

**PENERAPAN METODE *CLASSIFICATION AND REGRESSION  
TREES* TERHADAP KLASIFIKASI PENERIMAAN  
PESERTA DIDIK BARU**  
(Studi Kasus : MAN 2 Kota Probolinggo)

**SKRIPSI**

**OLEH:  
LULUK CHELAWATI CHODIROH  
NIM. 17610088**



**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2023**

**PENERAPAN METODE *CLASSIFICATION AND REGRESSION  
TREES* TERHADAP KLASIFIKASI PENERIMAAN  
PESERTA DIDIK BARU**  
(Studi Kasus : MAN 2 Kota Probolinggo)

**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:  
LULUK CHELAWATI CHODIROH  
NIM. 17610088**

**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI MAULANA MALIK IBRAHIM  
MALANG  
2023**

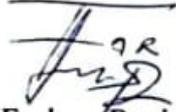
**PENERAPAN METODE *CLASSIFICATION AND REGRESSION TREES* TERHADAP KLASIFIKASI PENERIMAAN PESERTA DIDIK BARU**  
(Studi Kasus : MAN 2 Kota Probolinggo)

**SKRIPSI**

Oleh:  
**Luluk Chelawati Chodiroh**  
NIM. 17610088

Telah Disetujui Untuk Diuji  
Malang, 27 Desember 2023

Dosen Pembimbing I



Dr. Fachrud Rozi, M.Si.  
NIP. 19800527 200801 1 012

Dosen Pembimbing II



Erna Herawati, M.Pd.  
NIP. 19760723 202321 2 006

Mengetahui,  
Ketia Program Studi Matematika



Dr. Lily Susanti, M.Sc.  
NIP. 19741129 200012 2 005

**PENERAPAN METODE *CLASSIFICATION AND REGRESSION TREES* TERHADAP KLASIFIKASI PENERIMAAN PESERTA DIDIK BARU**  
(Studi Kasus : MAN 2 Kota Probolinggo)

**SKRIPSI**

Oleh:  
**Luluk Chelawati Chodiroh**  
NIM. 17610088

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

Tanggal 30 Desember 2023

Ketua Penguji : Prof. Dr. Hj. Sri Harini, M.Si.

Anggota Penguji 1 : Ria Dhea Layla Nur Karisma, M.Si.

Anggota Penguji 2 : Dr. Fachrur Rozi, M.Si.

Anggota Penguji 3 : Erna Herawati, M.Pd.



Mengetahui,  
Ketua Program Studi Matematika  
  
Dr. Elly Susanti, M.Sc.  
NIP. 19741129 200012 2 005



## PERNYATAAN KEASLIAN TULISAN

Saya yang bertanda tangan di bawah ini:

Nama : Luluk Chelawati Chodiroh

NIM : 17610088

Program Studi : Matematika

Fakultas : Sains dan Teknologi

Judul Skripsi : Penerapan Metode *Classification and Regression Trees*  
Terhadap Klasifikasi Penerimaan Peserta Didik Baru

Menyatakan dengan sebenarnya bahwa skripsi yang saya tulis ini benar-benar merupakan hasil karya sendiri, bukan merupakan 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, 30 Desember 2023

Yang membuat pernyataan,



Luluk Chelawati Chodiroh

NIM. 17610088

## **MOTO DAN PERSEMBAHAN**

### **MOTO**

“Keajaiban adalah nama lain dari kerja keras”

### **PERSEMBAHAN**

Skripsi ini penulis persembahkan untuk:

Umik dan Alm. Aba tersayang sebagai tanda bakti dan terimakasih yang tak terhingga.

Suami dan anak tercinta yang selalu menjadi penyemangat.

Seluruh keluarga besar dan sahabat-sahabat yang selalu memberikan dukungan dan semangat.

Diri sendiri, yang telah berjuang menghadapi kesulitan-kesulitan dalam menyelesaikan skripsi ini.

## KATA PENGANTAR

Dengan memanjatkan puji syukur kehadirat Allah SWT yang telah melimpahkan rahmat, taufik, serta hidayah-Nya sehingga penulis dapat menyelesaikan skripsi ini sebagai salah satu syarat untuk menyelesaikan Program Sarjana (S1) Jurusan Matematika Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim Malang.

Penulis menyadari bahwa skripsi ini tidak mungkin terselesaikan tanpa adanya dukungan, bimbingan, dan nasehat dari berbagai pihak selama penyusunan skripsi ini. Pada kesempatan ini penulis menyampaikan terima kasih setulus-tulusnya kepada:

1. Prof. Dr. H. M. Zainuddin, M.A., selaku rektor Universitas Islam Maulana Malik Ibrahim.
2. Prof. Dr. Hj. Sri Harini, M.Si., selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim.
3. Dr. Elly Susanti, M.Sc., selaku ketua Program Studi Matematika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim.
4. Dr. Fachrur Rozi, M.Si., selaku dosen pembimbing I yang telah membimbing dan memberikan arahan, nasehat, serta motivasi kepada penulis.
5. Erna Herawati, M.Pd., selaku dosen pembimbing II yang telah memberikan arahan dan ilmunya kepada penulis.
6. Prof. Dr. Hj. Sri Harini, M.Si., selaku Ketua Penguji dalam Ujian Skripsi.
7. Ria Dhea Layla Nur Karisma, M.Si., selaku Penguji I dalam Ujian Skripsi.
8. Seluruh dosen Program Studi Matematika, Fakultas Sains dan Teknologi, Universitas Islam Negeri Maulana Malik Ibrahim.
9. Kedua orang tua, suami, anak saya, kakak dan adik saya serta keluarga yang telah memberikan segala dukungan, kasih sayang, semangat, do'a, serta motivasi kepada penulis.
10. Seluruh teman-teman mahasiswa di Program Studi Matematika angkatan 2017.
11. Semua pihak yang tidak dapat disebutkan satu-persatu yang telah membantu penulis dalam menyelesaikan skripsi ini.

Semoga Allah SWT melimpahkan rahmat dan karunia-Nya kepada kita semua. Penulis berharap semoga skripsi ini dapat bermanfaat bagi penulis dan pembaca.

Malang, 30 Desember 2023

Luluk Chelawati Chodiroh

## DAFTAR ISI

<b>HALAMAN JUDUL</b> .....	<b>i</b>
<b>HALAMAN PENGANTAR</b> .....	<b>ii</b>
<b>HALAMAN PERSETUJUAN</b> .....	<b>iii</b>
<b>HALAMAN PENGESAHAN</b> .....	<b>iv</b>
<b>PERNYATAAN KEASLIAN TULISAN</b> .....	<b>v</b>
<b>MOTO DAN PERSEMBAHAN</b> .....	<b>vi</b>
<b>KATA PENGANTAR</b> .....	<b>vii</b>
<b>DAFTAR ISI</b> .....	<b>ix</b>
<b>DAFTAR TABEL</b> .....	<b>xi</b>
<b>DAFTAR GAMBAR</b> .....	<b>xii</b>
<b>DAFTAR SIMBOL</b> .....	<b>xiii</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xiv</b>
<b>ABSTRAK</b> .....	<b>xv</b>
<b>ABSTRACT</b> .....	<b>xvi</b>
<b>مستخلص البحث</b> .....	<b>xvii</b>
<b>BAB I PENDAHULUAN</b> .....	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	5
1.3 Tujuan Penelitian .....	5
1.4 Manfaat Penelitian .....	6
1.5 Batasan Masalah .....	6
1.6 Definisi Istilah .....	7
<b>BAB II KAJIAN TEORI</b> .....	<b>8</b>
2.1 Teori Pendukung .....	8
2.1.1 <i>Classification and Regression Trees (CART)</i> .....	8
2.1.2 Ukuran Ketepatan Klasifikasi .....	14
2.1.3 <i>Variable Importance</i> .....	16
2.1.4 Penerimaan Peserta Didik Baru .....	17
2.2 Kajian Integrasi Topik dengan Al-Qur'an/Hadits .....	18
2.3 Kajian Topik dengan Teori Pendukung .....	21
<b>BAB III METODE PENELITIAN</b> .....	<b>23</b>
3.1 Jenis Penelitian .....	23
3.2 Data dan Sumber Data .....	23
3.3 Lokasi Penelitian .....	25
3.4 Teknik Pengumpulan Data .....	25
3.5 Teknik Analisis Data .....	25
3.6 Flowchart .....	27
<b>BAB IV HASIL DAN PEMBAHASAN</b> .....	<b>28</b>
4.1 Analisis Data .....	28
4.2 Analisis <i>Classification and Regression Trees (CART)</i> .....	32
4.2.1 Penentuan Persentase Pembagian Data <i>Learning</i> dan Data <i>Testing</i> .....	32
4.2.2 Pembentukan Pohon Klasifikasi .....	33
4.2.3 Pemangkasan Pohon Klasifikasi Maksimal .....	36
4.2.4 Pohon Klasifikasi Optimal .....	37
4.2.5 Ketepatan Klasifikasi <i>Classification and Regression Trees</i> ...	39

4.3 Variabel Penting Pada <i>Classification and Regression Trees</i> .....	43
4.4 Kajian Islam pada Hasil Penelitian .....	46
<b>BAB V PENUTUP</b> .....	<b>48</b>
5.1 Kesimpulan .....	48
5.2 Saran .....	49
<b>DAFTAR PUSTAKA</b> .....	<b>50</b>
<b>LAMPIRAN</b> .....	<b>52</b>
<b>RIWAYAT HIDUP</b>	

## DAFTAR TABEL

Tabel 2.1 Tabel <i>Confusion Matrix</i> .....	14
Tabel 4.1 Deskriptif Karakteristik.....	30
Tabel 4.2 Ketepatan Klasifikasi Pada Kombinasi Persentase Data .....	32
Tabel 4.3 Pemilihan Pemilah Variabel Prediktor.....	33
Tabel 4.4 Urutan Variabel Terpenting .....	34
Tabel 4.5 Nilai <i>Complexity Parameter</i> .....	36
Tabel 4.6 Ketepatan Klasifikasi CART .....	40
Tabel 4.7 Nilai Ketepatan Klasifikasi CART .....	42
Tabel 4.8 Ketepatan Klasifikasi CART Menggunakan Data <i>Learning</i> .....	42
Tabel 4.9 <i>Variable Importance</i> .....	45

## DAFTAR GAMBAR

Gambar 2.1 Struktur Pohon Klasifikasi .....	9
Gambar 3.1 <i>Flow Chart Classification and Regression Trees</i> .....	27
Gambar 4.1 Diagram Lingkaran Jumlah Calon Peserta Didik Baru .....	28
Gambar 4.2 Perbandingan Berdasarkan Jenis Kelamin .....	29
Gambar 4.3 Perbandingan Berdasarkan Asal Sekolah .....	30
Gambar 4.4 Pohon Klasifikasi Maksimal .....	35
Gambar 4.5 Pohon Klasifikasi Optimal Penerimaan Peserta Didik Baru ..	38
Gambar 4.6 Kurva ROC.....	42
Gambar 4.7 Bar Plot <i>Variable Importance</i> .....	44

## DAFTAR SIMBOL

$i(t)$	: Indeks gini pada simpul $t$
$p(j t)$	: Proporsi kelas ke $j$ pada simpul $t$
$\Delta i(s, t)$	: Nilai <i>goodness of split</i>
$P_L$	: Nilai proporsi pengamatan simpul kiri
$P_R$	: Nilai proporsi pengamatan simpul kanan
$i(t_L)$	: Nilai fungsi heterogenitas pada simpul anak kiri
$i(t_R)$	: Nilai fungsi heterogenitas pada simpul anak kanan
$p(j_0 t)$	: Proporsi kelas $j$ pada simpul $t$
$N_j(t)$	: Jumlah pengamatan kelas $j$ pada simpul $t$
$N(t)$	: Jumlah pengamatan pada simpul $t$
$R_\alpha(T)$	: <i>Resubstitution</i> suatu pohon $T$ pada kompleksitas $\alpha$
$R(T)$	: <i>Resubstitution Estimate</i> (asumsi pengganti)
$\alpha$	: Parameter <i>cost complexity</i> bagi penambahan satu simpul akhir pohon $T$
$ \tilde{T} $	: Ukuran banyaknya simpul terminal pohon
TP	: Jumlah kasus positif, yang diklasifikasikan positif
FP	: Jumlah kasus negatif, yang diklasifikasikan positif
TN	: Jumlah kasus negatif, yang diklasifikasikan negatif
FN	: Jumlah kasus positif, yang diklasifikasikan negatif

## DAFTAR LAMPIRAN

Lampiran 1 Data Penerimaan Peserta Didik Baru .....	52
Lampiran 2 Statistik Deskriptif .....	59
Lampiran 3 Kombinasi Pembagian Data <i>Learning</i> dan Data <i>Testing</i> .....	60
Lampiran 4 Pembentukan Pohon Klasifikasi dan <i>Variable Importance</i> .....	62
Lampiran 5 Menghitung AUC dan Kurva ROC .....	63

## ABSTRAK

Chodiroh, Luluk Chelawati. 2023. **Penerapan Metode *Classification and Regression Trees Terhadap Klasifikasi Penerimaan Peserta Didik Baru***. Skripsi. Jurusan Matematika Fakultas Sains dan Teknologi. Universitas Islam Negeri Maulana Malik Ibrahim Malang. Pembimbing (1) Dr. Fachrur Rozi, M.Si. (2) Erna Herawati, M.Pd.

**Kata Kunci:** Penerimaan Peserta Didik Baru, Tingkat Akurasi, *Classification and Regression Trees (CART)*, *Variable importance*.

MAN 2 Kota Probolinggo merupakan sekolah terakreditasi A yang memiliki beberapa jurusan salah satunya yaitu Bahasa dengan program Bahasa Arab dan terdapat program akselerasi yang menjadi ikon sekolah, sehingga banyak peminat yang ingin belajar di MAN 2 Kota Probolinggo. Penelitian ini bertujuan untuk mengetahui tingkat akurasi klasifikasi dan variabel yang memiliki tingkat kepentingan tertinggi pada klasifikasi penerimaan peserta didik baru di MAN 2 Kota Probolinggo. Analisis ini menggunakan variabel respon diterima dan tidak diterima dengan variabel prediktor yaitu jenis kelamin, asal sekolah, nilai pengetahuan bahasa, nilai pengetahuan logika dan sains, serta nilai pengetahuan keagamaan dan kewarganegaraan. Metode yang digunakan pada penelitian ini yaitu *Classification and Regression Trees (CART)* dengan menghitung nilai *accuracy*, *sensitivity*, dan *specificity*. Tahap pertama dilakukan analisis deskriptif, selanjutnya membagi data menjadi data *learning* dan data *testing*. Kemudian pemilihan pemilah, membentuk pohon klasifikasi maksimal, pemangkasan pohon, penentuan pohon klasifikasi optimal serta mengukur tingkat *accuracy* dan menghitung tingkat kepentingan variabel. Hasil dari penelitian ini menunjukkan bahwa nilai *accuracy* CART sebesar 94,12% dengan *sensitivity* sebesar 88,89% dan *specificity* sebesar 100% serta nilai AUC yaitu 0,944 yang berarti bahwa model sangat baik dalam membedakan kelas diterima dan tidak diterima. *Variable importance* yang memiliki pengaruh terbesar dalam pohon klasifikasi yaitu nilai pengetahuan bahasa dengan tingkat kepentingan sebesar 34,76%.

## ABSTRACT

Chodiroh, Luluk Chelawati. 2023. **Application of the Classification and Regression Trees Method to the Classification of New Student Admissions**. Thesis. Mathematics Department. Faculty of Science and Technology. Universitas Islam Negeri Maulana Malik Ibrahim Malang. Advisors (1) Dr. Fachrur Rozi, M.Si. (2) Erna Herawati, M.Pd.

**Keyword:** New Student Admission, Level of Accuracy, Classification and Regression Trees (CART), Variable Importance.

MAN 2 Probolinggo is an A-accredited school which has several departments, one of which is language with an Arabic language program and there is an acceleration program which has become an icon for the school so that many interested people want to study at MAN 2 Probolinggo. This research aims to determine the level of classification accuracy and the variables that have the highest level of importance in the classification of new student admissions at MAN 2 Probolinggo. This analysis uses accepted and unaccepted response variables with predictor variables, namely gender, school origin, language knowledge scores, logic and science knowledge scores, as well as religious and civic knowledge scores. The method used in this research is Classification and Regression Trees (CART) by calculating accuracy, sensitivity and specificity values. The first stage carried out descriptive analysis, then divided the data into learning data and testing data. Then choose a classifier, form a maximum classification tree, prune the tree, determine the optimal classification tree and measure the level of accuracy and calculate the level of variable importance. The results of the research show that the CART accuracy value is 94.12% with a sensitivity of 88.89% and a specificity of 100% and an AUC value of 0.944, which means the model is very good at distinguishing accepted and not accepted classes. The variable of interest that has the greatest influence in the classification tree is the value of language knowledge with an importance level of 34.76%.

## مستخلص البحث

حاضره، لولوك حيلواتي. ٢٠٢٣. تطبيق أسلوب التصنيف وأشجار الانحدار على التصنيف القبول الطلاب الجدد. البحث الجامعي. قسم الرياضيات، كلية العلوم والتكنولوجيا، جامعة مولانا مالك إبراهيم الإسلامية الحكومية، مالانج. المشرف (1) د. فخر الرازي، الماجستير. (2) إيرنا هيراواتي، الماجستير.

الكلمات المفتاحية: قبول الطلاب الجدد، مستوى الدقة، شجرة التصنيف والانحدار (CART)، المتغيرات المهمة.

