.55 10 4 DITERIMA (0.60000000 0.40000000) *
11) IPA< 38.55 15 2 TIDAK DITERIMA (0.13333333 0.86666667) *
3) MTK< 57.5 233 15 TIDAK DITERIMA (0.06437768 0.93562232) *
    
```

Aturan pohon optimal CART presentase data 85%:15%

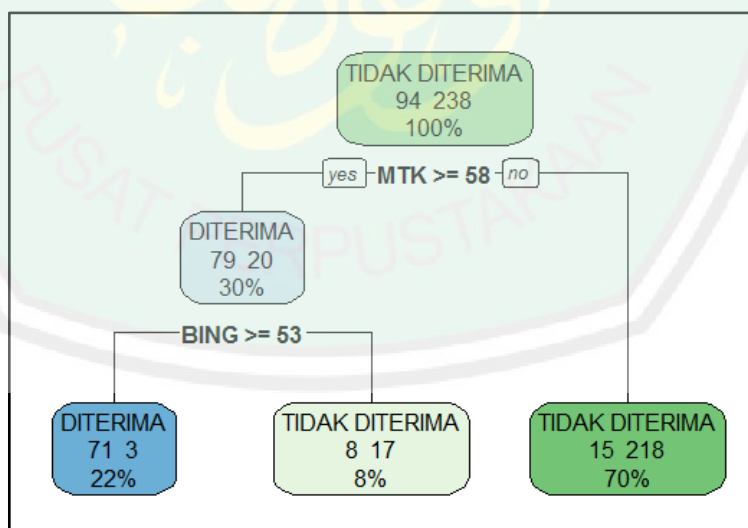
```

n= 332

node), split, n, loss, yval, (yprob)
  * denotes terminal node

1) root 332 94 TIDAK DITERIMA (0.28313253 0.71686747)
2) MTK>=57.5 99 20 DITERIMA (0.79797980 0.20202020)
4) BING>=53 74 3 DITERIMA (0.95945946 0.04054054) *
5) BING< 53 25 8 TIDAK DITERIMA (0.32000000 0.68000000) *
3) MTK< 57.5 233 15 TIDAK DITERIMA (0.06437768 0.93562232) *
    
```

Pohon optimal CART presentase data 85%:15%



Aturan pohon maksimal CART presentase data 90%:10%

```

n= 351
node), split, n, loss, yval, (yprob)
* denotes terminal node

1) root 351 99 TIDAK DITERIMA (0.28205128 0.71794872)
2) MTK>=57.5 106 24 DITERIMA (0.77358491 0.22641509)
4) BING>=53 81 6 DITERIMA (0.92592593 0.07407407) *
5) BING< 53 25 7 TIDAK DITERIMA (0.28000000 0.72000000)
10) MTK>=67.5 10 4 DITERIMA (0.60000000 0.40000000) *
11) MTK< 67.5 15 1 TIDAK DITERIMA (0.06666667 0.93333333) *
3) MTK< 57.5 245 17 TIDAK DITERIMA (0.06938776 0.93061224)
6) IPA>=44.3 25 11 TIDAK DITERIMA (0.44000000 0.56000000)
12) MTK>=47.5 11 2 DITERIMA (0.81818182 0.18181818) *
13) MTK< 47.5 14 2 TIDAK DITERIMA (0.14285714 0.85714286) *
7) IPA< 44.3 220 6 TIDAK DITERIMA (0.02727273 0.97272727) *
    
```

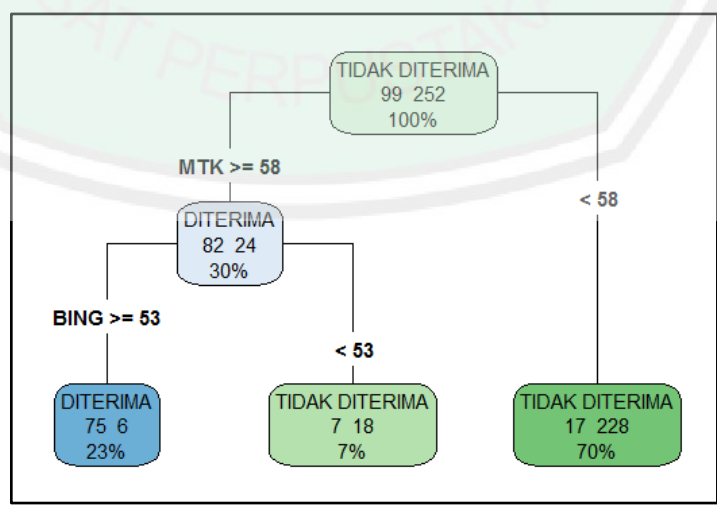
Aturan pohon optimal CART presentase data 90%:10%

```

n= 351
node), split, n, loss, yval, (yprob)
* denotes terminal node

1) root 351 99 TIDAK DITERIMA (0.28205128 0.71794872)
2) MTK>=57.5 106 24 DITERIMA (0.77358491 0.22641509)
4) BING>=53 81 6 DITERIMA (0.92592593 0.07407407) *
5) BING< 53 25 7 TIDAK DITERIMA (0.28000000 0.72000000) *
3) MTK< 57.5 245 17 TIDAK DITERIMA (0.06938776 0.93061224) *
    
```

Pohon optimal CART presentase data 90%:10%



**Lampiran 4 : Proses pemangkasan pohon secara rinci presentase 80%:20%****OUTPUT CART presentase 80%:20% dengan nilai CP**

