

TESIS

**DATA MINING UNTUK PENENTUAN MODEL
TINGKAT KESUKSESAN KELULUSAN MURID SMA
PADA PERGURUAN TINGGI NEGERI
(STUDI KASUS IAIN BONE)**

*DETERMINATION OF HIGH SCHOOL STUDENT
GRADUATION MODEL AT IAIN BONE STATE UNIVERSITY*

IRMAWATI

2017.13.0024



PROGRAM PASCASARJANA

STMIK HANDAYANI

MAKASSAR

2019

HALAMAN PENGAJUAN

**DATA MINING UNTUK PENENTUAN MODEL
TINGKAT KESUKSESAN KELULUSAN MURID SMA
PADA PERGURUAN TINGGI NEGERI
(STUDI KASUS IAIN BONE)**

Tesis

Sebagai Salah Satu Syarat Untuk Meraih Gelar Magister Komputer (**M.Kom**)

Konsentrasi:

SISTEM INFORMASI

Disusun dan Diajukan Oleh :

IRMAWATI

2017.13.0024

Kepada

PROGRAM PASCASARJANA

STMIK HANDAYANI

MAKASSAR

2019

TESIS


DATA MINING UNTUK PENENTUAN MODEL
TINGKAT KESUKSESAN KELULUSAN MURID SMA PADA
PERGURUAN TINGGI NEGERI
(STUDI KASUS IAIN BONE)

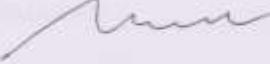
Disusun dan diajukan oleh :

IRMAWATI
NIM : 2017130024

Telah dipertahankan di depan Panitia Ujian Tesis
Pada Tanggal 5 Oktober 2019
Dan dinyatakan telah memenuhi syarat

Menyetujui,
Komisi Penasehat,



Dr. Ir. Zahir Zainuddin, M.Sc
Ketua


Dr. Eng. Yuyun, S.Kom., MT
Anggota

Ketua Program Studi
Sistem Komputer

Direktur Program Pascasarjana
STMIK Handayani


Dr. Ir. Zahir Zainuddin, M.Sc.


Dr. Eng Yuyun, S.Kom, MT.

LEMBAR PERNYATAAN KEASLIAN

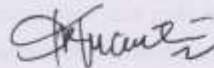
Yang bertanda-tangan dibawah ini :

Nama : IRMAWATI
Nomor Pokok Mahasiswa : 2017130024
Program Studi : SISTEM INFORMASI

Menyatakan dengan sebenarnya bahwa tesis yang saya tulis ini benar-benar merupakan hasil karya saya sendiri, bukan merupakan pengambilalihan tulisan atau pemikiran orang lain. Apabila dikemudian hari terbukti atau dapat dibuktikan bahwa sebagian atau keseluruhan tesis/disertasi ini hasil karya orang lain, saya bersedia menerima sanksi atas perbuatan tersebut.

Makassar, Oktober 2019

Yang menyatakan



IRMAWATI

PRAKATA

Puji syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa dengan selesainya tesis ini.

Gagasan yang melatarbelakangi permasalahan ini, timbul dari hasil pengamatan penulis terhadap penentuan kelulusan murid SMA pada Perguruan Tinggi Negeri IAIN Bone. Penulis bermaksud menyumbangkan beberapa konsep untuk menentukan model kelulusan dengan membuat pola atau aturan baru dari penentuan kelulusan pada IAIN Bone melalui penerimaan mahasiswa baru jalur mandiri.