المدرسة الثانوية الحكومية الثانية بروبولينجو هي المدرسة المعتمدة من الدرجة أ ولها عدة الأقسام في الدراسة ومنها: برنامج اللغة العربية و برنامج التسريع حتى أصبح أيقونة المدرسة، لذلك يرغب كثير من المحتمسين في الدحول ألى المدرسة الثانوية الحكومية الثانية بروبولينجو. ويهدف هذا البحث إلى مستوى دقة التصنيف والمتغيرات التي لها أعلى مستوى من الأهمية في تصنيف قبول الطلاب الجدد بل مدرسة الثانوية الحكومية الثانية بروبولينجو. ويستخدم هذا التحليل بمتغيرات الاستجابة بين المقبولة والمردودة مع المتغير المتنبى وهو الجنس، وأصل المدرسة، وقيمة المعرفة اللغوية، وقيمة المعرفة المنطق والعلوم، وقيمة المعرفة الديني والمواطنة. الطريقة المستخدمة في هذا البحث هي شجرة التصنيف والانحدار (CART) عن طريق حساب قيمة حساسية الدقة والخصوصية. قامت المرحلة الأولى بالتحليل الوصفي، ثم تقسيم البيانات إلى بيانات التعلم وبيانات الاختبار. ثم اختيار المصنف وتشكيل شجرة التصنيف القصوى وتقليم الشجرة وتحديد شجرة التصنيف الأمثل وقياس مستوى الدقة وحساب مستوى الأهمية المتغيرة. تظهر نتائج هذا البحث أن قيمة دقة CART هي 94.12% مع حساسية 88.89% ونوعية 100% وقيمة AUC 0.944 مما يعني أن النموذج جيد جداً في التمييز بين الفئات المقبولة والمردودة. ومتغير المهمة ذو التأثير الأكبر في شجرة التصنيف هو قيمة المعرفة اللغوية بمستوى أهمية 34.76%.

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Pendidikan merupakan unit penting bagi kemajuan suatu bangsa, salah satunya di Indonesia. Sebagaimana diatur dalam Undang-Undang No. 20 tahun 2003 tentang Sistem Pendidikan Nasional yang memaparkan fungsi dan tujuan pendidikan dalam pasal 3 yakni mengembangkan kemampuan dan membentuk watak serta peradaban bangsa yang bermartabat dalam rangka mencerdaskan kehidupan bangsa, dan bertujuan untuk berkembangnya potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, dan menjadi warga negara yang demokratis serta bertanggung jawab. Pengertian pendidikan menurut Thompson (dalam Neolaka, 2017) yaitu pengaruh lingkungan terhadap individu untuk menghasilkan perubahan-perubahan yang tetap dalam kebiasaan perilaku, pikiran serta sifat.

Pendidikan memiliki peranan yang sangat penting dalam kehidupan bahkan Al-Qur'an memberi isyarat mengenai pentingnya pendidikan serta meninggikan derajat bagi manusia yang mempunyai pengetahuan sebagaimana firman Allah SWT dalam QS. Al-Mujadalah:11 yang menyebutkan:

... وَإِذَا قِيلَ انشُرُوا فَانشُرُوا يَرْفَعِ اللَّهُ الَّذِينَ آمَنُوا مِنْكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ...  
(المجادلة)

*“...Dan apabila dikatakan, “Berdirilah kamu” maka berdirilah, niscaya Allah akan mengangkat (derajat) orang-orang yang beriman di antaramu dan orang-orang yang diberi ilmu beberapa derajat...” (QS. Al-Mujadalah:11)*

Ayat di atas menunjukkan bahwa Allah SWT mengangkat derajat orang yang beriman dan berilmu pengetahuan, yang mana derajat tersebut berarti kedudukan

atau kelebihan dari makhluk lainnya. Hanya Allah SWT yang mengetahui mengenai bentuk dan jenisnya dan siapa yang akan ditinggikan derajatnya. Mencari ilmu juga merupakan kewajiban bagi setiap muslim sebagaimana sabda Rasulullah SAW yang berbunyi:

حَدَّثَنَا هِشَامُ بْنُ عَمَّارٍ حَدَّثَنَا حَفْصُ بْنُ سُلَيْمَانَ حَدَّثَنَا كَثِيرُ بْنُ شَنْظِيرٍ عَنْ مُحَمَّدِ بْنِ سِيرِينَ عَنْ أَنَسِ بْنِ مَالِكٍ قَالَ قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ طَلَبُ الْعِلْمِ فَرِيضَةٌ عَلَى كُلِّ مُسْلِمٍ وَوَاضِعُ الْعِلْمِ عِنْدَ غَيْرِ أَهْلِهِ كَمُقَدِّدِ الْخَنَازِيرِ الْجَوْهَرَ وَاللُّؤْلُؤَ وَالذَّهَبَ  
(رواه ابن ماجه)

*“Hisyam bin Ammar telah menceritakan kepada kami, ia berkata Hafs bin Sulaiman telah menceritakan kepada kami, ia berkata Katsir bin Syinzir telah menceritakan kepada kami dari Muhammad bin Sirin dari Anas bin Malik ia berkata : Rasulullah SAW bersabda: “Mencari ilmu adalah kewajiban setiap muslim. Dan orang yang meletakkan ilmu bukan pada ahlinya, seperti seorang yang mengalungkan mutiara, intan dan emas ke leher babi.”” (HR Ibnu Majah)*

Hadits tersebut menjelaskan bahwa mencari ilmu itu wajib bagi setiap muslim, baik laki-laki maupun perempuan. Kewajiban tersebut yang berarti apabila dikerjakan mendapat pahala dan apabila ditinggalkan akan mendapat dosa.

Di Indonesia pemerintah memiliki program wajib belajar 12 tahun, untuk mensukseskan program tersebut diadakan penerimaan peserta didik baru setiap tahunnya. Penerimaan peserta didik baru adalah kegiatan instansi pendidikan yang melakukan penerimaan peserta didik baru untuk menyaring calon peserta didik baru yang akan mendaftar ke sekolah yang dituju dan biasanya dilaksanakan pada awal tahun ajaran baru. Peserta didik baru harus memenuhi kriteria yang ditentukan oleh sekolah agar dapat diterima di sekolah yang dituju (Farlina & Hudin, 2017).

Penerimaan peserta didik baru didasarkan oleh beberapa hal, yang pertama yaitu kebijakan penerimaan peserta didik, yang mana seorang peserta didik yang diterima harus memenuhi persyaratan-persyaratan yang telah ditentukan. Kedua

yaitu sistem seleksi penerimaan peserta didik baru. Terdapat dua sistem seleksi penerimaan peserta didik baru yang sering digunakan yaitu sistem seleksi dan promosi. Ketiga yaitu kriteria penerimaan peserta didik baru. Kriteria tersebut digunakan sebagai penentu peserta didik diterima atau tidak diterima di suatu sekolah. Terakhir yaitu prosedur penerimaan peserta didik baru (Suking, 2015).

Salah satu sekolah yang setiap tahunnya melaksanakan penerimaan peserta didik baru yaitu Madrasah Aliyah Negeri atau MAN 2 Kota Probolinggo. MAN 2 Kota Probolinggo merupakan sekolah yang sudah terakreditasi A. Terdapat beberapa jurusan di MAN 2 Kota Probolinggo yaitu Bahasa (dengan program Bahasa Arab), IPA, IPS, dan Agama. Di samping itu terdapat program kelas Akselerasi yang menjadi satu-satunya ikon sekolah, khususnya di Kota Probolinggo. Seiring berjalannya waktu, MAN 2 Kota Probolinggo mendapatkan banyak prestasi dan penghargaan. Berdasarkan profil MAN 2 Kota Probolinggo menunjukkan bahwa sekolah tersebut berprestasi di berbagai bidang, sehingga banyak peminat yang ingin belajar di MAN 2 Kota Probolinggo.

Terdapat beberapa faktor yang dapat mempengaruhi penerimaan peserta didik baru di MAN 2 Kota Probolinggo antara lain yaitu nilai Ujian Nasional (UN), nilai raport, nilai *Intelligence Quotient* (IQ), dan asal sekolah peserta didik. Sehingga untuk meneliti pengaruh tersebut, dapat menggunakan salah satu metode statistika. Metode statistika yang dapat menggambarkan hubungan antara variabel dependen dengan satu atau pun lebih dari satu variabel independen yaitu metode CART (*Classification and Regression Trees*) (Hartati, 2012).

Metode CART merupakan salah satu metode statistika nonparametrik yang berguna untuk mendapatkan kelompok data yang akurat dan digunakan sebagai

penentu dari proses klasifikasi. Metode tersebut memiliki kelebihan yaitu hasil yang diperoleh mudah diinterpretasikan, lebih akurat dan perhitungannya cepat, tetapi juga memiliki kekurangan yaitu apabila terjadi perubahan data *learning* maka pohon klasifikasi akan mengalami perubahan besar atau tidak stabil yang berpengaruh pada hasil akurasi prediksi (Suwardika, dkk., 2019).

Penelitian menggunakan metode CART telah dilakukan oleh Susiana, dkk. (2019) untuk mengklasifikasi korban kecelakaan lalu lintas di Kota Palu yang menghasilkan nilai akurasi pada CART sebesar 77,9%. Pada penelitian tersebut juga menghasilkan nilai *sensitivity* dan *specificity*. Nilai *sensitivity* yang diperoleh pada metode CART yaitu sebesar 99,20% sedangkan nilai *specificity*nya yaitu 16,19%.

Klasifikasi menggunakan metode CART juga pernah dilakukan oleh Suwardika, dkk. (2019) pada ketidaktepatan waktu kelulusan mahasiswa Universitas Terbuka yang menghasilkan nilai akurasi pada CART sebesar 65,41% dengan nilai *sensitivity* dan *specificity* sebesar 18,87% dan 96,25%. Pada penelitian yang dilakukan oleh Arrahimi, dkk. (2019) untuk klasifikasi masa studi mahasiswa menghasilkan nilai akurasi menggunakan metode CART yaitu sebesar 75,592%. Hasil tersebut menunjukkan bahwa CART memiliki nilai akurasi yang baik dalam klasifikasi.

Berdasarkan beberapa uraian di atas, hal tersebut melatarbelakangi penulis melakukan penelitian mengenai penerapan metode *Classification and Regression Trees* (CART) yang diaplikasikan pada faktor-faktor yang dapat mempengaruhi penerimaan peserta didik baru di MAN 2 Kota Probolinggo. Metode yang digunakan dalam penelitian ini adalah metode *Classification and Regression*

*Trees* (CART) yang hasilnya dapat digunakan sebagai penentu dari proses klasifikasi. Hasil dari penelitian ini diharapkan dapat memberikan kemudahan dalam proses klasifikasi penerimaan peserta didik baru di MAN 2 Kota Probolinggo.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang tersebut diperoleh rumusan masalah sebagai berikut:

1. Bagaimana hasil penerapan dan tingkat akurasi metode *Classification and Regression Trees* pada klasifikasi penerimaan peserta didik baru?
2. Bagaimana tingkat kepentingan variabel-variabel yang mempengaruhi penerimaan peserta didik baru menggunakan metode *Classification and Regression Trees*?

## **1.3 Tujuan Penelitian**

Berdasarkan rumusan masalah tersebut diperoleh tujuan penelitian sebagai berikut:

1. Mendeskripsikan hasil penerapan dan tingkat akurasi metode *Classification and Regression Trees* pada klasifikasi penerimaan peserta didik baru.
2. Mendeskripsikan tingkat kepentingan variabel-variabel yang mempengaruhi penerimaan peserta didik baru menggunakan metode *Classification and Regression Trees*.

#### 1.4 Manfaat Penelitian

Berdasarkan rumusan masalah dan tujuan penelitian, maka manfaat pada penelitian ini adalah:

1. Penelitian ini berguna untuk memberikan informasi kepada masyarakat yang berminat untuk mendaftar di MAN 2 Kota Probolinggo terkait klasifikasi penerimaan peserta didik baru. Selain itu, dengan menggunakan *Classification and Regression Trees* diharapkan dapat memberi kemudahan dalam proses penerimaan peserta didik baru di MAN 2 Kota Probolinggo.
2. Penelitian ini diharapkan dapat menambah wawasan dan pengetahuan mengenai *Classification and Regression Trees*.

#### 1.5 Batasan Masalah

Batasan masalah pada penelitian ini sebagai berikut:

1. Data yang digunakan merupakan data penerimaan peserta didik baru jalur reguler di MAN 2 Kota Probolinggo tahun 2020 dengan variabel prediktor yang diduga mempengaruhi yaitu jenis kelamin, asal sekolah, nilai pengetahuan Bahasa, nilai pengetahuan Logika dan Sains, serta nilai pengetahuan Keagamaan dan Kewarganegaraan.
2. Metode yang digunakan yaitu *Classification and Regression Trees* dengan menggunakan *software R*.

## 1.6 Definisi Istilah

Penelitian ini memiliki batasan-batasan pengertian supaya tidak terjadi kesalahpahaman dalam penafsiran istilah yaitu sebagai berikut:

1. *Accuracy* adalah suatu ukuran untuk mengukur seberapa tepat suatu model klasifikasi dalam memprediksi kelas data
2. *Sensitivity* adalah ukuran untuk mendeteksi sebagian besar positif yang sebenarnya dalam dataset
3. *Specitifity* adalah ukuran untuk mendeteksi sebagian besar negatif yang sebenarnya dalam dataset
4. *Data Testing* adalah data yang digunakan untuk mengukur kebaikan atau kelayakan model
5. *Data Learning* adalah data pelatihan yang digunakan untuk melatih CART
6. *Root Node* adalah simpul utama yang menjadi titik awal pada pohon klasifikasi
7. *Internal Node* adalah simpul dalam pada pohon klasifikasi
8. *Terminal Node* adalah simpul akhir pada pohon klasifikasi

## BAB II

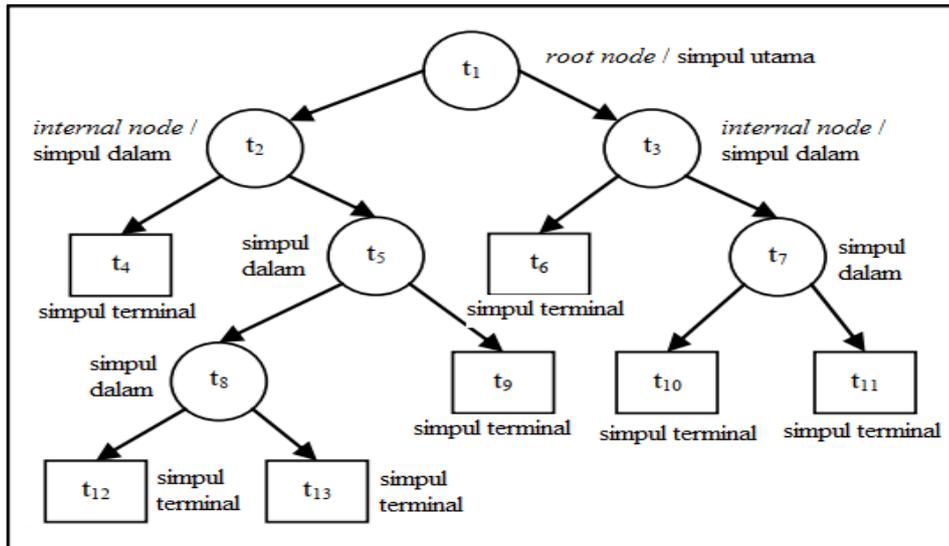
### KAJIAN TEORI

#### 2.1 Teori Pendukung

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

*Classification and Regression Trees (CART)* merupakan salah satu algoritma *decision tree*. Terdapat beberapa algoritma dalam pembentukan *decision tree* diantaranya yaitu C5.0, dan CART. Algoritma C5.0 merupakan pohon keputusan non biner yang mana cabang pohon bisa lebih dari dua, sedangkan CART merupakan pohon keputusan biner yang mana cabang pohon hanya terdiri dari dua cabang. Pada algoritma C5.0 pemilihan atribut diproses dengan menggunakan ukuran *gain ratio*, sedangkan pada CART menggunakan *indeks gini*. Metode CART menganalisis klasifikasi pada variabel dependen yang berupa nominal, ordinal, atau kontinu. CART akan menghasilkan pohon klasifikasi atau pohon regresi. Jika variabel dependennya kategorik, maka akan menghasilkan pohon klasifikasi. Sedangkan apabila variabel dependennya kontinu, maka akan memperoleh suatu pohon regresi (Arrahmi, dkk, 2019).

Metode CART sudah banyak diterapkan dalam penyelesaian masalah di berbagai bidang dan mampu mengatasi berbagai kondisi data. Adapun kelebihan dari metode tersebut yaitu mampu mengeksplorasi data berdimensi tinggi, hasil yang diperoleh mudah diinterpretasi, serta dapat digunakan pada data kontinu maupun kategorik (Suwardika, dkk., 2019). Pada penelitian ini CART yang digunakan yaitu data kategorik yang akan menghasilkan pohon klasifikasi. Berikut ini adalah struktur pohon klasifikasi.



**Gambar 2.1** Struktur Pohon Klasifikasi

CART juga dapat dikatakan sebagai metode pemilahan *rekursif biner*, yang berarti bahwa sekelompok data dalam simpul pohon dipilah menjadi dua simpul anak, dari masing-masing simpul anak kemudian dipilah menjadi dua anak simpul lagi dan berlanjut seperti itu sampai berhenti karena kriteria tertentu (Prianti, 2020). Pemilahan data dipilah menjadi dua, yaitu kelompok simpul kiri dan kanan kemudian dilakukan pada tiap simpul sampai diperoleh simpul terminal atau yang disebut simpul akhir.