```

Node number 2: 94 observations,      complexity param=0.07954545
predicted class=DITERIMA      expected loss=0.212766  P(node) =0.3012821
class counts:   74   20
probabilities: 0.787 0.213
left son=4 (69 obs) right son=5 (25 obs)
Primary splits:
  BING < 53   to the right, improve=12.4331300, (0 missing)
  IPA < 38.55 to the right, improve= 7.4267730, (0 missing)
  MTK < 62.5  to the right, improve= 6.4296520, (0 missing)
  NILAI.IQ < 101 to the right, improve= 2.7577080, (0 missing)
  AS < 3.5   to the right, improve= 0.6461185, (0 missing)
Surrogate splits:
  IPA < 24.3  to the right, agree=0.755, adj=0.08, (0 split)
  NILAI.IQ < 93.5 to the right, agree=0.745, adj=0.04, (0 split)

Node number 3: 218 observations,      complexity param=0.01136364
predicted class=TIDAK DITERIMA expected loss=0.06422018  P(node) =0.6987179
class counts:   14  204
probabilities: 0.064 0.936
left son=6 (15 obs) right son=7 (203 obs)
Primary splits:
  BING < 71   to the right, improve=5.2179270, (0 missing)
  IPA < 44.3  to the right, improve=4.1362830, (0 missing)
  MTK < 47.5  to the right, improve=3.4872780, (0 missing)
  NILAI.IQ < 105 to the right, improve=1.3301810, (0 missing)
  AS < 3.5   to the right, improve=0.1001033, (0 missing)

Node number 4: 69 observations
predicted class=DITERIMA      expected loss=0.05797101  P(node) =0.2211538
class counts:   65   4
probabilities: 0.942 0.058

```

```

Node number 5: 25 observations,      complexity param=0.05681818
predicted class=TIDAK DITERIMA expected loss=0.36  P(node) =0.08012821
class counts:   9   16
probabilities: 0.360 0.640
left son=10 (7 obs) right son=11 (18 obs)
Primary splits:
  MTK < 72.5  to the right, improve=4.8057140, (0 missing)
  IPA < 38.55 to the right, improve=3.8533330, (0 missing)
  NILAI.IQ < 120.5 to the left, improve=1.2994120, (0 missing)
  BING < 41   to the right, improve=0.9168254, (0 missing)
  JK < 1.5   to the right, improve=0.1482051, (0 missing)
Surrogate splits:
  IPA < 44.3  to the right, agree=0.8, adj=0.286, (0 split)

Node number 6: 15 observations
predicted class=TIDAK DITERIMA expected loss=0.4666667  P(node) =0.04807692
class counts:   7   8
probabilities: 0.467 0.533

Node number 7: 203 observations,      complexity param=0.01136364
predicted class=TIDAK DITERIMA expected loss=0.03448276  P(node) =0.650641
class counts:   7  196
probabilities: 0.034 0.966
left son=14 (20 obs) right son=15 (183 obs)
Primary splits:
  IPA < 44.3  to the right, improve=3.12817000, (0 missing)
  MTK < 52.5  to the right, improve=1.73559900, (0 missing)
  BING < 59   to the right, improve=1.32951100, (0 missing)
  NILAI.IQ < 105 to the right, improve=0.68390800, (0 missing)
  AS < 3.5   to the right, improve=0.02398985, (0 missing)

```

```

Node number 10: 7 observations
predicted class=DITERIMA      expected loss=0.1428571  P(node) =0.0224359
class counts:      6      1
probabilities: 0.857 0.143

Node number 11: 18 observations
predicted class=TIDAK DITERIMA expected loss=0.1666667  P(node) =0.05769231
class counts:      3      15
probabilities: 0.167 0.833

Node number 14: 20 observations,      complexity param=0.01136364
predicted class=TIDAK DITERIMA expected loss=0.3  P(node) =0.06410256
class counts:      6      14
probabilities: 0.300 0.700
left son=28 (9 obs) right son=29 (11 obs)
Primary splits:
  MTK < 47.5 to the right, improve=4.4000000, (0 missing)
  BING < 53 to the right, improve=2.4000000, (0 missing)
  NILAI.IQ < 105 to the right, improve=1.9384620, (0 missing)
  JK < 1.5 to the left, improve=0.6828283, (0 missing)
  AS < 3.5 to the left, improve=0.5318681, (0 missing)
Surrogate splits:
  JK < 1.5 to the left, agree=0.70, adj=0.333, (0 split)
  AS < 3.5 to the left, agree=0.70, adj=0.333, (0 split)
  IPA < 50 to the right, agree=0.70, adj=0.333, (0 split)
  NILAI.IQ < 126 to the right, agree=0.65, adj=0.222, (0 split)
  BING < 45 to the right, agree=0.60, adj=0.111, (0 split)

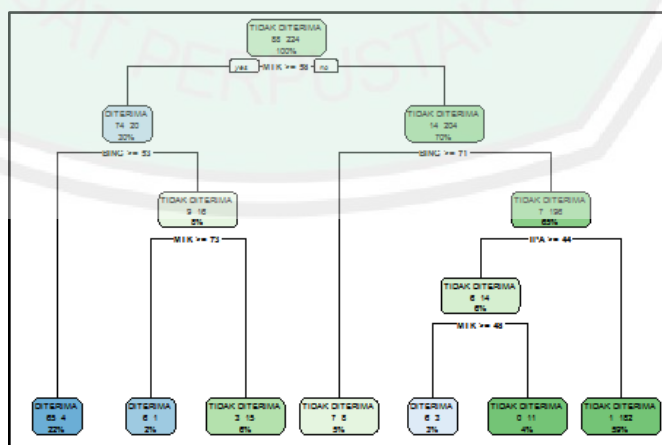
Node number 15: 183 observations
predicted class=TIDAK DITERIMA expected loss=0.005464481  P(node) =0.5865385
class counts:      1      182
probabilities: 0.005 0.995
    