Banyak kendala yang dihadapi penulis dalam rangka penyusunan tesis ini, dan hanya berkat bantuan berbagai pihak, maka tesis ini selesai pada waktunya. Dalam kesempatan ini, penulis dengan tulus menyampaikan terima kasih kepada Dr. Ir. Zahir Zainuddin, M.Sc sebagai Ketua Komisi Penasihat dan Dr. Eng. Yuyun, S.Kom., MT sebagai anggota komisi penasehat atas bantuan dan bimbingan yang telah diberikan terhadap penulisan tesis ini. Terima kasih juga penulis sampaikan kepada Kepala Unit Teknologi Informasi dan Pangkalan Data (UTIPD) IAIN Bone Dr. Zakaria, S.Pd., M.Si yang telah banyak membantu dalam rangka pengumpulan data dan informasi, serta kepada Nuranis, S.Kom yang telah banyak membantu dalam pengolahan data pada aplikasi dan yang terakhir saya ucapkan terima kasih kepada suami dan anakku Achmad Naufal Amzari yang rela meninggalkan sekolah demi kelancaran pembuatan tesis ini. Juga kepada mereka yang namanya tidak tercantum tetapi telah banyak membantu penulis dalam menyelesaikan tesis ini.

Makassar, Oktober 2019

IRMAWATI

ABSTRAK

IRMAWATI. Data Mining Untuk Penentuan Model Tingkat Kesuksesan Kelulusan Murid SMA pada Perguruan Tinggi Negeri IAIN Bone (dibimbing oleh Dr. Ir. Zahir Zainuddin, M.Sc dan Dr. Eng. Yuyun, S.Kom., MT).

Penelitian ini bertujuan (1) Memprediksi tingkat keberhasilan kelulusan murid pada Perguruan Tinggi Negeri dengan menggunakan algoritma Naive Bayes dan C.45. (2) Mengetahui kriteria yang menjadi penentu kesuksesan kelulusan murid pada IAIN Bone. (3) Menemukan pola atau aturan untuk penentuan kelulusan murid pada IAIN Bone. (4) Mengetahui tingkat akurasi algoritma Naive Bayes dan C.45 dalam penentuan kelulusan murid pada IAIN Bone.

Penelitian ini dilaksanakan di Institut Agama Islam Negeri (IAIN) Bone. Metode yang digunakan dalam penelitian adalah pengolahan data dalam bentuk data mining dengan menggunakan Algoritma Naive Bayes dan C.45. Pengambilan data dilakukan dengan menarik data dari Database Sistem Informasi Akademik (SISFO) IAIN Bone dari 3 (tiga) sekolah yaitu SMAN 4 Watampone, MAN 1 Bone, SMKN 1 Bone . Data diolah dengan menggunakan Microsoft Excel dan olah data pada aplikasi dengan menggunakan bahasa PHP dan database MySql yang dilanjutkan dengan melakukan pengujian nilai Precision, Recall dan Akurasi serta pengujian blackbox untuk aplikasi.

Hasil penelitian menunjukkan bahwa dari tujuh variabel yang ada yakni Tahun Lulus, Pendidikan Orang Tua, Penghasilan Orang Tua, Nilai Ujian Nasional, Nilai Tes, Nilai Wawancara dan Nilai Baca Tulis Huruf Qur'an (BTHQ) menunjukkan bahwa Nilai BTHQ menjadi penentu utama dalam kelulusan murid SMA pada IAIN Bone sesuai dengan hasil olahan data training sebanyak 170 dan olahan data testing sebanyak 10. Hasil perbandingan algoritma Naive Bayes dan C.45 menghasilkan Nilai Precision sebesar 100%, Recall 100%, Akurasi 100%. Selain itu, ditemukan pola sekuensial baru sebanyak 72 dengan melakukan pengujian ulang berdasarkan hasil urutan kejadian dari variabel Nilai Ujian Nasional, Nilai Tes, Nilai Wawancara dan Nilai BTHQ.

Kata Kunci : Data Mining, Naive Bayes, C.45, Presisi, Recall, Akurasi

ABSTRACT

IRMAWATI. Data Mining For Determination of the Model of Success Level of High School Graduation Students at IAIN Bone State University (supervised by Dr. Ir. Zahir Zainuddin, M.Sc and Dr. Eng. Yuyun, S.Kom., MT).