Terdapat beberapa simpul dalam pohon klasifikasi, diantaranya yaitu simpul utama (*root node*), simpul dalam (*internal node*), dan simpul terminal (*terminal node*) yang dapat dilihat pada gambar 2.1. Simpul utama atau yang bisa disebut dengan *root node* dinotasikan sebagai  $t_1$  dan simpul dalam atau *internal node* dinotasikan sebagai  $t_2, t_3, t_5, t_7$  dan  $t_8$ . Adapun simpul terminal atau yang disebut dengan *terminal node* dinotasikan sebagai  $t_4, t_6, t_9, t_{10}, t_{11}, t_{12}$  dan  $t_{13}$  yang mana tidak ada lagi pemilahan. Menghitung kedalaman pohon dimulai dari simpul utama atau  $t_1$  yang berada pada kedalaman 1,  $t_2$  dan  $t_3$  ada pada kedalaman 2, begitu pula seterusnya sampai pada simpul akhir  $t_{12}$  dan  $t_{13}$  yang berada pada

kedalaman 5 (Pratiwi & Zain, 2014). Adapun tahapan penerapan Algoritma CART sebagai berikut:

### 1. Pembentukan Pohon Klasifikasi

Tahap awal dalam pembentukan pohon klasifikasi yakni menentukan variabel pemilah yang mana nilai dari variabel tersebut atau yang disebut dengan nilai *threshold* akan dijadikan pemilah pada tiap simpul. Pembentukan pohon klasifikasi terdiri dari tiga tahap sebagai berikut (Breiman, 1998):

#### a. Pemilahan

Pemilahan pada tiap simpul dilakukan untuk memperoleh pemilah yang memiliki nilai variabel dengan tingkat homogen yang paling besar. Pemilihan menggunakan sampel data *learning* yang masih bersifat heterogen. Berdasarkan kriteria *goodness of split*, proses pemilahan dilakukan untuk mengurangi keheterogenan pada simpul utama (*root node*) dan memaksimalkan kehomogenan pada simpul dalam (*internal node*).

Suatu metode pemilahan yang dapat mengukur tingkat keheterogenan kelas suatu simpul dalam pohon klasifikasi yaitu dengan menggunakan metode *impurity measure*  $i(t)$ . Metode tersebut dapat menemukan fungsi pemilah paling optimal yaitu dengan menggunakan indeks gini. Data *impurity* dapat didefinisikan dengan menggunakan ukuran varians. Misalkan 1 adalah pengamatan pada simpul  $t$  kelas ke- $j$  dan 0 untuk lainnya. Estimasi varian untuk simpul  $t$  pada pengamatan sebagai berikut ini (Breiman, 1998):

$$p(j|t)(1 - p(j|t)) \quad (2.1)$$

Indeks gini pada simpul  $t$  kelas ke- $j$  dapat dituliskan sebagai berikut:

$$i(t) = \sum_{j=1}^k p(j|t)(1 - p(j|t))$$

$$\begin{aligned}
&= \sum_{j=1}^k p(j|t) - p^2(j|t) \\
&= \sum_{j=1}^k p(j|t) - \sum_{j=1}^k p^2(j|t) \\
&= 1 - \sum_{j=1}^k p^2(j|t) \tag{2.2}
\end{aligned}$$

dimana:

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

$p(j|t)$  : Proporsi kelas ke  $j$  pada simpul  $t$

*Goodness of split* adalah suatu pengevaluasian pemilahan  $s$  pada simpul  $t$  sebagai penurunan keheterogenan. Misal pemilahan  $s$  yang memilah simpul  $t$  menjadi 2 simpul yaitu simpul kiri dan simpul kanan. Simpul kiri  $t_L$  dengan proporsi  $P_L$  dan simpul kanan  $t_R$  dengan proporsi  $P_R$  yang mana diketahui  $P_R = 1 - P_L$ . Maka *goodness of split* dapat didefinisikan sebagai berikut (Andriyashin, 2005):

$$\Delta i(s, t) = i(t) - P_L \cdot i(t_L) - P_R \cdot i(t_R) \tag{2.3}$$

dimana

$\Delta i(s, t)$  : Nilai *goodness of split*

$i(t)$  : Nilai fungsi heterogenitas pada simpul  $t$

$P_L$  : Nilai proporsi pengamatan simpul kiri

$P_R$  : Nilai proporsi pengamatan simpul kanan

$i(t_L)$  : Nilai fungsi heterogenitas pada simpul anak kiri

$i(t_R)$  : Nilai fungsi heterogenitas pada simpul anak kanan

Pemilahan yang memperoleh hasil  $\Delta i(s, t)$  tertinggi adalah pemilah terbaik. Pengembangan pohon dilakukan pada simpul utama  $t_1$  lalu dipilah menjadi  $t_2$  dan  $t_3$  dan seterusnya.

### b. Penentuan Simpul Terminal

Penentuan simpul  $t$  dapat menjadi simpul terminal atau tidak dengan cara dipilah, apabila pada simpul  $t$  sudah tidak ada penurunan keheterogenan. Adapun kriteria pada penentuan simpul  $t$  dapat menjadi simpul terminal yaitu jika pada simpul  $t$  terdapat  $n > 5$  dan pohon yang dibentuk akan berhenti jika telah mencapai level yang ditentukan pada kriteria pohon maksimal (Breiman, 1998).

### c. Penandaan Label Kelas

Sesuai aturan jumlah terbanyak dilakukanlah penandaan label kelas pada simpul terminal. Adapun label kelas simpul terminal  $t$  yaitu  $j_0$  yang mana memberi nilai dugaan kesalahan pengklasifikasian simpul  $t$  terbesar. Pembentukan pohon klasifikasi akan berhenti apabila ada satu pengamatan dalam tiap simpul anak atau jika terdapat batasan minimum  $n$ , semua pengamatan dalam tiap simpul anak identik, serta terdapat batasan jumlah level atau pun kedalaman pohon maksimal.

$$p(j_0|t) = \max_j p(j|t) = \max_j \frac{N_j(t)}{N(t)} \quad (2.4)$$

dimana:

$p(j_0|t)$  : Proporsi kelas  $j$  pada simpul  $t$

$N_j(t)$  : Jumlah pengamatan kelas  $j$  pada simpul  $t$

$N(t)$  : Jumlah pengamatan pada simpul  $t$

## 2. Pemangkasan Pohon Klasifikasi

Pohon klasifikasi yang optimal diperoleh dari pemangkasan pohon yang kurang penting. Adapun ukuran pemangkasan *Cost complexity minimum*

digunakan untuk memperoleh ukuran pohon yang layak. Ukuran *cost complexity* dengan  $\alpha > 0$  menggunakan berikut ini:

$$R_\alpha(T) = R(T) + \alpha |\tilde{T}| \quad (2.5)$$

dimana:

$R_\alpha(T)$  : *Resubstitution* suatu pohon  $T$  pada kompleksitas  $\alpha$

$R(T)$  : *Resubstitution Estimate* (asumsi pengganti)

$\alpha$  : Parameter *cost complexity* bagi penambahan satu simpul akhir pohon  $T$

$|\tilde{T}|$  : Ukuran banyaknya simpul terminal pohon

### 3. Penentuan Pohon Klasifikasi Optimal

Ukuran pohon yang besar dapat berakibat pada nilai kompleksitas yang tinggi karena struktur data yang cenderung kompleks, sehingga harus dipilih pohon optimal yang berukuran sederhana dan memberikan nilai penduga pengganti cukup kecil. Terdapat dua jenis penduga pengganti yaitu penduga sampel uji atau disebut *test sample estimate* dan penduga validasi silang lipat  $V$  atau dikenal dengan *cross validation V-fold estimate* (Pratiwi & Zain, 2014).

*Cross Validation V-Fold* atau penduga validasi silang lipat  $V$  merupakan salah satu penduga pengganti yang sering digunakan untuk jumlah sampel yang kecil atau kurang dari 3000 (Pratiwi dan Zain, 2014). Data pengamatan dibagi secara acak dengan ukuran yang sama pada setiap bagian (*fold*) dan saling lepas. Pada prosedur pembuatan pohon klasifikasi, salah satu dari setiap bagian (*fold*) yang sudah terbentuk akan menjadi data *learning* dan lainnya akan menjadi data *testing*. Setelah terbentuk pohon sebanyak *V-Fold* pengulangan data, kemudian menghitung rata-rata klasifikasi masing-masing bagian (*fold*) (Lewis, 2000).

Metode *cross validation* dapat mengatasi permasalahan *overlapping* pada data *testing*. Pada metode tersebut nilai *V* yang sering digunakan yaitu 10 maka menjadi *10 fold cross validation*. Nilai tersebut merupakan nilai terbaik karena dapat menghasilkan estimasi error yang rendah. *Cross Validation V-Fold* dapat membuat pembagian data menjadi *balance*. Nilai akurasi yang diperoleh dari data yang *balance* lebih akurat (Ian H. Witten, 2011). Sehingga nilai *10 fold cross validation* digunakan sebagai standar klasifikasi pembagian data *learning* dan data *testing*.

### 2.1.2 Ukuran Ketepatan Klasifikasi

Mengevaluasi model dapat menggunakan metode *Confusion matrix* dan kurva *Receiver Operating Characteristics (ROC)*. *Confusion matrix* dapat memberikan rincian pada klasifikasi, kelas yang ditampilkan di bagian atas matrix yaitu kelas yang diprediksi sedangkan kelas yang diobservasi ditampilkan di bagian kiri. Adapun evaluasi model *Confusion matrix* menggunakan tabel di bawah ini:

**Tabel 2.1** Tabel *Confusion Matrix*

		<i>Actual Class</i>	
		<i>Yes</i>	<i>No</i>
<i>Predicted Class</i>	<i>Yes</i>	<i>True Positive (TP)</i>	<i>False Positive (FP)</i>
	<i>No</i>	<i>False Negative (FN)</i>	<i>True Negative (TN)</i>

Berikut ini adalah rumus untuk menghitung akurasi:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (2.9)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (2.10)$$

$$Specificity = \frac{TN}{TN+FP} \quad (2.11)$$

dimana:

$TP$  = Jumlah kasus positif, yang diklasifikasikan positif

$FP$  = Jumlah kasus negatif, yang diklasifikasikan positif

$TN$  = Jumlah kasus negatif, yang diklasifikasikan negatif

$FN$  = Jumlah kasus positif, yang diklasifikasikan negatif

(Bisri, 2015).

Kurva *Receiver Operating Characteristics* (ROC) merupakan suatu metode yang berfungsi untuk membuat suatu visualisasi dan menyeleksi model klasifikasi terbaik berdasarkan tingkat performansinya. Pada kurva ROC, *True Positive* (TP) *rate* diplot pada sumbu Y sedangkan *False Positive* (FP) *rate* diplot pada sumbu X. ROC dapat dinyatakan sebagai klasifikasi prediksi yang mendekati atau menjadi negatif apabila semakin rendah titik ke kiri (0,0), sedangkan jika semakin ke atas titik ke kanan (1,1) maka dinyatakan sebagai klasifikasi prediksi yang mendekati atau menjadi positif. Untuk titik yang bernilai 1 maka dinyatakan sebagai tingkat *True Positive* (TP), sedangkan titik yang bernilai 0 maka dinyatakan sebagai tingkat *False Positive* (FP). Titik (0,1) adalah klasifikasi prediksi yang sempurna karena semua kasus baik positif maupun negatif dinyatakan dengan benar, sedangkan titik (1,0) adalah klasifikasi prediksi yang dinyatakan tidak benar (Bisri, 2015).

ROC memiliki suatu area yang dinamakan *Area Under Curve* (AUC) yang berguna untuk membandingkan perfomansi dari beberapa model klasifikasi untuk mengetahui model yang terbaik. Nilai AUC berkisar antara 0 sampai dengan 1 yang mana apabila mendekati 1 maka dapat dikatakan model mampu

mengklasifikasi dengan baik. Adapun perhitungan AUC didefinisikan sebagai berikut:

$$AUC = \frac{1}{2} \left( \frac{TP}{TP+FN} + \frac{TN}{TN+FP} \right) \quad (2.12)$$

Nilai dalam pengklasifikasian AUC dibagi dalam beberapa kelompok yaitu sebagai berikut:

- a. 0.90 - 1.00 = Klasifikasi sangat baik
- b. 0.80 – 0.90 = Klasifikasi baik
- c. 0.70 – 0.80 = Klasifikasi cukup
- d. 0.60 – 0.70 = Klasifikasi buruk
- e. 0.50 – 0.60 = Klasifikasi salah

(Sari, dkk, 2020).

### 2.1.3 Variable Impotance

Pada penerimaan peserta didik baru terdapat beberapa indikator (variabel) yang dapat mempengaruhi diterima atau tidak diterimanya peserta didik baru. Salah satu variabel yang dapat mempengaruhi klasifikasi penerimaan peserta didik baru yaitu asal sekolah peserta didik, jenis kelamin, dan nilai pengetahuan dari peserta didik. Setelah melakukan klasifikasi pada *Classification and Regression Trees* (CART) akan diketahui variabel yang memiliki pengaruh besar pada proses pengklasifikasian yang disebut dengan variabel penting.

Variabel penting atau yang dikenal dengan *variable importance* merupakan variabel yang memiliki dampak besar atau yang paling berpengaruh dalam suatu klasifikasi. Variabel penting dapat diperoleh dari nilai pada data *overall* yang

diperoleh dari pohon klasifikasi yang dibagi dengan keseluruhan data (Bobbitt, 2020).

$$VarImp = \left( \frac{\text{nilai overall tiap variabel}}{\text{total nilai overall}} \right) \times 100\%$$

#### **2.1.4 Penerimaan Peserta Didik Baru**

Berdasarkan Undang-Undang Republik Indonesia Nomor 20 tahun 2003 tentang Sistem Pendidikan Nasional yang diatur dalam pasal 1 memaparkan pengertian dari peserta didik yaitu anggota masyarakat yang berusaha mengembangkan potensi diri melalui proses pendidikan pada jalur, jenjang, dan jenis pendidikan tertentu. Peserta didik merupakan anggota masyarakat yang berhak mendapatkan pendidikan formal maupun non-formal dari jenjang pendidikan usia dini, pendidikan dasar, pendidikan menengah, dan pendidikan tinggi (Prastiyo, 2019).

Penerimaan peserta didik baru memiliki kebijakan operasional yang berisi aturan mengenai jumlah peserta didik yang bisa diterima oleh suatu sekolah. Terdapat beberapa faktor dalam penentuan jumlah peserta didik diantaranya yaitu daya tampung kelas, kriteria peserta didik baru yang dapat diterima, anggaran yang tersedia, sarana dan prasarana, tenaga kependidikan, dan lain sebagainya. Kebijakan penerimaan peserta didik baru dilaksanakan berdasarkan petunjuk dari Dinas Pendidikan (Kristiawan, 2017).

Terdapat dua sistem yang biasanya digunakan dalam penerimaan peserta didik baru yaitu sistem promosi dan sistem seleksi. Sistem promosi ialah penerimaan peserta didik baru tanpa melakukan seleksi, peserta didik yang mendaftar di suatu sekolah secara langsung dinyatakan lolos atau diterima.

Sedangkan sistem seleksi ialah penerimaan peserta didik baru yang melakukan seleksi terlebih dahulu, seperti seleksi berdasarkan nilai epta murni, penelusuran minat bakat dan kemampuan, dan melalui tes masuk.

Kriteria yang menentukan peserta didik baru diterima atau tidaknya ada tiga macam yang pertama yaitu kriteria acuan yang merupakan suatu penerimaan peserta didik baru berdasarkan acuan yang telah ditentukan sebelumnya. Kedua yaitu kriteria acuan norma yang merupakan suatu penerimaan peserta didik yang berdasarkan pada keseluruhan prestasi peserta didik yang mengikuti seleksi. Terakhir yaitu kriteria berdasarkan daya tampung sekolah (Kristiawan, 2017).

## **2.2 Kajian Integrasi Topik dengan Al-Qur'an/Hadits**

Pendidikan merupakan usaha secara sadar serta terencana dalam mewujudkan proses pembelajaran supaya peserta didik secara aktif mengembangkan potensi diri yang memiliki kepribadian, kecerdasan, akhlak mulia serta keterampilan yang diperlukan untuk diri sendiri, masyarakat, bangsa, dan negara. Menurut Ki Hajar Dewantara pendidikan adalah menuntun segala kekuatan kodrat agar mereka dapat mencapai keselamatan dan kebahagiaan setinggi-tingginya sebagai manusia dan masyarakat (Marbun, 2018). Adapun upaya agar manusia mengenal, memahami, mempercayai ajaran islam serta menghormati agama lain dalam hubungan antar umat beragama agar tercipta persatuan dan kesatuan bangsa merupakan pengertian pendidikan islam (Rohman & Hairudin, 2018).

Pendidikan islam ialah suatu proses edukatif yang mengarahkan manusia kepada terbentuknya kepribadian yang sesuai nilai-nilai agama islam. Pendidikan Islam menurut Hasan Langgulung merupakan suatu proses spiritual, akhlak, intelektual serta sosial yang dapat membimbing manusia dan memberi nilai-nilai,

prinsip-prinsip serta teladan dalam kehidupan untuk mempersiapkan kehidupan dunia dan akhirat. Sedangkan menurut Moh. Fadil Al-Djamali pendidikan islam ialah suatu proses yang mengarahkan manusia menuju kehidupan yang baik serta mengangkat derajat manusia sesuai kemampuan dasar dan kemampuan ajar (Mulyasana, dkk, 2020).

Tujuan pendidikan islam menurut Al-Ghazali yaitu membentuk insan purna yang bertujuan mendekatkan diri kepada Allah SWT menuju kebahagiaan hidup di dunia maupun di akhirat. Menurut Al-Ghazali tujuan pendidikan islam yakni kesempurnaan manusia di dunia dan di akhirat. Manusia bisa mencapai kesempurnaan itu melalui ilmu, dengan ilmu manusia akan memperoleh kebahagiaan di dunia dan kebahagiaan yang hakiki dengan mendekatkan diri kepada Allah SWT (Rohman & Hairudin, 2018).