```

```

Node number 28: 9 observations
predicted class=DITERIMA      expected loss=0.3333333  P(node) =0.02884615
class counts:      6      3
probabilities: 0.667 0.333

Node number 29: 11 observations
predicted class=TIDAK DITERIMA expected loss=0  P(node) =0.03525641
class counts:      0      11
probabilities: 0.000 1.000
    
```

Pohon Maksimal Klasifikasi CART Presentase 80%:20%



## Aturan Pemangkasan Pohon CART Presentase 80%:20%

```
> summary(prunetree)
Call:
rpart(formula = HASIL ~ ., data = IQC.learning, method = "class")
n= 312

      CP nsplit rel error   xerror   xstd
1 0.61363636  0 1.0000000 1.0000000 0.09032446
2 0.07954545  1 0.3863636 0.4545455 0.06710494
3 0.05681818  2 0.3068182 0.4659091 0.06781357
4 0.01136364  3 0.2500000 0.4318182 0.06564582

Variable importance
      MTK      BING      IPA NILAI.IQ
      63      16      14      8

Node number 1: 312 observations, complexity param=0.6136364
predicted class=TIDAK DITERIMA expected loss=0.2820513 P(node) =1
class counts: 88 224
probabilities: 0.282 0.718
left son=2 (94 obs) right son=3 (218 obs)
Primary splits:
      MTK < 57.5 to the right, improve=68.6677800, (0 missing)
      BING < 57 to the right, improve=36.6952800, (0 missing)
      IPA < 44.3 to the right, improve=33.7918000, (0 missing)
      NILAI.IQ < 105 to the right, improve=21.7392800, (0 missing)
      JK < 1.5 to the right, improve= 0.9465469, (0 missing)

Surrogate splits:
      IPA < 47.3 to the right, agree=0.760, adj=0.202, (0 split)
      NILAI.IQ < 120.5 to the right, agree=0.737, adj=0.128, (0 split)
      BING < 65 to the right, agree=0.724, adj=0.085, (0 split)

Node number 2: 94 observations, complexity param=0.07954545
predicted class=DITERIMA expected loss=0.212766 P(node) =0.3012821
class counts: 74 20
probabilities: 0.787 0.213
left son=4 (69 obs) right son=5 (25 obs)
Primary splits:
      BING < 53 to the right, improve=12.4331300, (0 missing)
      IPA < 38.55 to the right, improve= 7.4267730, (0 missing)
      MTK < 62.5 to the right, improve= 6.4296520, (0 missing)
      NILAI.IQ < 101 to the right, improve= 2.7577080, (0 missing)
      AS < 3.5 to the right, improve= 0.6461185, (0 missing)
Surrogate splits:
      IPA < 24.3 to the right, agree=0.755, adj=0.08, (0 split)
      NILAI.IQ < 93.5 to the right, agree=0.745, adj=0.04, (0 split)

Node number 3: 218 observations
predicted class=TIDAK DITERIMA expected loss=0.06422018 P(node) =0.6987179
class counts: 14 204
probabilities: 0.064 0.936
```

```

Node number 4: 69 observations
predicted class=DITERIMA      expected loss=0.05797101  P(node) =0.2211538
class counts:      65      4
probabilities: 0.942 0.058

Node number 5: 25 observations,      complexity param=0.05681818
predicted class=TIDAK DITERIMA  expected loss=0.36  P(node) =0.08012821
class counts:      9      16
probabilities: 0.360 0.640
left son=10 (7 obs) right son=11 (18 obs)
Primary splits:
  MTK < 72.5 to the right, improve=4.8057140, (0 missing)
  IPA < 38.55 to the right, improve=3.8533330, (0 missing)
  NILAI.IQ < 120.5 to the left, improve=1.2994120, (0 missing)
  BING < 41 to the right, improve=0.9168254, (0 missing)
  JK < 1.5 to the right, improve=0.1482051, (0 missing)
Surrogate splits:
  IPA < 44.3 to the right, agree=0.8, adj=0.286, (0 split)

Node number 10: 7 observations
predicted class=DITERIMA      expected loss=0.1428571  P(node) =0.0224359
class counts:      6      1
probabilities: 0.857 0.143

Node number 11: 18 observations
predicted class=TIDAK DITERIMA  expected loss=0.1666667  P(node) =0.05769231
class counts:      3      15
probabilities: 0.167 0.833

```



**Lampiran 5** : Perhitungan akurasi setiap kombinasi pembagian data

Nilai akurasi kombinasi 75%:25% data *testing*

```
> table(IQC75.testing[,7],pt75)
      pt75
      DITERIMA TIDAK DITERIMA
DITERIMA      20           7
TIDAK DITERIMA 6          64
> rata75
[1] 0.8659794
```

Nilai akurasi kombinasi 75%:25% data *learning*

```
> table(IQC75.learning[,7],pt75l)
      pt75l
      DITERIMA TIDAK DITERIMA
DITERIMA      72          11
TIDAK DITERIMA 12         198
> rata75l
[1] 0.9488055
```

Nilai akurasi kombinasi 80%:20% data *testing*

```
> table(IQC.testing[,7],pt)
      pt
      DITERIMA TIDAK DITERIMA
DITERIMA      17           5
TIDAK DITERIMA 2          54
> rata8
[1] 0.9230769
```

Nilai akurasi kombinasi 80%:20% data *learning*

```
> table(IQC.learning[,7],ptl)
      ptl
      DITERIMA TIDAK DITERIMA
DITERIMA      71          17
TIDAK DITERIMA 5          219
> rata8l
[1] 0.9294872
```

Nilai akurasi kombinasi 85%:15% data *testing*