This study aims to determine (1) Predict the success rate of student graduation at State Universities using the Naive Bayes algorithm and C.45. (2) Knowing the criteria that determine the success of student graduation at IAIN Bone. (3) Finding patterns or rules for determining student graduation at IAIN Bone. (4) Knowing the accuracy of the Naive Bayes and C.45 algorithms in determining student graduation at IAIN Bone.

This research was conducted at the State Islamic Institute of Religion (IAIN) Bone. The method used in this research is data processing in the form of data mining using the Naive Bayes Algorithm and C.45. Data is collected by pulling data from the IAIN Bone Academic Information System Database (SISFO) from 3 (three) schools, namely SMAN 4 Watampone, MAN 1 Bone, SMK 1 Bone. Data is processed using Microsoft Excel and data processing in the application using the PHP language and MySql database, followed by testing the value of Precision, Recall and Accuracy and testing the blackbox for the application.

The results showed that of the eight variables, Graduation Years, Parental Education, Parents' Earnings, National Examination Values, Test Values, Interview Values and Literacy Letter Reading (BTHQ) scores indicate that BTHQ Values are determinants. The main result in high school students graduation at IAIN Bone is in accordance with the results of processed training data of 170 and processed data testing of 10. The results of the comparison of Naive Bayes and C.45 algorithms produce a Precision Value of 100%, Recall 100%, Accuracy of 100%. In addition, 72 new sequential patterns were found by re-testing based on the results of the sequence of events from the variable National Examination Score, Test Score, Interview Value and BTHQ Value.

Keywords : Data Mining, Naive Bayes, C.45, Precision, Recall, Accuracy

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PENGAJUAN	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN	iv
LEMBAR PERNYATAAN KEASLIAN	v
PRAKATA	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR	xiv
DAFTAR LAMPIRAN	xv
DAFTAR SINGKATAN	xvi
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah	1
B. Rumusan Masalah	4
C. Tujuan Penelitian	4
D. Manfaat Penelitian	5
E. Batasan Masalah	5
BAB II TINJAUAN PUSTAKA	6

A. Perguruan Tinggi	6
B. Data	7
C. Data Mining	8
D. Proses Data Mining	11
E. Tahap-Tahap Data Mining	11
F. Algoritma Klasifikasi	14
G. Penelitian Terkait	15
H. Kerangka Pikir	17
BAB III METODE PENELITIAN	18
A. Tahapan Penelitian	18
B. Waktu dan Lokasi Penelitian	28
C. Jenis Penelitian	29
D. Sumber Data	29
E. Instrumen Penelitian	29
BAB IV HASIL DAN PEMBAHASAN	31
A. Implementasi Sistem	31
B. Pengolahan Data	31
1. Algoritma Naïve Bayes	31
2. Algoritma C.45	33
C. Hasil Peengujian Precision, Recall dan Accuracy	49
D. Perbandingan Algoritma	52
E. Pola Sekuensial Yang Terbentuk	53

F. Implementasi Algoritma Dalam Aplikasi	59
G. Pengujian Blackbox	63
H. Listing Program Perhitungan Algoritma	64

BAB V KESIMPULAN DAN SARAN

A. Kesimpulan	71
B. Saran	72

DAFTAR PUSTAKA

DAFTAR TABEL

Nomor	Halaman
1. Kategori Tahun Lulus	19
2. Kategori Pendidikan Orang Tua	20
3. Kategori Penghasilan Orang Tua	20
4. Kategori Nilai Ujian Nasional	21
5. Kategori Nilai Tes	21
6. Kategori Nilai Wawancara	22
7. Kategori Nilai Baca Tulis Huruf Qur'an	22
8. Persamaan Precision, Recall dan Akurasi	27
9. Perhitungan Algoritma Naive Bayes	32
10. Penentuan Calon Cabang	34
11. Hasil Perhitungan Nilai Entrophy	35
12. Perhitungan Algoritma C.45	43
13. Hasil Perhitungan Entrophy Gain Tertinggi	45
14. Perhitungan Algoritma C.45 Gain Tertinggi	48
15. Data Testing.....	49
16. Tabel Prediksi	50
17. Tabel Perbandingan	50
18. Persamaan Precision, Recall dan Akurasi Data Testing	51
19. Pola Sekuensial Yang Terbentuk	53