Adapun tujuan dari pendidikan ialah menciptakan seseorang yang berkarakter dan berkualitas agar mempunyai pandangan yang luas ke depannya untuk mencapai cita-cita yang diinginkan dan mampu beradaptasi di berbagai lingkungan dengan tepat (Marbun, 2018). Menurut islam tujuan pendidikan sendiri yaitu terwujudnya muslim yang *kaffah*. Muslim yang *kaffah* merupakan muslim yang sehat, kuat jasmani, cerdas, dan hatinya dipenuhi iman kepada Allah SWT. Beriman kepada Allah SWT, mengikuti perintah-Nya dan menjauhi larangan-Nya secara penuh, bukan hanya setengah-setengah atau mencampuradukkan dengan agama lain merupakan pengertian dari islam yang *kaffah*.

Terdapat beberapa sifat dan corak dalam ilmu pendidikan islam. Pertama yaitu corak normatif yakni ilmu pendidikan yang berbasis pada ajaran yang terkandung

dalam Al-Qur'an dan hadits. Adapun yang kedua yaitu corak filosofis, ilmu pendidikan yang berbasis pada penalaran yang dilakukan oleh sarjana muslim. Ketiga, corak historis yakni ilmu pendidikan islam yang berpedoman pada sejarah. Terakhir yaitu corak aplikatif, ilmu pendidikan yang berdasar pada sistem dan cara penerapannya. Berdasarkan hal tersebut, maka dalam Al-Qur'an terdapat ayat mengenai pentingnya belajar dan pembelajaran:

اقْرَأْ بِاسْمِ رَبِّكَ الَّذِي خَلَقَ الْإِنْسَانَ مِنْ عَلَقٍ اقْرَأْ وَرَبُّكَ الْأَكْرَمُ الَّذِي عَلَّمَ بِالْقَلَمِ  
عَلَّمَ الْإِنْسَانَ مَا لَمْ يَعْلَمُ (العلق)

*“Bacalah dengan (menyebut) nama Tuhanmu yang menciptakan, Dia telah menciptakan manusia dari segumpal darah, Bacalah, dan Tuhanmulah Yang Maha Mulia, Yang mengajar (manusia) dengan pena, Dia mengajarkan manusia apa yang tidak diketahuinya.” (QS. Al-Alaq : 1-5)*

Ayat di atas menjelaskan tentang belajar dan pembelajaran, kata iqra' mempunyai berbagai makna seperti membaca, mengumpulkan, menganalisa, menyampaikan, mendalami, menelaah, meneliti, tetapi tidak menyebutkan suatu objek. Buku tafsir karangan Ahmad Izzan dan Saehudin menerangkan bahwa dalam wahyu pertama yang diturunkan kepada Nabi Muhammad SAW tersebut tidak dijelaskan sesuatu yang harus dibaca. Hal tersebut yang berarti Al-Qur'an menghendaki umatnya supaya membaca segala macam pengetahuan yang bermanfaat bagi manusia dalam menunjang kelangsungan hidup.

Nabi Muhammad SAW yang buta huruf aksara diperintahkan agar belajar membaca. Objek yang dibaca beragam, salah satunya yaitu ayat *al-qur'aniyah* (ayat-ayat tertulis) dan ayat *al-kauniyah* (ayat-ayat tidak tertulis). Belajar membaca ayat *qur'aniyah* dapat memperoleh ilmu agama seperti tauhid, fikih, akhlak. Sedangkan belajar membaca ayat *kauniyah* dapat memperoleh ilmu sains seperti biologi, fisika, kimia. Sehingga dapat disimpulkan bahwa ilmu yang

bersumber dari ayat *qur'aniyah* dan *kauniyah* harus diperoleh melalui proses belajar membaca.

Perintah membaca atau kata *iqra'* dalam QS. Al-Alaq terulang dua kali, yaitu terdapat pada ayat 1 dan 3. Menurut M. Quraish Shihab perintah membaca pertama sebagai perintah belajar mengenai sesuatu ataupun hal yang belum diketahui, sedangkan perintah membaca kedua sebagai perintah untuk mengajarkan ilmu kepada orang lain. Hal tersebut menunjukkan bahwa dalam proses belajar dan pembelajaran harus ada usaha yang maksimal kemudian harus mengajarkan ilmu tersebut kepada orang lain.

### **2.3 Kajian Topik dengan Teori Pendukung**

Pemerintah di Indonesia sangat menjunjung tinggi pendidikan, yaitu dengan memberlakukannya program wajib belajar selama 12 tahun. Untuk mensukseskan program tersebut maka diadakannya penerimaan peserta didik baru setiap tahunnya. Penerimaan peserta didik baru yaitu proses administrasi yang dilakukan setiap tahun untuk seleksi calon peserta didik agar dapat melanjutkan pendidikan ke jenjang yang lebih tinggi. Pada proses penerimaan peserta didik baru terdapat beberapa faktor yang dapat mempengaruhi peserta didik baru diterima atau tidak di sekolah tersebut.

Adapun metode yang dapat mengklasifikasi peserta didik baru dapat diterima atau tidak dapat menggunakan salah satu metode statistika yaitu metode *Classification and Regression Trees*. Metode CART (*Classification and Regression Trees*) dapat diaplikasikan pada berbagai macam penelitian. Pada penelitian ini, metode *Classification and Regression Trees* digunakan untuk mengklasifikasi penerimaan peserta didik baru. Sebelum mengklasifikasi

menggunakan metode *Classification and Regression Trees* perlu dilakukan analisis deskriptif terlebih dahulu. Selanjutnya memilih pemilah dengan menggunakan nilai *goodness of split* yang paling besar dan dilanjutkan dengan menentukan simpul terminal.

Setelah menentukan simpul terminal kemudian dilakukan penandaan label kelas yang dapat dilihat dari semua simpul terminal lalu melakukan pemangkasan pohon. Menghitung nilai akurasi dapat menggunakan beberapa cara salah satunya dengan menggunakan *confusion matrix* dan Kurva *Receiver Operating Characteristics* (ROC). ROC dapat dinyatakan sebagai klasifikasi prediksi mendekati atau menjadi negatif apabila semakin rendah titik ke kiri (0,0), sedangkan apabila semakin ke atas titik ke kanan (1,1) maka dinyatakan sebagai klasifikasi prediksi yang mendekati atau menjadi positif.

## **BAB III**

### **METODE PENELITIAN**

#### **3.1 Jenis Penelitian**

Pendekatan penelitian yang digunakan pada penelitian ini adalah pendekatan deskriptif kuantitatif. Pendekatan ini dimulai dari pengumpulan data, penafsiran terhadap data, dan penampilan hasil yang diperoleh. Deskriptif kuantitatif menggunakan teori-teori umum kemudian dilakukan penelitian sehingga menghasilkan sebuah kesimpulan.

#### **3.2 Data dan Sumber Data**

Data dalam penelitian ini menggunakan data sekunder penerimaan peserta didik baru di MAN 2 Kota Probolinggo. Pengambilan data dilakukan pada hari Senin tanggal 19 April 2021 di MAN 2 Kota Probolinggo. Data yang diperoleh dari hasil tes masuk jalur reguler tahun 2020 berjumlah 172 data dengan 79 peserta didik laki-laki dan 93 peserta didik perempuan. Data yang digunakan sebagai data *learning* sebanyak 90% dari data atau 155 data, sedangkan 10% atau 17 data lainnya digunakan sebagai data *testing*.

Variabel respon (Y) pada penelitian ini merupakan data diterima dan tidak diterimanya peserta didik di MAN 2 Kota Probolinggo yang berskala nominal. Variabel ini dibedakan menjadi 2 kategori yaitu:

Y=1, untuk peserta didik diterima

Y=2, untuk peserta didik tidak diterima

Variabel prediktor (X) yang digunakan pada penelitian ini yaitu variabel yang diduga berpengaruh terhadap penentuan hasil tes masuk di MAN 2 Kota

Probolinggo yang terbagi menjadi variabel kualitatif dan variabel kuantitatif. Variabel kualitatif terdiri dari dua variabel yaitu jenis kelamin dan asal sekolah, sedangkan variabel kuantitatif terdiri dari tiga variabel yaitu nilai pengetahuan bahasa, nilai pengetahuan logika dan sains, dan nilai pengetahuan keagamaan dan kewarganegaraan. Penelitian ini menggunakan variabel prediktor sebagai berikut:

1. Jenis Kelamin (X1)

Variabel ini dibedakan menjadi dua kategori yaitu:

X1=1, untuk peserta didik laki-laki

X1=2, untuk peserta didik perempuan

2. Asal Sekolah (X2)

Variabel asal sekolah dibagi menjadi dua kategori yaitu:

X2=1, untuk peserta didik yang berasal dari Sekolah Menengah Pertama (SMP) umum

X2=2, untuk peserta didik yang berasal dari Sekolah Menengah Pertama (SMP) berbasis islam

3. Nilai Pengetahuan Bahasa (X3)

Variabel nilai pengetahuan Bahasa ini diperoleh dari nilai rata-rata dua mata pelajaran yaitu Bahasa Indonesia dan Bahasa Inggris. Variabel ini menggunakan skala rasio 0-100.

4. Nilai Pengetahuan Logika dan Sains (X4)

Variabel nilai pengetahuan Logika dan Sains merupakan nilai yang diperoleh dari rata-rata nilai Matematika dan Ilmu Pengetahuan Alam (IPA) yang berskala rasio 0-100.

#### 5. Nilai Pengetahuan Keagamaan dan Kewarganegaraan (X5)

Variabel ini adalah nilai rata-rata dari mata pelajaran Pendidikan Agama Islam, Akidah Akhlak, dan Pendidikan Kewarganegaraan (PKN). Variabel X5 menggunakan skala rasio 0-100.

### 3.3 Lokasi Penelitian

Lokasi penelitian yaitu tempat dimana penelitian akan dilakukan. Penelitian ini dilakukan di MAN 2 Kota Probolinggo yang berlokasi di Jalan Soekarno-Hatta Kecamatan Kanigaran, Kota Probolinggo.

### 3.4 Teknik Pengumpulan Data

Pada penelitian ini penulis memperoleh data sebagai acuan dalam penelitian. Pengumpulan data menggunakan metode studi dokumen dimana data diperoleh dengan mengumpulkan berbagai data yang sudah ada.

### 3.5 Teknik Analisis Data

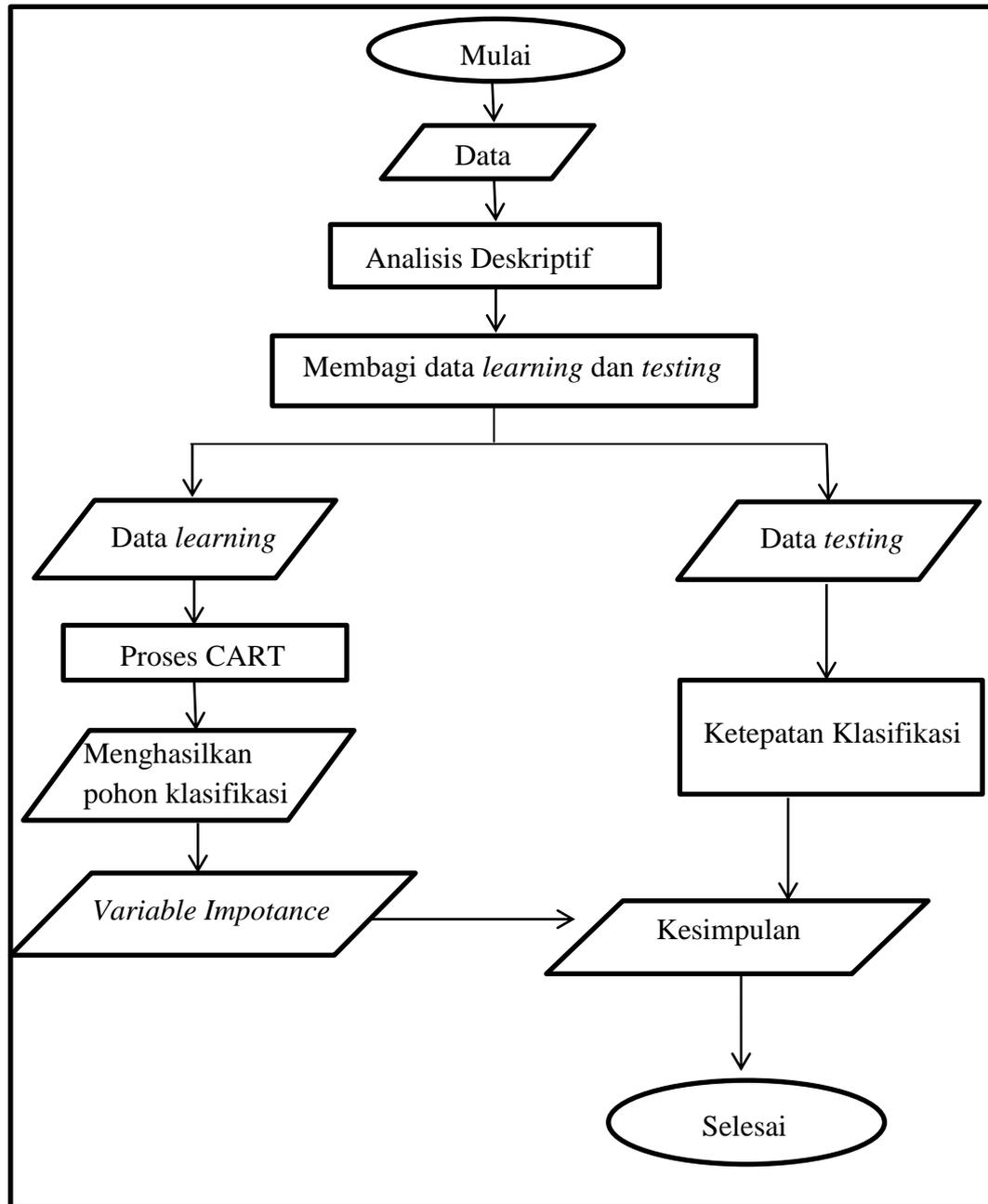
Penelitian ini menggunakan metode *Classification and Regression Trees* (CART) dengan bantuan *software* R tahap-tahap sebagai berikut:

1. Mengumpulkan data sekunder hasil seleksi tes masuk MAN 2 Kota Probolinggo.
2. Menentukan variabel respon dan variabel prediktor.
3. Melakukan statistik deskriptif pada masing-masing variabel yang digunakan.
4. Membagi data menjadi dua, yaitu data *learning* dan data *testing* dengan mencari kombinasi terbaik dari nilai ketepatan klasifikasi perbandingan 75%:25%, 80%:20%, 85%:15%, dan 90%:10%.

5. Membentuk pohon klasifikasi dengan menggunakan metode *Classification and Regression Trees* dengan tahap-tahap dibawah ini:
  - Memilih pemilah
  - Menentukan simpul terminal.
  - Melakukan penandaan label kelas yang dapat dilihat dari semua simpul terminal.
  - Melakukan pemangkasan pohon.
  - Melakukan pembentukan pohon optimal
6. Melakukan prediksi data menggunakan data *testing* dengan algoritma pohon CART
7. Menghitung nilai akurasi ketepatan pohon klasifikasi *Classification and Regression Trees* (CART)
8. Menentukan variabel yang mempengaruhi klasifikasi penerimaan peserta didik baru dari interpretasi hasil akhir *Classification and Regression Trees* (CART)

### 3.6 Flowchart

Tahapan pada penelitian ini dapat digambarkan dengan diagram alir sebagai berikut:



**Gambar 3.1** Flow Chart Classification and Regression Trees

## BAB IV

### HASIL DAN PEMBAHASAN

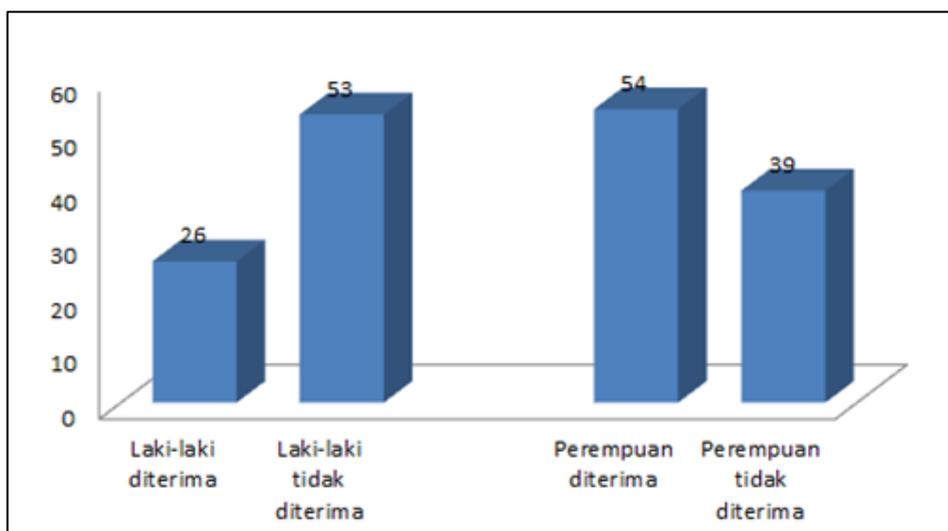
#### 4.1 Analisis Data

Jumlah banyaknya calon peserta didik baru di MAN 2 Kota Probolinggo tahun ajaran 2020-2021 sebanyak 172 peserta. Berdasarkan jumlah tersebut sebanyak 80 calon peserta didik baru dinyatakan telah diterima sebagai peserta didik baru di MAN 2 Kota Probolinggo dan sebanyak 92 dinyatakan tidak diterima. Diterima atau tidak diterimanya calon peserta didik baru menjadi variabel respon pada penelitian ini. Pada penelitian ini terdapat indikator-indikator yang diduga dapat mempengaruhi penerimaan peserta didik baru di MAN 2 Kota Probolinggo diantaranya yaitu jenis kelamin (laki-laki atau perempuan), asal sekolah (SMP umum dan SMP berbasis islam), nilai pengetahuan Bahasa, nilai pengetahuan Logika dan Sains, dan nilai pengetahuan Keagamaan dan Kewarganegaraan. Indikator-indikator tersebut dapat membantu untuk mengklasifikasi penerimaan peserta didik baru di MAN 2 Kota Probolinggo.