```
> table(IQC85.testing[,7],pt85)
      pt85
      DITERIMA TIDAK DITERIMA
DITERIMA      11          5
TIDAK DITERIMA  3         39
> rata85
[1] 0.862069
```

Nilai akurasi kombinasi 85%:15% data *learning*

```
> table(IQC85.learning[,7],pt85l)
      pt85l
      DITERIMA TIDAK DITERIMA
DITERIMA      71         23
TIDAK DITERIMA  3        235
> rata85l
[1] 0.9216867
```

Nilai akurasi kombinasi 90%:10% data *testing*

```
> table(IQC70.testing[,7],pt70)
      pt70
      DITERIMA TIDAK DITERIMA
DITERIMA       7          4
TIDAK DITERIMA  0         28
> rata70
[1] 0.8974359
```

Nilai akurasi kombinasi 90%:10% data *learning*

```
> table(IQC70.learning[,7],pt70l)
      pt70l
      DITERIMA TIDAK DITERIMA
DITERIMA      90          9
TIDAK DITERIMA 12        240
> rata70l
[1] 0.9401709
```

**Lampiran 6** : Tabel prediksi data *learning* dari setiap *fold**Fold 1*

ptk011		
	DITERIMA	TIDAK DITERIMA
DITERIMA	73	26
TIDAK DITERIMA	5	247

*Fold 2*

ptk021		
	DITERIMA	TIDAK DITERIMA
DITERIMA	75	24
TIDAK DITERIMA	6	246

*Fold 3*

ptk031		
	DITERIMA	TIDAK DITERIMA
DITERIMA	74	25
TIDAK DITERIMA	5	247

*Fold 4*

ptk041		
	DITERIMA	TIDAK DITERIMA
DITERIMA	75	24
TIDAK DITERIMA	5	247

*Fold 5*

ptk051		
	DITERIMA	TIDAK DITERIMA
DITERIMA	75	24
TIDAK DITERIMA	4	248

*Fold 6*

ptk061		
	DITERIMA	TIDAK DITERIMA
DITERIMA	74	25
TIDAK DITERIMA	6	246

*Fold 7*

ptk071		
	DITERIMA	TIDAK DITERIMA
DITERIMA	73	26
TIDAK DITERIMA	6	246

*Fold 8*

ptk081		
	DITERIMA	TIDAK DITERIMA
DITERIMA	74	25
TIDAK DITERIMA	6	246

*Fold 9*

ptk091		
	DITERIMA	TIDAK DITERIMA
DITERIMA	73	26
TIDAK DITERIMA	4	248

*Fold 10*

ptk101		
	DITERIMA	TIDAK DITERIMA
DITERIMA	73	26
TIDAK DITERIMA	6	246

Lampiran 7 : Output tabel prediksi data *testing*

		ptgbag85	
		DITERIMA	TIDAK DITERIMA
DITERIMA		15	1
TIDAK DITERIMA		1	41



**Lampiran 8 : Data testing**

<b>DATA</b>	<b>JK</b>	<b>AS</b>	<b>MTK</b>	<b>IPA</b>	<b>BING</b>	<b>IQ</b>	<b>HASIL</b>
1	2	4	75	57.1	70	127	DITERIMA
2	1	3	80	62.9	70	106	DITERIMA
3	1	3	80	54.3	42	120	DITERIMA
4	1	4	40	31.4	66	120	TIDAK DITERIMA
5	1	3	40	37.1	36	113	TIDAK DITERIMA
6	2	4	30	34.3	44	98	TIDAK DITERIMA
7	1	4	35	45.7	60	125	TIDAK DITERIMA
8	2	3	55	34.3	62	125	TIDAK DITERIMA
9	2	3	35	28.6	54	98	TIDAK DITERIMA
10	2	3	40	25.7	62	113	TIDAK DITERIMA
11	2	3	20	37.1	44	91	TIDAK DITERIMA
12	2	4	60	37.1	68	121	DITERIMA
13	1	3	55	45.7	62	106	DITERIMA
14	2	3	55	25.7	68	98	TIDAK DITERIMA
15	2	3	20	17.1	18	68	TIDAK DITERIMA
16	2	3	70	51.4	72	127	DITERIMA
17	1	3	20	28.6	32	84	TIDAK DITERIMA
18	1	3	5	22.9	40	88	TIDAK DITERIMA
19	1	3	30	11.4	18	98	TIDAK DITERIMA
20	1	3	50	25.7	46	84	TIDAK DITERIMA
21	1	3	60	31.4	80	112	DITERIMA
22	2	3	65	45.7	66	112	DITERIMA
23	1	4	60	37.1	48	106	TIDAK DITERIMA
24	2	3	55	51.4	64	121	DITERIMA

**Lampiran 8: Lanjutan**

<b>DATA</b>	<b>JK</b>	<b>AS</b>	<b>MTK</b>	<b>IPA</b>	<b>BING</b>	<b>IQ</b>	<b>HASIL</b>
25	2	3	55	28.6	46	112	TIDAK DITERIMA
26	1	3	25	28.6	36	96	TIDAK DITERIMA
27	1	3	50	48.6	38	91	TIDAK DITERIMA
28	1	3	60	37.1	54	120	TIDAK DITERIMA
29	1	3	30	22.9	62	120	TIDAK DITERIMA
30	1	3	85	45.7	46	120	DITERIMA
31	1	3	50	42.9	48	104	TIDAK DITERIMA