20. Tabel Prediksi Pola Sekuensial	56
21. Test Case Pengujian Blackbox	63

DAFTAR GAMBAR

Nomor	Halaman
1. Contoh Atribut dan Object	7
2. Tahapan Proses Dalam Data Mining	11
3. Kerangka Pikir	17
4. Alur Metode Penelitian	18
5. Pohon Keputusan	44
6. Pohon Keputusan Terakhir	47
7. Form Login	59
8. Tampilan Halaman Depan	60
9. Menu Algoritma Naive Bayes	60
10. Tampilan Proses Algoritma Naive Bayes	61
11. Menu Algoritma C.45	61
12. Tampilan Proses Algoritma C.45	62
13. Tampilan Form Entry Data Uji	62
14. Tampilan Hasil Olah Data Uji	63

DAFTAR LAMPIRAN

Nomor	Halaman
1. Instrumen Penelitian	75
2. Data Training	76
3. Pengolahan Data Untuk Data Testing	91
4. Database Aplikasi	98

DAFTAR ARTI LAMBANG DAN SINGKATAN

Lambang/Singkatan	Arti dan Keterangan
R	Data yang belum diketahui kelasnya
S	Hipotesis pada data R yang merupakan class khusus Untuk Algoritma Naive Bayes
$P(R/S)$	Nilai probabilitas pada hipotesis R yang berdasarkan kondisi S
$P(R)$	Nilai probabilitas pada hipotesis R
$P(S/R)$	Nilai probabilitas S yang berdasarkan dengan kondisi hipotesis R
$P(S)$	Nilai probabilitas S
S	Himpunan kasus untuk Algoritma C.45
k	Jumlah partisi S
p_j	Jumlah kasus pada partisi ke-j
A	Atribut dari dataset
TP	True Positive, Nilai Sebenarnya yang bernilai sama dengan Nilai Prediksi
TN	True Negative, Nilai Sebenarnya yang bernilai benar, tetapi dari Nilai Prediksi bernilai salah
FP	False Positive, Nilai Sebenarnya yang bernilai salah, tetapi dari Nilai Prediksi bernilai benar
FN	False Negative, Nilai Sebenarnya dan Nilai Prediksi sama- sama menghasilkan nilai salah

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Pendidikan merupakan kebutuhan pokok bagi siapa saja dalam menghadapi era revolusi industri. Dimana setiap sumber daya manusia dituntut untuk lebih berkualitas dan mampu menghadapi perubahan-perubahan global. Tuntutan ini pun membuat murid SMA semakin peduli dengan kualitas pendidikan yang akan mereka tempuh selanjutnya setelah lulus.

Perguruan tinggi adalah tahap akhir opsional pada pendidikan formal. Biasanya dalam bentuk universitas, akademik, colleges, seminari, sekolah musik dan institus teknologi. Berdasarkan kepemilikannya, perguruan tinggi dibagi menjadi dua yaitu perguruan tinggi negeri dan perguruan tinggi swasta [1]. Perguruan Tinggi Negeri (PTN) merupakan tempat kuliah yang paling diminati oleh siapa saja yang akan lulus dari SMA atau sederajat. Namun, tidak semua yang memiliki minat tersebut, mampu lulus pada Perguruan Tinggi Negeri. Hal ini disebabkan banyak faktor seperti umur, nilai, waktu belajar, dan bakat/ minat.