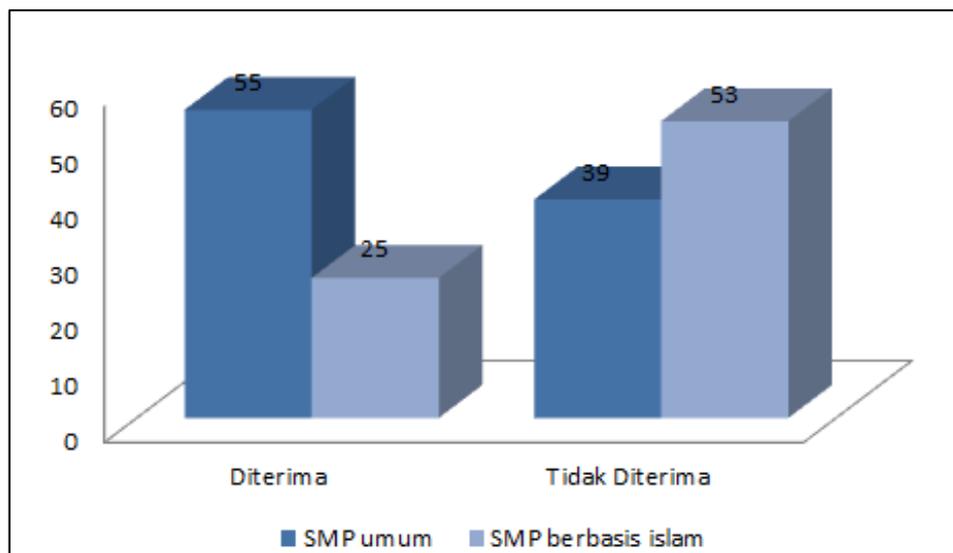
**Gambar 4.1** Diagram Lingkaran Jumlah Calon Peserta Didik Baru

Berdasarkan gambar 4.1 dapat diketahui bahwa dari 172 calon peserta didik baru yang mendaftar, sebesar 47% diterima atau sebanyak 80 peserta didik baru. Sedangkan peserta didik yang tidak diterima yaitu sebesar 53% atau sebanyak 92 peserta didik baru. Indikator pertama yang diduga dapat mempengaruhi penerimaan peserta didik baru di MAN 2 Kota Probolinggo yaitu jenis kelamin calon peserta didik baru.



**Gambar 4.2** Perbandingan Berdasarkan Jenis Kelamin

Pada gambar 4.2 menunjukkan perbandingan peserta didik baru diterima dan tidak diterima berdasarkan jenis kelamin. Didapatkan informasi bahwa di MAN 2 Kota Probolinggo lebih banyak perempuan yang diterima yaitu sebanyak 54 peserta didik daripada peserta didik laki-laki yaitu sebanyak 26. Peserta didik yang tidak diterima yang berjenis kelamin perempuan lebih sedikit yaitu sebanyak 39 peserta didik, sedangkan peserta didik laki-laki yang tidak diterima yaitu sebanyak 53. Indikator kedua yang diduga dapat mempengaruhi penerimaan peserta didik baru di MAN 2 Kota Probolinggo yaitu asal sekolah peserta didik.



**Gambar 4.3** Perbandingan Berdasarkan Asal Sekolah

Gambar 4.3 memberikan informasi terkait diterima dan tidak diterimanya peserta didik baru berdasarkan asal sekolah. Peserta didik baru yang berasal dari SMP umum yang diterima sebanyak 55 peserta didik dan yang tidak diterima sebanyak 39 peserta didik. Peserta didik baru diterima yang berasal dari SMP berbasis islam sebanyak 25 peserta didik sedangkan yang tidak diterima sebanyak 53 peserta didik. Indikator selanjutnya yaitu nilai pengetahuan Bahasa, nilai pengetahuan Logika dan Sains, serta nilai pengetahuan Keagamaan dan Kewarganegaraan yang berskala rasio. Berikut ini adalah deskriptif karakteristiknya:

**Tabel 4.1** Deskriptif Karakteristik

Variabel	Kategori	Rata-rata	Standar Deviasi	Maximum	Minimum
Pengetahuan Bahasa	Diterima	83,654	2,773	90,800	79,100
	Tidak Diterima	78,243	1,855	82,300	73,700
Pengetahuan Logika dan Sains	Diterima	81,931	3,751	92,000	76,300
	Tidak Diterima	75,060	8,096	82,800	5,810
Pengetahuan Keagamaan & Kewarganegaraan	Diterima	85,996	3,271	93,200	79,800
	Tidak Diterima	78,670	8,383	87,300	5,700

Pada tabel 4.1 diperoleh informasi mengenai indikator nilai pengetahuan Bahasa, nilai pengetahuan Logika dan Sains, serta nilai pengetahuan Keagamaan dan Kewarganegaraan terhadap klasifikasi penerimaan peserta didik baru yaitu diterima atau tidak diterima di MAN 2 Kota Probolinggo. Hasil rata-rata nilai pengetahuan Bahasa yang diterima yaitu sebesar 83,654 dengan standar deviasi sebesar 2,773, untuk rata-rata nilai pengetahuan Logika dan Sains yang diterima sebesar 81,931 yang memiliki standar deviasi sebesar 3,751, sedangkan rata-rata nilai pengetahuan Keagamaan dan Kewarganegaraan yang diterima sebesar 85,996 dan standar deviasinya sebesar 3,271. Dapat diketahui nilai rata-rata peserta didik dengan rentang antara 81,931 sampai dengan 85,996 mempunyai peluang untuk dapat diterima.

Hasil rata-rata nilai pengetahuan Bahasa, nilai pengetahuan Logika dan Sains, serta nilai pengetahuan Keagamaan dan Kewarganegaraan yang tidak diterima adalah sebesar 78,243, 75,060, dan 78,670. Nilai pengetahuan Bahasa dengan rentang antara 79,1 sampai dengan 90,8 mempunyai peluang untuk dapat diterima tetapi dengan mempunyai nilai minimum belum tentu dapat diterima, kecuali ditunjang dengan indikator lainnya. Nilai pengetahuan Logika dan Sains antara 76,3 sampai dengan 92 dinyatakan diterima. Nilai pengetahuan Keagamaan dan Kewarganegaraan dengan rentang antara 79,8 sampai dengan 93,2 dapat dinyatakan diterima. Jika ada peserta didik baru yang mempunyai nilai minimum, tidak menutup kemungkinan untuk tidak diterima karena nilai maksimum yaitu 87,3 bisa dikategorikan dalam peserta tidak diterima.

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

Analisis *Classification and Regression Trees* (CART) dilakukan untuk mengetahui klasifikasi penerimaan peserta didik baru di MAN 2 Kota Probolinggo yang dibagi menjadi dua kategori yaitu peserta didik diterima dan tidak diterima. Tahap pertama pada penelitian ini yaitu membagi data menjadi dua yakni data *learning* dan data *testing* dengan beberapa macam persentase terhadap 172 data penerimaan peserta didik baru di MAN 2 Kota Probolinggo tahun ajaran 2020/2021.

### 4.2.1 Penentuan Persentase Pembagian Data *Learning* dan Data *Testing*

Penelitian ini menggunakan pembagian persentase data *learning* dan data *testing* menjadi empat macam kombinasi yaitu (75%,25%), (80%, 20%), (85%,15%), dan (90%,10%). Selanjutnya menghitung ketepatan klasifikasi pada setiap persentase pembagian data. Setelah memperoleh hasil, melakukan perbandingan dari hasil ketepatan klasifikasi tersebut yang akan dipilih apabila nilai ketepatan klasifikasi yang mempunyai data *learning* dan data *testing* paling baik. Data *learning* digunakan untuk pembentukan pohon klasifikasi sedangkan data *testing* digunakan untuk pengukuran kebaikan (kelayakan) model.

**Tabel 4.2** Ketepatan Klasifikasi Pada Kombinasi Persentase Data

No.	Persentase Data (%)		Ketepatan Klasifikasi	
	<i>Learning</i>	<i>Testing</i>	<i>Learning</i>	<i>Testing</i>
1	75	25	0,9534884	0,8372093
2	80	20	0,9565217	0,8235294
3	85	15	0,9520548	0,8076923
4	90	10	<b>0,9483871</b>	<b>0,9411765</b>

Tabel 4.2 menunjukkan nilai ketepatan klasifikasi pada setiap kombinasi persentase data. Nilai ketepatan klasifikasi terbesar dari keempat kombinasi persentase data yaitu pada persentase 90% data *learning* dan 10% data *testing* dengan nilai 0,9483871 data *learning* dan 0,9411765 data *testing*. Persentase data yang akan digunakan pada analisis *Classification and Regression Trees* selanjutnya adalah 90% data *learning* dan 10% data *testing*.

#### 4.2.2 Pembentukan Pohon Klasifikasi

Terdapat beberapa tahap dalam pembentukan pohon klasifikasi CART. Tahap awal dalam pembentukan pohon klasifikasi yaitu menentukan banyak pemilah pada setiap variabel prediktor. Variabel prediktor pada penelitian ini menggunakan skala nominal dan rasio. Variabel yang berskala nominal yaitu variabel Jenis Kelamin (X1) dan variabel Asal Sekolah (X2), sedangkan Variabel Nilai Pengetahuan Bahasa (X3), Nilai Pengetahuan Logika dan Sains (X4), serta Nilai Pengetahuan Keagamaan dan Kewarganegaraan (X5) menggunakan skala rasio. Perhitungan untuk menentukan kemungkinan pemilah pada setiap variabel prediktor ditunjukkan pada tabel sebagai berikut:

**Tabel 4.3** Pemilihan Pemilah Variabel Prediktor

Variabel	Nama Variabel	Skala Data	Banyaknya Kategori	Kemungkinan Pemilah
X1	Jenis Kelamin	Nominal	2	$2^{2-1} - 1 = 1$
X2	Asal Sekolah	Nominal	2	$2^{2-1} - 1 = 1$
X3	Nilai Bahasa	Rasio	1-100	$100-1 = 99$
X4	Nilai Logika dan Sains	Rasio	1-100	$100-1 = 99$
X5	Nilai Keagamaan dan Kewarganegaraan	Rasio	1-100	$100-1 = 99$

Tabel 4.3 menunjukkan kemungkinan pemilah pada setiap variabel prediktor. Variabel Jenis Kelamin (X1) dan Asal Sekolah (X2) terdapat satu kemungkinan pemilah. Variabel Nilai Pengetahuan Bahasa (X3), Nilai Pengetahuan Logika dan Sains (X4) serta Nilai Pengetahuan Keagamaan dan Kewarganegaraan (X5) memiliki data yang berskala rasio 1 sampai dengan 100 sehingga terdapat 99 kemungkinan pemilah pada masing-masing variabel prediktor tersebut.

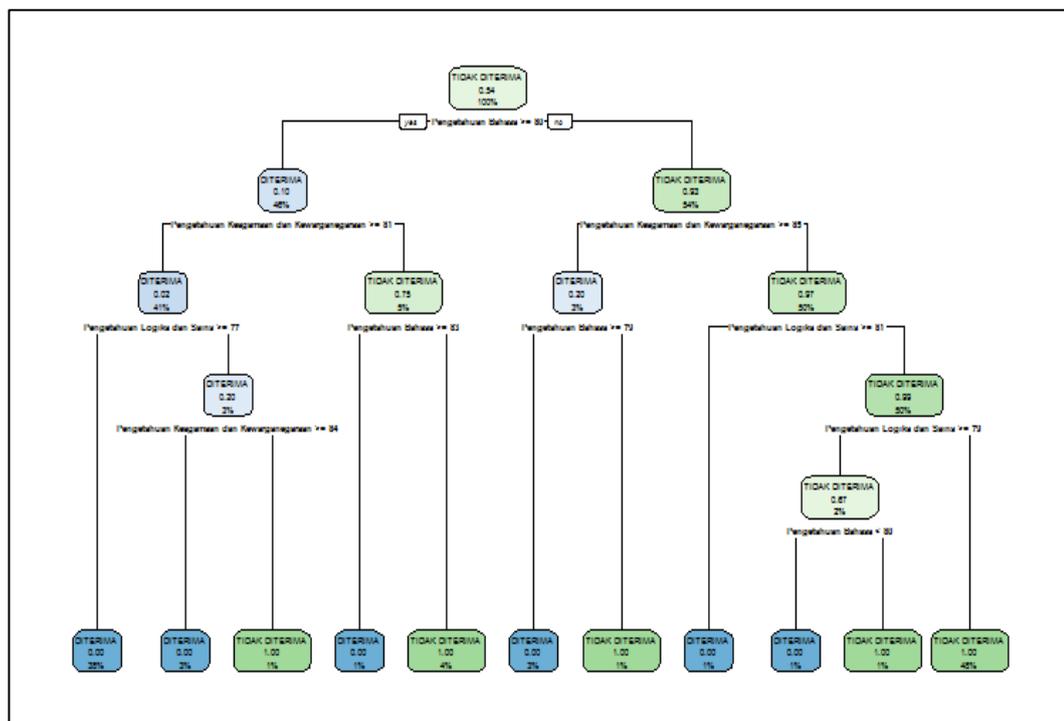
Langkah selanjutnya ialah menghitung nilai *indeks gini* untuk memperoleh nilai keheterogenan simpul. *Indeks gini* melakukan pemilahan simpul pada simpul kanan dan simpul kiri. Nilai *indeks gini* digunakan sebagai penentu dari *goodness of split*. Variabel pemilah yang memiliki nilai *goodness of split* paling tinggi yang akan menjadi pemilah terbaik. Berikut ini merupakan besarnya kontribusi variabel sebagai pemilah terbaik pada pohon klasifikasi yang ditunjukkan pada Tabel 4.4.

**Tabel 4.4** Urutan Variabel Terpenting

<b>Variabel</b>	<b>Nama Variabel</b>	<b>Nilai Variabel</b>
X3	Nilai Pengetahuan Bahasa	35
X5	Nilai Pengetahuan Keagamaan dan Kewarganegaraan	28
X4	Nilai Pengetahuan Logika dan Sains	23
X1	Jenis Kelamin	9
X2	Asal Sekolah	5

Tabel 4.4 menunjukkan bahwa variabel terpenting dan menjadi pemilah utama yaitu Variabel Nilai Pengetahuan Bahasa (X3) dengan nilai tertinggi yaitu

35. Variabel pemilah berikutnya yaitu Nilai Pengetahuan Keagamaan dan Kewarganegaraan (X5) dengan nilai 28. Setelah melakukan beberapa kali perulangan pada pembentukan pohon klasifikasi, *root node* yang dihasilkan yaitu Nilai Pengetahuan Bahasa. Setelah mendapatkan pemilah terbaik yakni variabel X3, maka pada simpul utama terdapat 155 peserta didik dari data *learning*. Selanjutnya akan dipilah untuk memecah simpul menjadi dua bagian yaitu simpul kiri ( $t_L$ ) dan simpul kanan ( $t_R$ ). Berikut ini adalah pohon klasifikasi maksimal pada klasifikasi penerimaan peserta didik baru:



**Gambar 4.4** Pohon Klasifikasi Maksimal

Gambar 4.4 menunjukkan pohon klasifikasi maksimal pada klasifikasi penerimaan peserta didik baru. Pohon klasifikasi maksimal merupakan pohon yang terbentuk dengan memiliki *terminal node* terbanyak dengan kedalaman yang tinggi. Pohon klasifikasi maksimal yang diperoleh memiliki tingkat kedalaman sebesar 6.