32	1	4	50	28.6	48	104	TIDAK DITERIMA
33	1	3	15	28.6	46	84	TIDAK DITERIMA
34	1	4	20	42.9	24	91	TIDAK DITERIMA
35	1	3	45	37.1	62	128	TIDAK DITERIMA
36	1	3	65	14.3	54	91	TIDAK DITERIMA
37	1	3	25	34.3	22	88	TIDAK DITERIMA
38	2	3	60	25.7	64	98	TIDAK DITERIMA
39	2	3	60	34.3	48	127	TIDAK DITERIMA
40	2	3	35	25.7	58	120	TIDAK DITERIMA
41	2	4	45	37.1	54	120	TIDAK DITERIMA
42	2	3	40	31.4	56	104	TIDAK DITERIMA
43	2	3	45	31.4	66	98	TIDAK DITERIMA
44	2	3	35	28.6	48	112	TIDAK DITERIMA
45	1	3	40	25.7	42	96	TIDAK DITERIMA

**Lampiran 8: Lanjutan**

<b>DATA</b>	<b>JK</b>	<b>AS</b>	<b>MTK</b>	<b>IPA</b>	<b>BING</b>	<b>IQ</b>	<b>HASIL</b>
46	2	3	50	31.4	40	104	TIDAK DITERIMA
47	1	3	20	22.9	48	91	TIDAK DITERIMA
48	1	4	35	20.0	62	96	TIDAK DITERIMA
49	1	3	90	42.9	74	113	DITERIMA
50	2	3	75	48.6	66	113	DITERIMA
51	2	3	30	28.6	38	91	TIDAK DITERIMA
52	2	3	65	25.7	70	112	DITERIMA
53	2	3	30	34.3	36	98	TIDAK DITERIMA
54	1	4	25	20.0	42	81	TIDAK DITERIMA
55	2	3	45	22.9	52	88	TIDAK DITERIMA
56	1	4	80	54.3	80	128	DITERIMA
57	2	3	35	51.4	74	112	DITERIMA
58	2	3	65	37.1	78	113	DITERIMA

**Lampiran 9:** Prediksi Data *Testisng Fold 5* Disetiap Pohon

<b>DATA</b>	<b>POHON 1</b>	<b>POHON 2</b>	<b>POHON 3</b>	<b>....</b>	<b>POHON 100</b>
1	DITERIMA	DITERIMA	DITERIMA	....	DITERIMA
2	DITERIMA	DITERIMA	DITERIMA	....	DITERIMA
3	DITERIMA	DITERIMA	DITERIMA	....	DITERIMA
4	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
5	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
6	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
7	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
8	DITERIMA	DITERIMA	DITERIMA	....	TIDAK DITERIMA
9	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	....	TIDAK DITERIMA
10	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
11	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
12	DITERIMA	DITERIMA	TIDAK DITERIMA	....	DITERIMA
13	DITERIMA	DITERIMA	DITERIMA	....	DITERIMA
14	TIDAK DITERIMA	DITERIMA	DITERIMA	....	TIDAK DITERIMA
15	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	....	TIDAK DITERIMA
16	DITERIMA	DITERIMA	TIDAK DITERIMA	....	DITERIMA
17	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	....	TIDAK DITERIMA
18	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
19	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
20	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
21	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
22	DITERIMA	DITERIMA	DITERIMA	....	DITERIMA
23	TIDAK DITERIMA	DITERIMA	DITERIMA	....	TIDAK DITERIMA

**Lampiran 9:** Lanjutan

<b>DATA</b>	<b>POHON 1</b>	<b>POHON 2</b>	<b>POHON 3</b>	<b>....</b>	<b>POHON 100</b>
24	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	....	DITERIMA
25	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
26	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
27	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	DITERIMA
28	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	....	DITERIMA
29	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	....	TIDAK DITERIMA
30	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
31	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
32	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	DITERIMA
33	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
34	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
35	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	....	TIDAK DITERIMA
36	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
37	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	....	DITERIMA
38	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
39	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
40	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
41	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
42	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
43	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA

**Lampiran 9:** Lanjutan

<b>DATA</b>	<b>POHON 1</b>	<b>POHON 2</b>	<b>POHON 3</b>	<b>....</b>	<b>POHON 100</b>
44	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
45	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
46	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
47	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
48	TIDAK DITERIMA	DITERIMA	DITERIMA	....	DITERIMA
49	DITERIMA	DITERIMA	DITERIMA	....	DITERIMA
50	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
51	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	....	DITERIMA
52	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
53	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
54	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
55	TIDAK DITERIMA	DITERIMA	DITERIMA	....	DITERIMA
56	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	....	TIDAK DITERIMA
57	TIDAK DITERIMA	DITERIMA	DITERIMA	....	DITERIMA
58	DITERIMA	DITERIMA	DITERIMA	....	DITERIMA

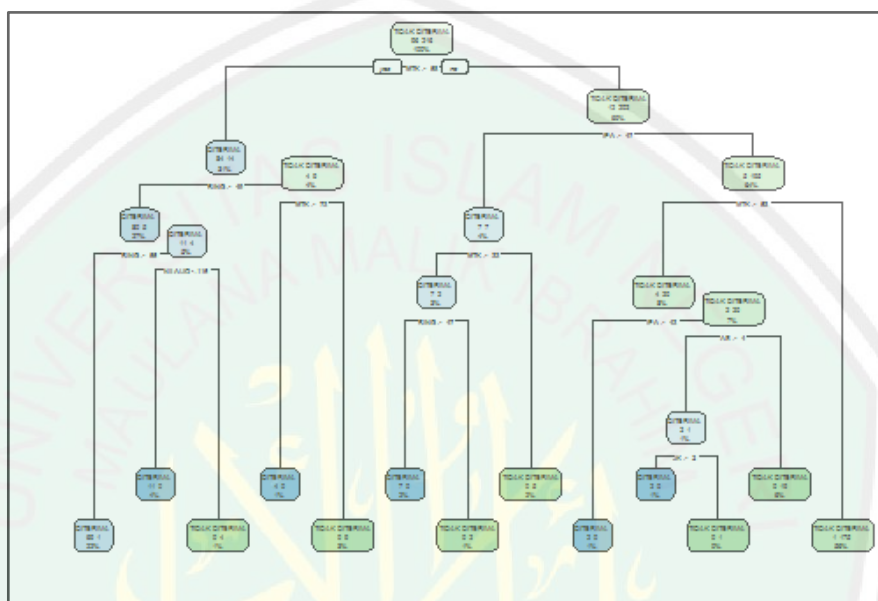
**Lampiran 10 : Hasil Majority Vote Data Testing**

<b>DATA</b>	<b>HASIL</b>
1	DITERIMA
2	DITERIMA
3	DITERIMA
4	TIDAK DITERIMA
5	TIDAK DITERIMA
6	TIDAK DITERIMA
7	TIDAK DITERIMA
8	TIDAK DITERIMA
9	TIDAK DITERIMA
10	TIDAK DITERIMA
11	DITERIMA
12	DITERIMA
13	TIDAK DITERIMA
14	TIDAK DITERIMA
15	DITERIMA
16	TIDAK DITERIMA
17	TIDAK DITERIMA
18	TIDAK DITERIMA
19	TIDAK DITERIMA
20	DITERIMA
21	DITERIMA
22	DITERIMA
23	TIDAK DITERIMA
24	DITERIMA
25	TIDAK DITERIMA
26	TIDAK DITERIMA
27	TIDAK DITERIMA
28	DITERIMA
29	TIDAK DITERIMA

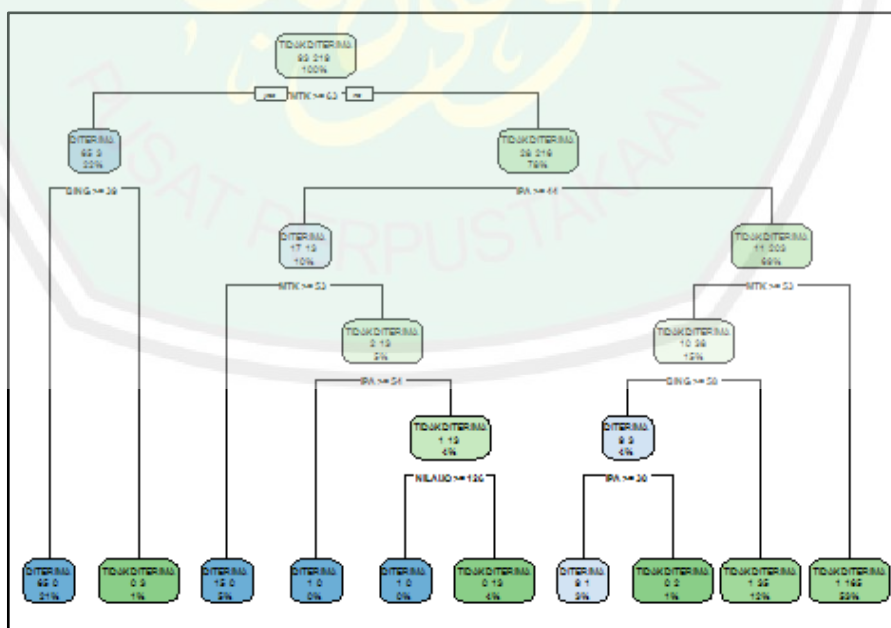
<b>DATA</b>	<b>HASIL</b>
30	DITERIMA
31	TIDAK DITERIMA
32	TIDAK DITERIMA
33	TIDAK DITERIMA
34	TIDAK DITERIMA
35	TIDAK DITERIMA
36	TIDAK DITERIMA
37	TIDAK DITERIMA
38	TIDAK DITERIMA
39	TIDAK DITERIMA
40	TIDAK DITERIMA
41	TIDAK DITERIMA
42	TIDAK DITERIMA
43	TIDAK DITERIMA
44	TIDAK DITERIMA
45	TIDAK DITERIMA
46	TIDAK DITERIMA
47	TIDAK DITERIMA
48	TIDAK DITERIMA
49	DITERIMA
50	DITERIMA
51	TIDAK DITERIMA
52	DITERIMA
53	TIDAK DITERIMA
54	TIDAK DITERIMA
55	TIDAK DITERIMA
56	DITERIMA
57	TIDAK DITERIMA
58	DITERIMA

**Lampiran 11:** Pohon 1, 2,3, ..., 98, 99, 100 Bootstrap Aggregating Classification and Regression Tree

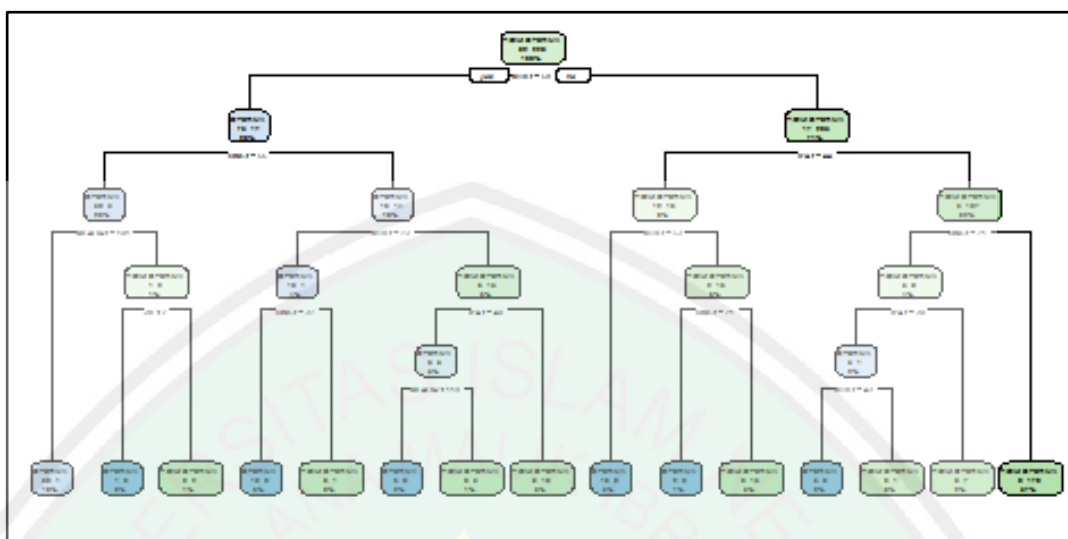
Pohon 1



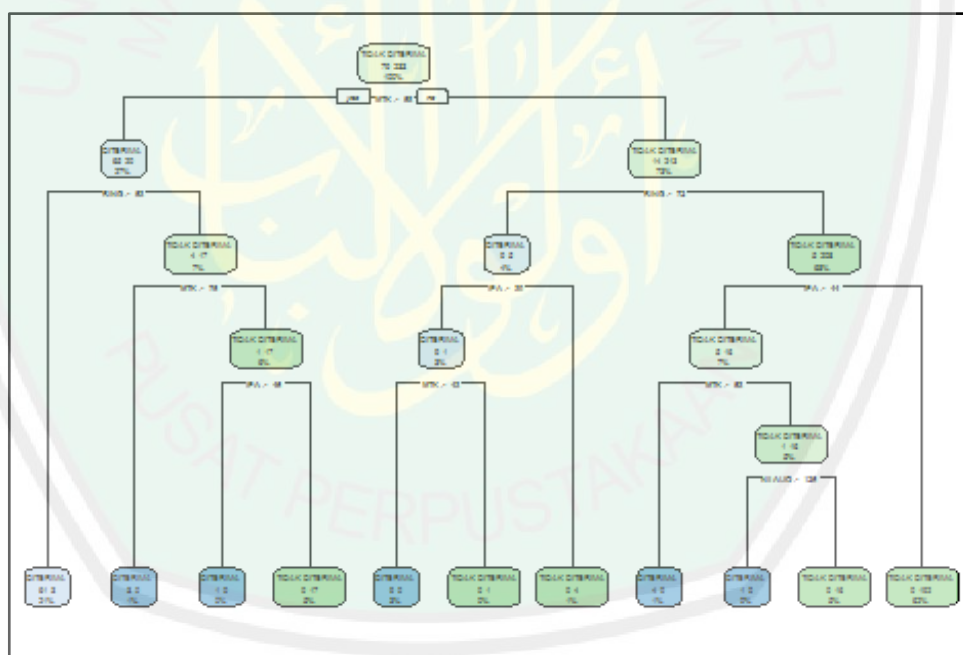
Pohon 2



Pohon 3



Pohon 98





## Lampiran 12: Prediksi dugaan Data *Testing* di Pohon 1,2,3...,98,99,100 *Bagging* CART

### Pohon 1