### 4.2.3 Pemangkasan Pohon Klasifikasi Maksimal

Setelah diperoleh pohon klasifikasi maksimal, selanjutnya yaitu akan dilakukan pemangkasan pohon supaya diperoleh pohon klasifikasi yang optimal. Pemangkasan pohon atau yang biasa disebut *pruning* dilakukan untuk mengurangi kompleksitas pohon supaya menjadi lebih sederhana dan terbentuk pohon klasifikasi optimal. Pemangkasan pohon juga dapat menghindari *overfitting* dan *underfitting*. *Overfitting* yaitu nilai akurasi yang diperoleh tidak sesuai atau bisa melebihi nilai sebenarnya, sedangkan *underfitting* adalah nilai akurasi yang diperoleh kurang dari nilai sebenarnya. Pohon yang baik dapat dihasilkan dengan melakukan pemangkasan dengan menggunakan *cost complexity minimum*. Berdasarkan nilai *complexity parameter* yang dapat meminimalkan nilai *cross validation error* (CV Error). Nilai *complexity parameter* dapat mengontrol ukuran pohon keputusan dan memilih ukuran pohon klasifikasi optimal. Berikut ini merupakan tabel nilai *complexity parameter* pada pohon klasifikasi maksimal:

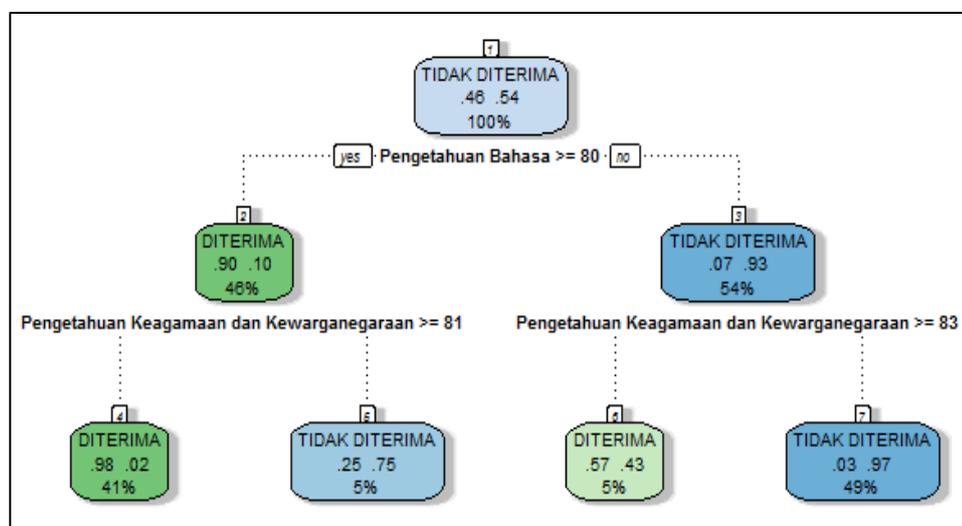
**Tabel 4.5** Nilai *Complexity Parameter*

No	CP	Nsplit	Rel error	Xerror	Xstd
1	0,817	0	1,000	1,000	0,087
2	0,056	1	0,183	0,183	0,049
3	0,042	2	0,127	0,183	0,049
4	0,028	3	0,085	0,155	0,045
5	0,141	4	0,056	0,113	0,039
6	0,007	6	0,028	0,099	0,036
7	0,000	10	0,000	0,085	0,034

Tabel 4.5 menunjukkan nilai CP (*Complexity Parameter*) dari pohon yang terkecil (tanpa pemecahan) sampai pohon yang terbesar (10 pemecahan). Setiap baris pada tabel tersebut mewakili kedalaman pohon yang berbeda. Semakin tinggi tingkatan pohon dapat memperoleh nilai kesalahan klasifikasi yang semakin rendah. Berdasarkan Tabel 4.5 dapat diketahui nilai *complexity parameter* untuk setiap split berturut-turut yaitu 0,817; 0,056; 0,042; 0,028; 0,141; 0,007; 0,000. Berdasarkan nilai *cross validation error* maka diperoleh nilai *complexity parameter* yang digunakan untuk pemangkasan pohon yaitu 0,817.

#### 4.2.4 Pohon Klasifikasi Optimal

Pohon klasifikasi optimal diperoleh setelah melakukan proses pemangkasan pohon klasifikasi maksimal, berikut ini adalah pohon klasifikasi optimal dari penerimaan peserta didik baru:



**Gambar 4.5** Pohon Klasifikasi Optimal Penerimaan Peserta Didik Baru

Gambar 4.4 menunjukkan pohon klasifikasi penerimaan peserta didik baru dengan *root node* ( $t_1$ ) yang hasilnya tidak diterima dengan indikator pengetahuan bahasa. Variabel penting yang berpengaruh pada proses

pengklasifikasian CART yaitu nilai pengetahuan bahasa dan nilai pengetahuan keagamaan dan kewarganegaraan. Pohon klasifikasi di atas menggambarkan terdapat dua *internal node* yaitu  $t_2$  dan  $t_3$  serta terdapat *terminal node* yaitu  $t_4, t_5, t_6$  dan  $t_7$ .

Berdasarkan pohon klasifikasi pada gambar 4.4 diketahui bahwa *root node* dengan kategori tidak diterima memiliki peluang 100%. Apabila nilai pengetahuan bahasa lebih dari sama dengan 80 maka peluang tidak diterima sebesar 0,46 sedangkan apabila nilai pengetahuan bahasa kurang dari sama dengan 80 maka peluang tidak diterima sebesar 0,54. Pada simpul dalam (*internal node*)  $t_2$  dengan kategori diterima memiliki peluang sebesar 46%. Apabila nilai pengetahuan keagamaan dan kewarganegaraan lebih dari sama dengan 81 maka peluang diterima yaitu sebesar 0,90, sedangkan apabila nilai pengetahuan keagamaan dan kewarganegaraan kurang dari sama dengan 81 maka peluang diterimanya yaitu 0,10. Simpul dalam (*internal node*)  $t_3$  yang berkategori tidak diterima memiliki peluang 54%. Apabila nilai pengetahuan keagamaan dan kewarganegaraan lebih dari sama dengan 83 maka peluang tidak diterima yaitu sebesar 0,07, sedangkan apabila nilai pengetahuan keagamaan dan kewarganegaraan kurang dari sama dengan 83 maka peluang tidak diterima sebesar 0,93.

Simpul terminal atau *terminal node* pada gambar 4.4 ditunjukkan dalam  $t_4, t_5, t_6$  dan  $t_7$ . Pada *terminal node*  $t_4$  dengan kategori diterima berpeluang sebesar 41%. Apabila nilai pengetahuan keagamaan dan kewarganegaraan lebih dari sama dengan 81 maka peluang diterima yaitu sebesar 0,98, sedangkan apabila nilai pengetahuan keagamaan dan kewarganegaraan kurang dari sama

dengan 81 maka peluang diterima sebesar 0,02. Simpul terminal (*terminal node*)  $t_5$  yang berkategori tidak diterima memiliki peluang sebesar 5%. Jika nilai pengetahuan keagamaan dan kewarganegaraan lebih dari sama dengan 81 maka peluang tidak diterima sebesar 0,25, sedangkan apabila nilai pengetahuan keagamaan dan kewarganegaraan kurang dari sama dengan 81 maka peluang tidak diterima sebesar 0,75.

Adapun *terminal node*  $t_6$  dengan kategori diterima memiliki peluang sebesar 5%. Jika nilai pengetahuan keagamaan dan kewarganegaraan lebih dari sama dengan 83 maka peluang diterimanya sebesar 0,57, sedangkan jika nilai pengetahuan keagamaan dan kewarganegaraan kurang dari sama dengan 83 maka peluang diterima yaitu sebesar 0,43. Simpul terminal (*terminal node*)  $t_7$  dengan kategori tidak diterima memiliki peluang sebesar 49%. Jika nilai pengetahuan keagamaan dan kewarganegaraan lebih dari sama dengan 83 maka peluang tidak diterima sebesar 0,03, sedangkan apabila nilai pengetahuan keagamaan dan kewarganegaraan kurang dari sama dengan 83 maka peluang tidak diterima sebesar 0,97.

#### **4.2.5 Ketepatan Klasifikasi *Classification and Regression Trees***

Ketepatan klasifikasi *Classification and Regression Trees* (CART) pada penerimaan peserta didik baru dalam menentukan diterima dan tidak diterimanya peserta didik baru dapat dilihat dari nilai *accuracy*, *sensitivity* dan *specificity*. *Sensitivity* dapat melihat tingkat akurasi suatu model dalam menentukan diterimanya peserta didik baru, sedangkan *specificity* dapat melihat tingkat akurasi suatu model dalam menentukan tidak diterimanya peserta didik baru.

Perhitungan tingkat akurasi ketepatan klasifikasi CART pada penerimaan peserta didik baru ialah sebagai berikut:

**Tabel 4.6** Ketepatan Klasifikasi CART Menggunakan Data *Learning*

<i>Predicted Class</i>	<i>Actual Class</i>		<b>Total</b>
	<b>Diterima</b>	<b>Tidak Diterima</b>	
<b>Diterima</b>	67	4	71
<b>Tidak Diterima</b>	4	80	84
<b>Total</b>	71	84	155

Tabel 4.6 menunjukkan bahwa ketepatan klasifikasi CART menggunakan data *learning* dengan jumlah total data sampel sebanyak 155 peserta didik. Terdapat data yang diprediksi diterima dan memang benar bahwa peserta didik tersebut diterima yaitu sebanyak 67 peserta didik, sedangkan terdapat 4 peserta didik yang diprediksi tidak diterima tetapi sebenarnya diterima. Peserta didik yang diprediksi tidak diterima dan sebenarnya memang tidak diterima sebanyak 80 peserta didik dan 4 peserta didik diprediksi diterima tetapi sebenarnya tidak diterima.

**Tabel 4.7** Ketepatan Klasifikasi CART

<i>Predicted Class</i>	<i>Actual Class</i>		<b>Total</b>
	<b>Diterima</b>	<b>Tidak Diterima</b>	
<b>Diterima</b>	8	0	8
<b>Tidak Diterima</b>	1	8	9
<b>Total</b>	9	8	17

Tabel 4.7 menunjukkan bahwa ketepatan klasifikasi CART dengan jumlah total data sampel sebanyak 17 peserta didik. Terdapat data yang diprediksi diterima yaitu sebanyak 8 peserta didik yang memang benar bahwa peserta didik tersebut diterima, sedangkan terdapat 1 peserta didik yang diprediksi tidak

diterima tetapi sebenarnya diterima. Peserta didik yang diprediksi tidak diterima sebanyak 8 peserta didik dan sebenarnya memang tidak diterima dan 0 peserta didik diprediksi diterima yang sebenarnya tidak diterima. Berikut ini perhitungan tingkat akurasi menggunakan *accuracy*, *sensitivity*, dan *specificity* pada pohon klasifikasi penerimaan peserta didik baru:

$$\begin{aligned}
 Accuracy &= \left( \frac{TP + TN}{TP + TN + FP + FN} \right) \times 100\% \\
 &= \left( \frac{8 + 8}{8 + 8 + 0 + 1} \right) \times 100\% \\
 &= \frac{16}{17} \times 100\% \\
 &= 0,9412 \times 100\% \\
 &= 94,12\%
 \end{aligned}$$

$$\begin{aligned}
 Sensitivity &= \frac{TP}{TP + FN} \times 100\% \\
 &= \frac{8}{8 + 1} \times 100\% \\
 &= \frac{8}{9} \times 100\% \\
 &= 0,8889 \times 100\% \\
 &= 88,89\%
 \end{aligned}$$

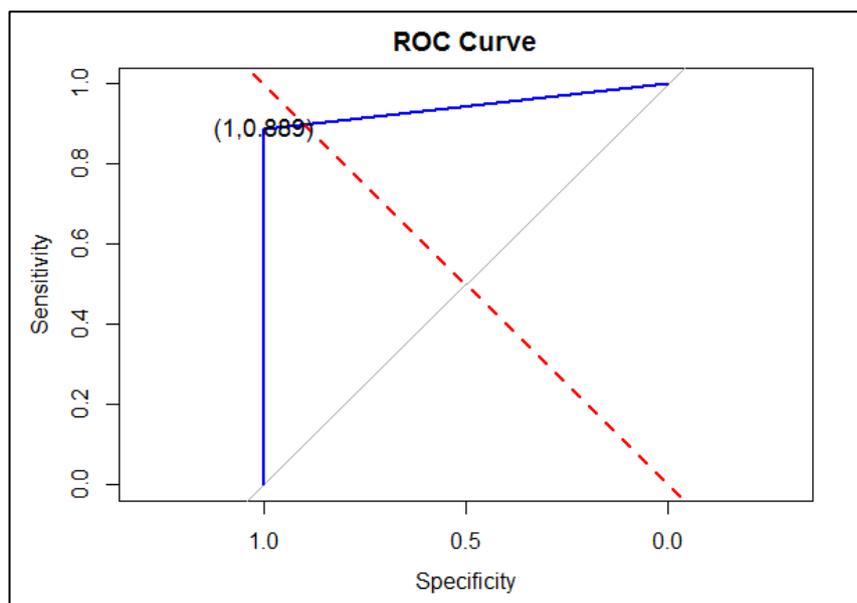
$$\begin{aligned}
 Specificity &= \frac{TN}{TN + FP} \times 100\% \\
 &= \frac{8}{8 + 0} \times 100\% \\
 &= \frac{8}{8} \times 100\% \\
 &= 1 \times 100\% \\
 &= 100\%
 \end{aligned}$$

**Tabel 4.7** Nilai Ketepatan Klasifikasi CART

<i>Accuracy</i>	<i>Sensitivity</i>	<i>Specificity</i>
94,12%	88,89%	100%

Berdasarkan perhitungan di atas diperoleh nilai akurasi sebesar 94,12% yang memiliki nilai *sensitivity* sebesar 88,89% dan nilai *specificity* sebesar 100%. Tingkat akurasi di atas menunjukkan bahwa *Classification and Regression Trees* (CART) sangat baik dalam mengklasifikasi penerimaan peserta didik baru di MAN 2 Kota Probolinggo.

Kurva ROC pada klasifikasi penerimaan peserta didik baru ditunjukkan pada Gambar 4.6:

**Gambar 4.6** Kurva ROC

Gambar 4.6 menunjukkan kurva ROC pada penerimaan peserta didik baru, garis berwarna biru pada kurva ROC tersebut menunjukkan titik dari nilai *sensitivity* dan *specificity*. Nilai dari *sensitivity* dan *specificity* yaitu sebesar 0,889 dan 1. Adapun perhitungan AUC (*Area Under Curve*) pada penerimaan peserta didik baru yaitu sebagai berikut:

$$AUC = \frac{1}{2} \left( \frac{TP}{TP + FN} + \frac{TN}{TN + FP} \right)$$

$$AUC = \frac{1}{2} \left( \frac{8}{8 + 1} + \frac{8}{8 + 0} \right)$$

$$AUC = \frac{1}{2} \left( \frac{8}{9} + \frac{8}{8} \right)$$

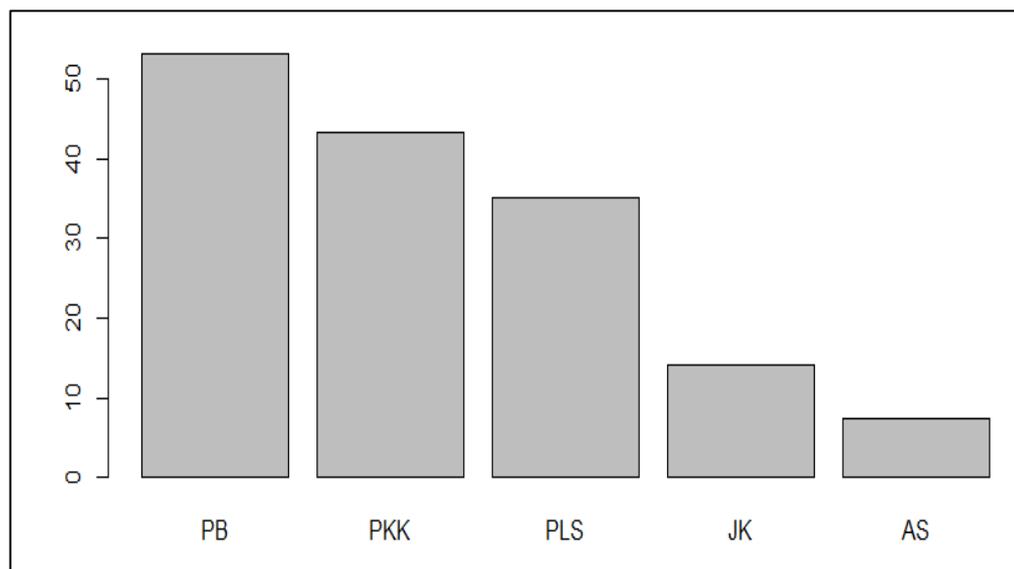
$$AUC = \frac{1}{2} \left( \frac{136}{72} \right)$$

$$AUC = 0,944$$

Diperoleh nilai AUC pada penerimaan peserta didik baru di MAN 2 Kota Probolinggo yakni sebesar 0,944 yang berarti bahwa model sangat baik dalam membedakan kelas diterima dan tidak diterima.

### 4.3 Variabel Penting Pada *Classification and Regression Trees*

Berdasarkan proses klasifikasi CART pada penerimaan peserta didik baru di MAN 2 Kota Probolinggo diketahui variabel prediktor sebagai *variable importance*. Tingkat pengaruh variabel pada pohon klasifikasi digambarkan pada bar plot berikut ini:



**Gambar 4.7** Bar Plot *Variable Importance*

Gambar di atas merupakan bar plot *variabel importance* yang mana variabel PB atau Pengetahuan Bahasa merupakan variabel yang memiliki pengaruh paling besar dalam pohon keputusan dengan nilai sebesar 53,18. Variabel PKK atau Pengetahuan Keagamaan dan Kewarganegaraan merupakan variabel penting kedua yang memiliki nilai sebesar 43,24, variabel PLS atau Pengetahuan Logika dan Sains memiliki nilai sebesar 35,16. Variabel yang memiliki pengaruh yang kecil dalam pohon keputusan yaitu variabel JK atau Jenis Kelamin dan AS atau Asal Sekolah dengan nilai 14,03 dan 7,39.