```

Levels: DITERIMA TIDAK DITERIMA
> ptgbag851<-predict(gbag85[["mtrees"]][[1]][["btree"]], newdata=IQC85.testing, type = "class")
> ptgbag851

```

9	18	24	29	30	33
DITERIMA	DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
38	44	46	48	49	60
TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA
79	93	108	112	113	116
DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
121	136	149	153	158	165
TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
193	202	213	218	223	225
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
229	230	233	239	247	248
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA
254	262	267	274	276	288
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
291	306	313	327	328	329
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
337	341	349	357	359	360
DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
367	376	387	388		
TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA		

Levels: DITERIMA TIDAK DITERIMA

### Pohon 2

```

> ptgbag851<-predict(gbag85[["mtrees"]][[2]][["btree"]], newdata=IQC85.testing, type = "class")
> ptgbag851

```

9	18	24	29	30	33
DITERIMA	DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
38	44	46	48	49	60
DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA
79	93	108	112	113	116
DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
121	136	149	153	158	165
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	DITERIMA	DITERIMA
193	202	213	218	223	225
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA
229	230	233	239	247	248
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
254	262	267	274	276	288
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
291	306	313	327	328	329
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
337	341	349	357	359	360
DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
367	376	387	388		
TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA		

Levels: DITERIMA TIDAK DITERIMA

### Pohon 3

```

> ptgbag851<-predict(gbag85[["mtrees"]][[3]][["btree"]], newdata=IQC85.testing, type = "class")
> ptgbag851

```

9	18	24	29	30	33
DITERIMA	DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
38	44	46	48	49	60
DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA
79	93	108	112	113	116
DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
121	136	149	153	158	165
TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
193	202	213	218	223	225
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA
229	230	233	239	247	248
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
254	262	267	274	276	288
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
291	306	313	327	328	329
TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
337	341	349	357	359	360
DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
367	376	387	388		
TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA		

Levels: DITERIMA TIDAK DITERIMA

### Pohon 98

```

Levels: DITERIMA TIDAK DITERIMA
> ptgbag851<-predict(gbag85[["mtrees"]][[98]]["btree"], newdata=IQC85.testing, type = "class")
> ptgbag851

```

	9	18	24	29	30	33
	DITERIMA	DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	38	44	46	48	49	60
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA
	79	93	108	112	113	116
	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	121	136	149	153	158	165
TIDAK	DITERIMA	TIDAK DITERIMA	DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA
	193	202	213	218	223	225
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA
	229	230	233	239	247	248
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	254	262	267	274	276	288
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	291	306	313	327	328	329
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	337	341	349	357	359	360
	DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	367	376	387	388		
TIDAK	DITERIMA	DITERIMA	DITERIMA	DITERIMA		

Levels: DITERIMA TIDAK DITERIMA

### Pohon 99