Perhitungan persentase tingkat kepentingan variabel prediktor pada klasifikasi penerimaan peserta didik baru ialah sebagai berikut:

$$VarImp = \left( \frac{\text{nilai overall tiap variabel}}{\text{total nilai overall}} \right) \times 100\%$$

- Nilai Pengetahuan Bahasa

$$VarImp = \left( \frac{53,18}{153} \right) \times 100\%$$

$$VarImp = 0,3476 \times 100\%$$

$$VarImp = 34,76\%$$

- Nilai Pengetahuan Keagamaan dan Kewarganegaraan

$$VarImp = \left( \frac{43,24}{153} \right) \times 100\%$$

$$VarImp = 0,2826 \times 100\%$$

$$VarImp = 28,26\%$$

- Nilai Pengetahuan Logika dan Sains

$$VarImp = \left( \frac{35,16}{153} \right) \times 100\%$$

$$VarImp = 0,2298 \times 100\%$$

$$VarImp = 22,98\%$$

- Jenis Kelamin

$$VarImp = \left(\frac{14,03}{153}\right) \times 100\%$$

$$VarImp = 0,0917 \times 100\%$$

$$VarImp = 9,17\%$$

- Asal Sekolah

$$VarImp = \left(\frac{7,39}{153}\right) \times 100\%$$

$$VarImp = 0,0483 \times 100\%$$

$$VarImp = 4,83\%$$

**Tabel 4.9** *Variabel Importance*

Variabel	Overall	Tingkat Kepentingan
Pengetahuan Bahasa	53,18	34,76%
Pengetahuan Keagamaan dan Kewarganegaraan	43,24	28,26%
Pegetahuan Logika dan Sains	35,16	22,98%
Jenis Kelamin	14,03	9,17%
Asal Sekolah	7,39	4,83%

Tabel 4.9 menunjukkan bahwa variabel Pengetahuan Bahasa merupakan variabel penting terbesar atau variabel yang sangat berpengaruh yang memiliki tingkat kepentingan sebesar 34,76% dari keseluruhan nilai pada data klasifikasi penerimaan peserta didik baru di MAN 2 Kota Probolinggo.

#### 4.4 Kajian Islam pada Hasil Penelitian

Pendidikan merupakan usaha terencana dalam proses pembelajaran supaya peserta didik dapat secara aktif mengembangkan potensi diri serta memiliki kepribadian yang mumpuni, kecerdasan, akhlak, pengetahuan umum yang diperlukan. Kualitas pendidikan dapat dilihat dari tiga aspek yaitu *input*, proses, dan *output*. *Input* peserta didik dilakukan untuk menyeleksi peserta didik dan mengetahui kebutuhan peserta didik. Kebutuhan peserta didik harus difasilitasi oleh lembaga sekolah supaya menghasilkan *output* peserta didik yang baik.

Berdasarkan penjelasan berikut, Allah SWT memerintahkan kepada umat Nabi Muhammad SAW untuk mencari ilmu sesuai dengan minat dan bakat masing-masing agar dapat berperan serta dalam mengisi pembangunan bangsa ini. Selain itu juga yang lebih penting adalah untuk mencari ilmu-ilmu Agama. Hal ini seiring dengan firman Allah SWT dalam surat An-Nahl ayat 125 yang berbunyi:

أَدْخِ إِلَى سَبِيلِ رَبِّكَ بِالْحِكْمَةِ وَالْمَوْعِظَةِ الْحَسَنَةِ وَجَادِلْهُمْ بِالَّتِي هِيَ أَحْسَنُ...

" Serulah (manusia) ke jalan Tuhanmu dengan hikmah dan pengajaran yang baik serta debatlah mereka dengan cara yang lebih baik." (QS. An-Nahl:125)

Surat An-Nahl ayat 125 menjelaskan bahwa kewajiban setelah menuntut ilmu itu untuk menyebarkannya atau berdakwah dengan cara yang bijaksana sesuai dengan kebutuhan masyarakat dan sesuai dengan keilmuan kita yang telah kita peroleh di bangku sekolah. Dalam hal ini, manfaat adanya metode pengklasifikasian yang telah dilakukan oleh sekolah MAN 2 Probolinggo ini agar memudahkan dalam proses belajar dan mengajar serta menghasilkan *output* yang maksimal yang mampu berkiprah serta berdaya saing tinggi di masyarakat.

## BAB V

### PENUTUP

#### 5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan klasifikasi penerimaan peserta didik baru di MAN 2 Kota Probolinggo diperoleh kesimpulan sebagai berikut:

1. Tingkat akurasi ketepatan klasifikasi pada pengklasifikasian penerimaan peserta didik baru di MAN 2 Kota Probolinggo dengan menggunakan *Classification and Regression Trees* (CART) sebesar 0,9412 atau 94,12% dengan nilai *sensitivity* sebesar 88,89% dan nilai *specitifty* sebesar 100%. Adapun nilai AUC pada penerimaan peserta didik baru di MAN 2 Kota Probolinggo yakni sebesar 0,944 yang berarti bahwa model sangat baik dalam membedakan kelas diterima dan tidak diterima.
2. Berdasarkan analisis *Classification and Regression Trees* (CART) terdapat *variable importance* yang mempengaruhi klasifikasi penerimaan peserta didik baru diantaranya yaitu Nilai Pengetahuan Bahasa yang merupakan variabel dengan pengaruh terbesar yang memiliki tingkat kepentingan sebesar 34,76%. Variabel yang berpengaruh kedua setelah Nilai Pengetahuan Bahasa yaitu Nilai Pengetahuan Keagamaan dan Kewarganegaraan dengan tingkat kepentingan sebesar 28,26%. Nilai Pengetahuan Logika dan Sains memiliki tingkat kepentingan sebesar 22,98%. Variabel Jenis Kelamin dan Asal Sekolah memiliki pengaruh yang cukup kecil terhadap penerimaan peserta didik baru dengan tingkat kepentingan sebesar 9,17% dan 4,83%.

## 5.2 Saran

Saran yang dapat diberikan berdasarkan hasil penelitian yang telah dilakukan yaitu untuk penelitian selanjutnya dapat menambahkan faktor-faktor lain dan menerapkan pembagian data *learning* dan data *testing* dengan kombinasi yang berbeda. Peneliti juga dapat membentuk pohon klasifikasi dari beberapa tahun lalu pada data penerimaan peserta didik baru agar diperoleh karakteristik klasifikasi penerimaan peserta didik baru yang lebih akurat, karena pada kasus penerimaan peserta didik baru selalu berubah setiap tahunnya dan nilai dari setiap calon peserta didik berbeda sehingga dapat mengakibatkan hasil dari pohon klasifikasi yang berbeda juga.

## DAFTAR PUSTAKA

- Andriyashin, A. (2005). Financial Applications of *Classification and Regression Trees*. CASE-Center of Applied Statistics and Economics, 7.
- Arrahimi, A. R. (2019). Teknik Bagging dan Boosting Pada Algoritma CART Untuk Klasifikasi Masa Studi Mahasiswa. *Jurnal Sains dan Informatika*, 5(1), 21-30.
- Bisri, A. (2015). Penerapan Adaboost untuk Penyelesaian Ketidaksamaan Kelas pada Penentuan Kelulusan Mahasiswa dengan Metode Decision Tree. *Journal of Intelligent Systems*, 1(1), 27-32.
- Bobbitt, Z. (2020, November 23). How to Perform Bagging in R (Step-by-Step). Retrieved Juni 11, 2022, from Statistics Simplified Statology. <https://www.statology.org/bagging-in-r/>
- Breiman, L., Friedman, J. H., Olshen, R. A., dan Stone, C.J. (1998). *Classification and Regression Trees*. New York: Chapman & Hall/CRC.
- Farlina, Y., & Hudin, M. J. (2017). Kajian Kepuasan Pengguna Informasi Penerimaan Peserta Didik Baru (PPDB) Online. *Journal on Computer and Information Technology*, 2(2), 48-54.
- Frastian, N. (2018). Implementasi Komparasi Algoritma Klasifikasi Menentukan Kelulusan Mata Kuliah Algoritma Universitas Budi Luhur. *Jurnal String*, 3(1), 1-8.
- Hangkawidjaja, Aan Darmawan, Yosafat A. P. dan Judea J. J. (2021). Perancangan dan Realisasi Detektor Retak Permukaan Dinding Bangunan. Karanganyar: Yayasan Lembaga Gumun Indonesia.
- Hartati, A., Zain, I., & S.U, B. S. (2012). Analisis CART (*Classification And Regression Trees*) pada Faktor-Faktor yang Mempengaruhi Kepala Rumah Tangga di Jawa Timur Melakukan Urbanisasi. *Jurnal Sains dan Seni ITS*, 1(1).
- Huda, M. (t.thn.). *Algoritma Data Mining: Analisis Data dengan Komputer*. Bisakimia.
- Ian H. Witten, E. (2011). *Data Mining Practical Machine Learning Tools and Techniques 3rd Edition*. USA: Elsevier.
- Kristiawan, M., Safitri, D., & Lestari, R. (2017). *Manajemen Pendidikan*. Sleman: Deepublish.
- Lewis, R. (2000). *An Introduction to Classification and Regression Tree (CART) Analysis*. Fransisco, California: UCLA Medical Center.
- Marbun, S. M. (2018). *Psikologi Pendidikan*. Ponorogo: Uwais Inspirasi Indonesia.
- Mulyasa, E. (2005). *Pendidikan Agama Islam Berbasis Kompetensi*. Bandung: PT Remaja Rosdakarya.

- Mulyasana, D., Sodikin, O., Kurniawan, A., & Haris, A. (2020). *Khazanah Pemikiran Pendidikan Islam dari Wacana Lokal hingga Tatanan Global*. Bandung: CV Cendekia Press.
- Neolaka, A., & Neolaka, G. A. (2017). *Landasan Pendidikan Dasar Pengenalan Diri Sendiri Menuju Perubahan Hidup*. Depok: Kencana.
- Prastiyo, F. (2019). *Peningkatan Hasil Belajar Peserta Didik dengan Model Kooperatif Jigsaw pada Materi Pecahan di Kelas V SDN Sepanjang 2*. Surakarta: CV Kekata Group.
- Pratiwi, F. E., & Zain, I. (2014). Klasifikasi Pengangguran Terbuka Menggunakan CART (*Classification and Regression Trees*) di Provinsi Sulawesi Utara. *Jurnal Sains dan Seni Pomits*, 3(1), 54-59.
- Prayitno. (2009). *Dasar Teori dan Praktis Pendidikan*. Semarang: Grasindo.
- Prianti, A. I., Santoso, R., & Hakim, A. R. (2020). Perbandingan Metode K-Nearest Neighbor dan Adaptive Boosting pada Kasus Klasifikasi Multi Kelas. *Jurnal Gaussian*, 9(3), 346-354.
- Rohman, Miftahur & Hairudin. (2018). Konsep Pendidikan Islam Perspektif Nilai-nilai Sosial Kultural. *Jurnal Pendidikan Islam*. 9(1). 21-35.
- Sari, Veronica Retno, Feranandah Firdausi, & Yufis Azhar. (2020). Perbandingan Prediksi Kualitas Kopi Arabika dengan Menggunakan Algoritma SGD, Random Forest dan Naïve Bayes. *Jurnal Pendidikan Informatika*, 4(2), 1-9.
- Suking, A. (2015, Agustus 14-16). Sistem Penerimaan Siswa Baru di Sekolah Efektif. *Prosiding TEMILNAS*, hal. 504-512.
- Susiana, L., Utami, I. T., & Junaidi. (2019). Penerapan Metode Boosting Pada CART untuk Mengklasifikasikan Korban Kecelakaan Lalu Lintas di Kota Palu. *Journal of Science and Technology*, 8(2), 106-109.
- Suwardika, G., S, K. P., & Nanik H, N. P. (2019). Ketidaktepatan Waktu Kelulusan Mahasiswa Universitas Terbuka dengan Metode Boosting CART. *Jurnal Varian*, 2(2), 37-46.
- Wezel, M., & Potharst, R. (2007). Improved Customer Choice Predictions using Ensemble Methods. *Journal Oper. Res*, 181(1), 436-452.
- Zainuddin, dkk. (2009). *Pendidikan Islam*. Malang: UIN Malang Press.

## LAMPIRAN

### Lampiran 1: Data Penerimaan Peserta Didik Baru

Keterangan:

JK : Jenis Kelamin

AS : Asal Sekolah

PB : Nilai Pengetahuan Bahasa

PLS : Nilai Pengetahuan Logika dan Sains

PKK : Nilai Pengetahuan Keagamaan dan Kewarganegaraan

<b>JK</b>	<b>AS</b>	<b>PB</b>	<b>PLS</b>	<b>PKK</b>	<b>K</b>
1	1	90,8	92	93,2	DITERIMA
2	1	88,8	88,2	92,8	DITERIMA
2	1	89,5	88,7	91,6	DITERIMA
2	1	90,7	87,3	91,3	DITERIMA
2	2	87,5	88,7	91,6	DITERIMA
1	1	87,9	88,4	91,5	DITERIMA
2	1	88,8	88,8	89,7	DITERIMA
2	1	87,6	89,7	88,4	DITERIMA
1	1	86,8	87,5	91,4	DITERIMA
1	1	86,8	86,1	92,3	DITERIMA
2	1	86,7	87,6	88,9	DITERIMA
2	1	87,1	84,8	89,6	DITERIMA
2	1	85,1	86,3	89,4	DITERIMA
2	1	88	84,1	88,6	DITERIMA
1	1	87,7	84,6	87,3	DITERIMA
2	2	84,1	90,6	84,1	DITERIMA
2	2	85,3	84,2	88,9	DITERIMA
2	1	85,9	83,4	88,7	DITERIMA
2	1	85	83	89,5	DITERIMA
2	1	84	81,6	91,5	DITERIMA
2	1	85,3	84,2	87	DITERIMA

<b>JK</b>	<b>AS</b>	<b>PB</b>	<b>PLS</b>	<b>PKK</b>	<b>K</b>
1	1	85,9	83,9	86,2	DITERIMA
1	1	83,9	86,5	85,5	DITERIMA
2	1	85,5	83,8	86,1	DITERIMA
2	2	84,9	82,8	87,7	DITERIMA
2	1	85,1	84,7	85,5	DITERIMA
2	1	85,3	83	86,1	DITERIMA
2	2	85,8	84,2	84,2	DITERIMA
1	1	85,8	82,3	85,6	DITERIMA
2	1	82,8	83,3	86,4	DITERIMA
2	1	85,5	81,3	85,1	DITERIMA
2	2	81,9	81,4	88,4	DITERIMA
1	1	83	82	86,1	DITERIMA
1	1	84,5	80,7	85,4	DITERIMA
2	2	82,5	81	86,8	DITERIMA
2	2	82,3	77,7	90,2	DITERIMA
2	1	80,4	82	87,8	DITERIMA
1	1	82,5	81,9	85,7	DITERIMA
2	1	84,7	77,9	87	DITERIMA
2	1	81,2	82,4	85,4	DITERIMA
2	1	83,3	80,4	85,2	DITERIMA
2	1	83,1	83,9	81,8	DITERIMA
2	2	79,2	80,1	89,1	DITERIMA
2	1	83,9	76,9	87,3	DITERIMA
2	2	84,8	80,5	82,6	DITERIMA
2	2	83,1	80,1	84,6	DITERIMA
2	1	81,4	81,4	84,9	DITERIMA
1	1	81,8	80,2	85,6	DITERIMA
2	1	82,7	82,8	81,9	DITERIMA
2	2	83,4	80,3	83,5	DITERIMA

<b>JK</b>	<b>AS</b>	<b>PB</b>	<b>PLS</b>	<b>PKK</b>	<b>K</b>
2	1	82,5	77,8	86,8	DITERIMA
2	1	81,6	79,1	85,9	DITERIMA
2	2	82,4	76,4	87,4	DITERIMA
1	1	82,8	82,6	80,7	DITERIMA
2	2	82,4	78,6	84,8	DITERIMA
2	1	83	78,2	84,5	DITERIMA
2	2	85,4	80,2	79,8	DITERIMA
2	1	82,3	76,5	86,4	DITERIMA
2	1	79,1	78,3	87,8	DITERIMA
1	1	81,2	79,8	84,2	DITERIMA
1	1	81,7	81,2	82,1	DITERIMA
2	1	82,4	78,6	83,9	DITERIMA
1	1	81,4	80,9	81,9	DITERIMA
1	1	79,9	81,9	82,3	DITERIMA
2	2	81	79,7	82,9	DITERIMA
2	2	81,6	79,3	82,7	DITERIMA
2	2	81,6	80,7	81	DITERIMA
1	1	79,8	77,8	85,5	DITERIMA
1	1	80,8	78,8	83	DITERIMA
2	2	80,4	76,3	86,1	DITERIMA
2	2	80,6	78,6	83,4	DITERIMA
1	1	82	77,8	82,3	DITERIMA
1	1	79,8	76,6	85,6	DITERIMA
1	2	81,1	79,6	81,1	DITERIMA
1	2	81,9	77,6	82,2	DITERIMA
1	2	81,4	79,2	80,9	DITERIMA
2	1	80,4	76,5	84,6	DITERIMA
1	2	80	79,5	81,9	DITERIMA
2	2	81,6	77,8	81,9	DITERIMA

<b>JK</b>	<b>AS</b>	<b>PB</b>	<b>PLS</b>	<b>PKK</b>	<b>K</b>
1	1	80,6	79,4	81,1	DITERIMA
2	1	82,1	76,5	82,5	TIDAK DITERIMA
2	1	79,2	78,9	82,9	TIDAK DITERIMA
1	2	80,1	78,8	81,7	TIDAK DITERIMA
1	2	79	79,3	82,3	TIDAK DITERIMA
2	1	80,1	77,8	82,6	TIDAK DITERIMA
2	2	79,4	78,3	82,6	TIDAK DITERIMA
1	1	80,3	79,3	80,3	TIDAK DITERIMA
2	2	80,2	79,9	79,8	TIDAK DITERIMA
2	2	78,1	78,5	83,2	TIDAK DITERIMA
2	2	81,9	78,5	79,4	TIDAK DITERIMA
1	1	77,7	78,9	82,7	TIDAK DITERIMA
2	2	79	72,9	87,3	TIDAK DITERIMA
2	1	80	76	83	TIDAK DITERIMA
1	2	80,9	78,8	79,3	TIDAK DITERIMA
2	2	78,1	77,7	83,2	TIDAK DITERIMA
1	1	78,8	79	81,2	TIDAK DITERIMA
2	2	78,1	78,6	82,2	TIDAK DITERIMA
2	1	78,3	75,4	85	TIDAK DITERIMA
1	1	79,6	77,4	81,6	TIDAK DITERIMA
2	2	79,4	77,7	81,4	TIDAK DITERIMA
2	2	82,3	75,4	80,6	TIDAK DITERIMA
1	2	80,1	78	80,1	TIDAK DITERIMA
1	2	80,2	79,5	78,5	TIDAK DITERIMA
2	1	77,6	78,5	82,1	TIDAK DITERIMA
1	1	81,2	75,9	80,9	TIDAK DITERIMA
2	1	79,7	75,3	82,9	TIDAK DITERIMA
2	1	79,2	78,6	80,1	TIDAK DITERIMA
1	1	80,7	76,7	80,3	TIDAK DITERIMA