```

> ptgbag851<-predict(gbag85[["mtrees"]][[99]]["btree"], newdata=IQC85.testing, type = "class")
> ptgbag851

```

	9	18	24	29	30	33
	DITERIMA	DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	38	44	46	48	49	60
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA
	79	93	108	112	113	116
	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	121	136	149	153	158	165
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA
	193	202	213	218	223	225
TIDAK	DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA
	229	230	233	239	247	248
	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA
	254	262	267	274	276	288
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	291	306	313	327	328	329
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	337	341	349	357	359	360
	DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	367	376	387	388		
TIDAK	DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA		

Levels: DITERIMA TIDAK DITERIMA

### Pohon 100

```

> ptgbag851<-predict(gbag85[["mtrees"]][[100]]["btree"], newdata=IQC85.testing, type = "class")
> ptgbag851

```

	9	18	24	29	30	33
	DITERIMA	DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	38	44	46	48	49	60
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA
	79	93	108	112	113	116
	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	121	136	149	153	158	165
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA
	193	202	213	218	223	225
TIDAK	DITERIMA	TIDAK DITERIMA	DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA
	229	230	233	239	247	248
	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	DITERIMA
	254	262	267	274	276	288
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	291	306	313	327	328	329
TIDAK	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	337	341	349	357	359	360
	DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA	TIDAK DITERIMA	TIDAK DITERIMA
	367	376	387	388		
TIDAK	DITERIMA	DITERIMA	TIDAK DITERIMA	DITERIMA		

Levels: DITERIMA TIDAK DITERIMA

**Lampiran 13:** Kombinasi Prediksi Data *Testing*

```
> table( IQC85.testing[,7],ptgbag85)
      ptgbag85
      DITERIMA TIDAK DITERIMA
DITERIMA      15             1
TIDAK DITERIMA 1             41
> |
```



## RIWAYAT HIDUP



Penulis bernama Dwi Miftakhul Jannah yang akrab disapa dengan dudu, merupakan anak dari pasangan Samsul dan Siti Romlah. Penulis merupakan anak bungsu dari dua bersaudara yang dilahirkan di Kota Mojokerto pada tanggal 30 November 1998. Pendidikan formal yang ditempuh penulis diawali dengan 2 tahun di TK Al-Baqiyatus Sholikhah, tahun 2004 hingga 2010 penulis belajar formal di SDN Sawahan, kemudian melanjutkan ke jenjang berikutnya di SMPN 1 Mojosari selama 3 tahun dan lulus pada tahun 2013, ditahun yang sama penulis melanjutkan pendidikan menengah keatas di MBI Amanatul Ummah hingga 2016. Pada tahun 2016, penulis melanjutkan di Perguruan Tinggi Negeri UIN Maulana Malik Ibrahim Malang dan diterima sebagai mahasiswa matematika dengan NIM 16610096. Semasa kuliah penulis aktif dalam organisasi intra kampus sebagai bendahara umum II HMJ Matematika pada tahun 2017-2018 dan beberapa organisasi extra kampus. Penulis memiliki pengalaman magang di BPS Kota Mojokerto dan di PT. Telkomsel Grapari Malang. Penulis menerima segala kritikan, saran dan masukan yang bersifat membangun demi manfaat tugas akhir ini. Oleh karena itu, penulis berkenan membantu pembaca terkait metode pada penelitian ini, pembaca dapat menghubungi penulis melalui email : [dwimj55@gmail.com](mailto:dwimj55@gmail.com).



**KEMENTERIAN AGAMA RI  
UNIVERSITAS ISLAM NEGERI  
MAULANA MALIK IBRAHIM MALANG  
FAKULTAS SAINS DAN TEKNOLOGI  
Jl. Gajayana No. 50 Dinoyo Malang  
Telp./Fax.(0341)558933**

**BUKTI KONSULTASI SKRIPSI**

Nama : Dwi Miftakhul Jannah  
NIM : 16610096  
Fakultas/Jurusan : Sains dan Teknologi/ Matematika  
Judul Skripsi : Klasifikasi Penerimaan Peserta Didik Baru di Madrasah Bertaraf International Amanatul Ummah Pacet Menggunakan *Bootstrap Aggregating Classification and Regression Tree*  
Pembimbing I : Heni Widayani, M.Si  
Pembimbing II : Angga Dwi Mulyanto, M.Si

No	Tanggal	Hal	Tanda Tangan
1.	7 Februari 2020	Revisi Bab 1,2 (Pembimbing 1)	1.
2.	14 Februari 2020	Konsultasi Bab 1,2 (Pembimbing 2)	2.
3.	11 Maret 2020	Revisi Bab 3,4 (Pembimbing 1)	3.
4.	13 Maret 2020	Konsultasi Bab 3,4 (Pembimbing 2)	4.
5.	26 Maret 2020	ACC Untuk Sempro (Pembimbing 2)	5.
6.	27 Maret 2020	ACC Untuk Sempro (Pembimbing 1)	6.
7.	20 April 2020	Konsultasi Bab 4,5 (Pembimbing 2)	7.
8.	22 April 2020	Konsultasi Bab 4,5 (Pembimbing 1)	8.
9.	27 April 2020	Konsultasi Abstrak (Pembimbing 2)	9.
10.	30 April 2020	ACC Untuk Sidang (Pembimbing 1)	10.
11.	30 April 2020	ACC Untuk Sidang (Pembimbing 2)	11.

Malang, 1 Juli 2020  
Mengetahui,  
Ketua Jurusan Matematika

Usman Pagalay, M.Si  
NIP. 19650414 200312 1 001