<b>JK</b>	<b>AS</b>	<b>PB</b>	<b>PLS</b>	<b>PKK</b>	<b>K</b>
1	1	79,1	79	79,4	TIDAK DITERIMA
1	2	79	77,5	81	TIDAK DITERIMA
1	1	79,3	77,6	80,5	TIDAK DITERIMA
1	2	79,5	76,9	80,9	TIDAK DITERIMA
1	1	79,6	76,5	81,2	TIDAK DITERIMA
2	2	77,8	78,7	80,7	TIDAK DITERIMA
1	2	79,7	77	80,1	TIDAK DITERIMA
2	1	77	77,7	82,1	TIDAK DITERIMA
2	1	78,2	75,7	82,9	TIDAK DITERIMA
1	1	79	78	79,6	TIDAK DITERIMA
2	1	78,9	74	83,7	TIDAK DITERIMA
2	2	79	76,9	80,7	TIDAK DITERIMA
2	1	78,4	75,7	82,5	TIDAK DITERIMA
2	2	77,3	77,4	81,8	TIDAK DITERIMA
2	1	78,8	76,8	80,9	TIDAK DITERIMA
1	2	80	77,6	78,6	TIDAK DITERIMA
1	2	79,8	78,1	78,1	TIDAK DITERIMA
2	2	78,3	75,9	81,7	TIDAK DITERIMA
1	1	77,9	76,5	81,5	TIDAK DITERIMA
2	2	78,6	76,6	80,5	TIDAK DITERIMA
1	2	76,9	78,7	79,6	TIDAK DITERIMA
2	1	76,1	77,1	81,3	TIDAK DITERIMA
2	2	75,9	77	81,5	TIDAK DITERIMA
2	2	78,2	76,8	79,3	TIDAK DITERIMA
2	2	77,8	77,5	78,9	TIDAK DITERIMA
1	2	76,1	78,3	79,6	TIDAK DITERIMA
1	2	76,1	78,3	79,4	TIDAK DITERIMA
1	1	77,9	76,2	79,5	TIDAK DITERIMA
1	2	78,6	76,7	78,2	TIDAK DITERIMA

<b>JK</b>	<b>AS</b>	<b>PB</b>	<b>PLS</b>	<b>PKK</b>	<b>K</b>
2	2	79,6	72,7	80,7	TIDAK DITERIMA
1	1	75,8	75,6	81,3	TIDAK DITERIMA
1	2	77,1	77,6	77,8	TIDAK DITERIMA
1	1	76,2	76,3	79,7	TIDAK DITERIMA
1	1	76,3	77,1	78,4	TIDAK DITERIMA
1	1	79,1	72,8	79	TIDAK DITERIMA
1	1	78	73,6	79,1	TIDAK DITERIMA
1	2	76,4	75,9	77,4	TIDAK DITERIMA
2	2	76,5	75,5	77,2	TIDAK DITERIMA
1	1	75,3	76,8	77,1	TIDAK DITERIMA
1	1	75	74,6	79,4	TIDAK DITERIMA
1	2	76,6	74	78,3	TIDAK DITERIMA
1	2	76,7	75,4	76,7	TIDAK DITERIMA
1	2	75,4	75,1	78,3	TIDAK DITERIMA
2	2	73,7	76,6	78,5	TIDAK DITERIMA
1	2	76,7	75,3	76,7	TIDAK DITERIMA
1	2	76,7	72,2	79,5	TIDAK DITERIMA
2	2	75,1	75,5	77,1	TIDAK DITERIMA
1	2	74	73,8	78,7	TIDAK DITERIMA
1	1	76,3	72,4	77,4	TIDAK DITERIMA
2	2	80,7	82,8	62,5	TIDAK DITERIMA
2	2	76,3	72,2	76,4	TIDAK DITERIMA
1	1	75,8	72,6	75,9	TIDAK DITERIMA
2	2	76,8	72,2	74,5	TIDAK DITERIMA
1	2	75,3	69,8	77,3	TIDAK DITERIMA
1	1	71,7	74,1	76,3	TIDAK DITERIMA
1	2	72	74,4	75,1	TIDAK DITERIMA
1	1	75	72,7	73	TIDAK DITERIMA
1	2	75	65,7	79,3	TIDAK DITERIMA

<b>JK</b>	<b>AS</b>	<b>PB</b>	<b>PLS</b>	<b>PKK</b>	<b>K</b>
1	1	70,5	73,5	75,8	TIDAK DITERIMA
1	2	78,2	65,7	72,25	TIDAK DITERIMA
2	2	67,8	69,3	74,1	TIDAK DITERIMA
1	1	68,8	62,2	77,3	TIDAK DITERIMA
1	2	62,8	60,9	68,5	TIDAK DITERIMA
1	2	5,12	5,81	5,7	TIDAK DITERIMA

## Lampiran 2: Statistik Deskriptif Variabel Prediktor

### Statistik Deskriptif Nilai Pengetahuan Bahasa

**Descriptive Statistics: PB diterima; PB tidak diterima**

Variable	Mean	StDev	Minimum	Maximum
PB diterima	83,654	2,773	79,100	90,800
PB tidak diterima	78,243	1,855	73,700	82,300

### Statistik Deskriptif Nilai Pengetahuan Logika dan Sains

**Descriptive Statistics: PLS diterima; PSL tidak diterima**

Variable	Mean	StDev	Minimum	Maximum
PLS diterima	81,931	3,751	76,300	92,000
PSL tidak diterima	75,060	8,096	5,810	82,800

### Statistik Deskriptif Nilai Pengetahuan Keagamaan dan Kewarganegaraan

**Descriptive Statistics: PKK diterima; PKK tidak diterima**

Variable	Mean	StDev	Minimum	Maximum
PKK diterima	85,996	3,271	79,800	93,200
PKK tidak diterima	78,670	8,383	5,700	87,300

### Lampiran 3: Kombinasi Pembagian Data *Learning* dan Data *Testing*

#### Script dan *Output* Pembagian Data *Learning* dan Data *Testing*

##### 1. Persentase data 75% dan 25%

```
#Membuat data training dan testing
n <- round(nrow(a)*0.75);n
set.seed(99191)
samp=sample(1:nrow(a),n)
data.train = a[samp,]
dim(data.train)
data.test = a[-samp,]
dim(data.test)

#klasifikasi decision tree
fit.tree = rpart(Keterangan ~ ., data=data.train, method = "class", cp=0.008)
fit.tree

n= 129

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

1) root 129 57 TIDAK DITERIMA (0.44186047 0.55813953)
 2) Pengetahuan Bahasa>=80.35 60 7 DITERIMA (0.88333333 0.11666667)
   4) Pengetahuan Keagamaan dan Kewarganegaraan>=81.45 53 1 DITERIMA (0.98113208 0.01886792)
   *
     5) Pengetahuan keagamaan dan Kewarganegaraan< 81.45 7 1 TIDAK DITERIMA (0.14285714 0.85714
286) *
     3) Pengetahuan Bahasa< 80.35 69 4 TIDAK DITERIMA (0.05797101 0.94202899) *
```

##### 2. Persentase data 80% dan 20%

```
#Membuat data training dan testing
n <- round(nrow(a)*0.80);n
set.seed(99191)
samp=sample(1:nrow(a),n)
data.train = a[samp,]
dim(data.train)
data.test = a[-samp,]
dim(data.test)

#klasifikasi decision tree
fit.tree = rpart(Keterangan ~ ., data=data.train, method = "class", cp=0.008)
fit.tree

n= 138

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

1) root 138 61 TIDAK DITERIMA (0.44202899 0.55797101)
 2) Pengetahuan Bahasa>=80.35 63 7 DITERIMA (0.88888889 0.11111111)
   4) Pengetahuan Keagamaan dan Kewarganegaraan>=81.45 56 1 DITERIMA (0.98214286 0.01785714)
   *
     5) Pengetahuan Keagamaan dan Kewarganegaraan< 81.45 7 1 TIDAK DITERIMA (0.14285714 0.85714
286) *
     3) Pengetahuan Bahasa< 80.35 75 5 TIDAK DITERIMA (0.06666667 0.93333333)
     6) Pengetahuan Keagamaan dan Kewarganegaraan>=83.45 7 3 DITERIMA (0.57142857 0.42857143) *
     7) Pengetahuan Keagamaan dan Kewarganegaraan< 83.45 68 1 TIDAK DITERIMA (0.01470588 0.9852
9412) *
```

### 3. Persentase data 85% dan 15%

```
#Membuat data training dan testing
n <- round(nrow(a)*0.85);n
set.seed(99191)
samp=sample(1:nrow(a),n)
data.train = a[samp,]
dim(data.train)
data.test = a[-samp,]
dim(data.test)

#klasifikasi decision tree
fit.tree = rpart(Keterangan ~ ., data=data.train, method = "class", cp=0.008)
fit.tree

n= 146

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

1) root 146 66 TIDAK DITERIMA (0.45205479 0.54794521)
  2) Pengetahuan Bahasa>=80.35 67 7 DITERIMA (0.89552239 0.10447761)
    4) Pengetahuan Keagamaan dan Kewarganegaraan>=81.45 60 1 DITERIMA (0.98333333 0.01666667)
      *
    5) Pengetahuan Keagamaan dan Kewarganegaraan< 81.45 7 1 TIDAK DITERIMA (0.14285714 0.85714286) *
      3) Pengetahuan Bahasa< 80.35 79 6 TIDAK DITERIMA (0.07594937 0.92405063)
        6) Pengetahuan Keagamaan dan Kewarganegaraan>=83.45 7 3 DITERIMA (0.57142857 0.42857143) *
        7) Pengetahuan Keagamaan dan Kewarganegaraan< 83.45 72 2 TIDAK DITERIMA (0.02777778 0.9722222) *
```

### 4. Persentase data 90% dan 10%

```
#Membuat data training dan testing
n <- round(nrow(a)*0.90);n
set.seed(99191)
samp=sample(1:nrow(a),n)
data.train = a[samp,]
dim(data.train)
data.test = a[-samp,]
dim(data.test)

#klasifikasi decision tree
fit.tree = rpart(Keterangan ~ ., data=data.train, method = "class", cp=0.008)
fit.tree

n= 155

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

1) root 155 71 TIDAK DITERIMA (0.45806452 0.54193548)
  2) Pengetahuan Bahasa>=80.35 72 7 DITERIMA (0.90277778 0.09722222)
    4) Pengetahuan Keagamaan dan Kewarganegaraan>=80.95 64 1 DITERIMA (0.98437500 0.01562500)
      *
    5) Pengetahuan Keagamaan dan Kewarganegaraan< 80.95 8 2 TIDAK DITERIMA (0.25000000 0.75000000) *
      3) Pengetahuan Bahasa< 80.35 83 6 TIDAK DITERIMA (0.07228916 0.92771084)
        6) Pengetahuan Keagamaan dan Kewarganegaraan>=83.45 7 3 DITERIMA (0.57142857 0.42857143) *
        7) Pengetahuan Keagamaan dan Kewarganegaraan< 83.45 76 2 TIDAK DITERIMA (0.02631579 0.97368421) *
```

## Lampiran 4: Pembentukan Pohon Klasifikasi dan *Variable Importance*

### Script Pembentukan Pohon Klasifikasi Maksimal

```
fit=rpart(keterangan~.,method="class",data=data.train,control=rpart.control(minsplit=1,minbucket=1,cp=0,08))
plot(fit,uniform=TRUE,main="Maximal Classification Tree")
text(fit,use.n=TRUE,all=TRUE,cex=.8)
rpart.plot(fit)
summary(fit)
```

### Script Pembentukan Pohon Klasifikasi Optimal dan *Variable Importance*

```
#klasifikasi decision tree
fit.tree = rpart(keterangan ~ ., data=data.train, method = "class", cp=0.008)
fit.tree
summary(fit.tree)

#plot decision tree
library(rattle)
fancyRpartPlot(fit.tree)

#Confusion Matrix
prediksi<-predict(fit.tree, newdata=data.test,type="class")
prediksi
table(prediksi,data.test$keterangan)
akurasi=(8+8)/(17)
akurasi
sensitivity=8/9
sensitivity
specificity=8/8
specificity

#variabel penting
fit.tree$variable.importance
barplot(fit.tree$variable.importance)
```

## Lampiran 5: Menghitung AUC dan Kurva ROC

### Script Menghitung AUC dan Kurva ROC

```
# Konversi variabel target ke format biner
binary_truth <- ifelse(data.test$keterangan == "DITERIMA", 1, 0)
binary_prediction <- ifelse(class_predict == "DITERIMA", 1, 0)

# Hitung ROC
roc_curve <- roc(binary_truth, binary_prediction)
roc_curve

# Hitung AUC
auc_value <- auc(roc_curve)
auc_value

# Gambar kurva ROC
plot(roc_curve, main = "ROC Curve", col = "blue", lwd = 2)
abline(a = 0, b = 1, col = "red", lwd = 2, lty = 2) # Garis referensi diagonal
```

### Output ROC dan AUC

```
> roc_curve

call:
roc.default(response = binary_truth, predictor = binary_prediction)

Data: binary_prediction in 8 controls (binary_truth 0) < 9 cases (binary_truth 1).
Area under the curve: 0.9444
> # Hitung AUC
> auc_value <- auc(roc_curve)
> auc_value
Area under the curve: 0.9444
```

## RIWAYAT HIDUP



Luluk Chelawati Chodiroh yang biasa dipanggil Ella, lahir di Probolinggo pada tanggal 17 Juni 1998. Anak kedua dari tiga bersaudara dari pasangan Bapak Juliono A.F. Hasyim dan Ibu Titin Suhartiningsih. Pendidikan formal yang telah ditempuh dimulai dari taman kanak-kanak di TK Kusuma Bangsa Kota Probolinggo. Kemudian melanjutkan pendidikan dasar di SDN Triwung Kidul 2 Kota Probolinggo dan lulus pada tahun 2011. Setelah itu melanjutkan pendidikan menengah pertama di SMP Nurul Jadid Paiton, Kabupaten Probolinggo dan lulus pada tahun 2014. Selanjutnya menempuh pendidikan menengah atas di MAN 2 Kota Probolinggo dan lulus pada tahun 2017. Kemudian menempuh pendidikan tinggi di Universitas Islam Negeri Maulana Malik Ibrahim Malang dengan mengambil program studi Matematika. Selama kuliah, penulis aktif di berbagai organisasi dan kepanitiaan, salah satunya menjadi pengurus HMJ “Integral” Matematika. Penulis menerima saran dan masukan yang bersifat membangun demi manfaat tugas akhir ini. Pembaca dapat menghubungi penulis melalui email: [chelawatiii@gmail.com](mailto:chelawatiii@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 : Luluk Chelawati Chodiroh  
NIM : 17610088  
Fakultas/ Program Studi : Sains dan Teknologi/ Matematika  
Judul Skripsi : Penerapan Metode *Classification and Regression Trees* Terhadap Klasifikasi Penerimaan Peserta Didik Baru  
Pembimbing I : Dr. Fachrur Rozi, M.Si.  
Pembimbing II : Erna Herawati, M.Pd.

No	Tanggal	Hal	Tanda Tangan	
1	16 Januari 2023	Konsultasi Bab I, II, dan III	1.	
2	17 Januari 2023	Konsultasi Kajian Islam		2.
3	13 Februari 2023	Revisi Bab I dan II	3.	
4	14 Februari 2023	Revisi Kajian Islam		4.
5	17 Mei 2023	Revisi Bab III	5.	
6	28 Agustus 2023	ACC Seminar Proposal		6.
7	26 September 2023	Konsultasi Revisi Proposal	7.	
8	26 Oktober 2023	Konsultasi Kajian Islam Bab IV		8.
9	26 Oktober 2023	Konsultasi Bab IV	9.	
10	15 November 2023	Revisi Bab IV dan Konsultasi Bab V		10.
11	20 November 2023	Revisi Kajian Islam Bab IV	11.	

No	Tanggal	Hal	Tanda Tangan
12	29 November 2023	Revisi Bab V dan Format Penulisan	12. <i>Fp</i>
13	04 Desember 2023	Revisi Kajian Islam Bab IV	13. <i>Rea</i>
14	12 Desember 2023	ACC Seminar Hasil	14. <i>Fp</i>
15	12 Desember 2023	ACC Seminar Hasil Kajian Islam	15. <i>Fp</i>
16	22 Desember 2023	Revisi Seminar Hasil dan ACC Sidang Skripsi	16. <i>Fp</i>
17	30 Desember 2023	ACC Naskah Skripsi Keseluruhan	17. <i>Fp</i>


 30 Desember 2023  
 Ketua Program Studi Matematika  
*Susanti*  
 Susanti, M.Sc  
 NIP. 19741129 200012 2 005