Selain itu, pendidikan orang tua juga memiliki pengaruh terhadap pola perkembangan anak. Orang tua dengan tingkat pendidikan formal yang lebih tinggi mempunyai kemampuan lebih untuk membantu anak dalam belajar dibanding dengan orang tua dengan tingkat pendidikan yang lebih rendah [2]. Fenomena yang terjadi kebanyakan orang tua menginginkan anaknya menjadi orang yang sukses dalam

pendidikan maupun karirnya, sehingga di masa yang akan datang mereka dapat memperbaiki kualitas hidupnya menjadi lebih baik dari sebelumnya [2].

Undang-Undang Republik Indonesia Nomor 12 Tahun 2012 yang mengatur tentang Pendidikan Tinggi menjelaskan bahwa pendidikan tinggi sebagai bagian dari sistem pendidikan nasional memiliki peran strategis dalam mencerdaskan kehidupan bangsa dan memajukan ilmu pengetahuan dan teknologi dengan memperhatikan dan menerapkan nilai humaniora serta kebudayaan dan pemberdayaan bangsa Indonesia yang berkelanjutan. Selain itu, dalam undang-undang ini pun dijelaskan pada pasal 1 ayat (1) bahwa pendidikan adalah usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat, bangsa dan negara. Pada ayat (2) dan ayat (7) dijelaskan pula bahwa pendidikan tinggi adalah jenjang pendidikan setelah pendidikan menengah yang mencakup program diploma, program sarjana, program magister, program doktor, dan program profesi serta program spesialis, yang diselenggarakan oleh perguruan tinggi berdasarkan kebudayaan bangsa Indonesia dan perguruan tinggi negeri yang selanjutnya disingkat PTN adalah Perguruan Tinggi yang didirikan dan/atau diselenggarakan oleh Pemerintah. Pada Pasal 5 dalam undang-undang ini pun semakin mempertegas bahwa pendidikan tinggi bertujuan (a) berkembangnya potensi mahasiswa agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa dan berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, terampil, kompeten dan berbudaya untuk kepentingan bangsa [3].

Tingginya minat murid pada Perguruan Tinggi Negeri juga dapat dipertegas pada bunyi dari Pasal 31 ayat (1) Undang-Undang Dasar 1945, yaitu : “Setiap warga Negara berhak mendapatkan pendidikan”. Oleh karena itu, tingkat kelulusan murid perlu dilakukan penelitian untuk memudahkan mereka membenahi diri sebelum mengikuti seleksi ujian masuk Perguruan Tinggi Negeri. Tentunya dalam hal ini dapat diklasifikasikan faktor-faktor yang dapat menjuang kelulusan [4].

Institut Agama Islam Negeri (IAIN) Bone merupakan satu-satunya Perguruan Tinggi Keagamaan Islam Negeri yang berada di Kabupaten Bone dan berkedudukan di Ibu Kota Watampone. Perguruan Tinggi ini terdiri dari 4 (empat) fakultas yakni Fakultas Tarbiyah, Fakultas Syariah dan Hukum Islam, Fakultas Ushuluddin Dakwah, dan Fakultas Ekonomi dan Bisnis Islam dimana dari semua fakultas memiliki 12 (dua belas) Program Studi. Perguruan Tinggi ini menjadi salah satu yang paling diminati oleh Murid SMA yang ada di Kabupaten Bone ketika telah selesai menempuh pendidikan jalur formal. Tingginya minat murid ini, membuat penulis untuk membuat model kelulusan yang dikelola dalam bentuk data mining

Penentuan model kelulusan murid SMA pada Perguruan Tinggi Negeri, dapat dilakukan dengan cara menerapkan data mining yang memiliki beberapa metode salah satunya klasifikasi. Metode ini akan membangun suatu model yang bisa mengklasifikasikan suatu objek berdasarkan atribut-atributnya. Algoritma pengelolaan data yang digunakan dalam tesis ini adalah Naive Bayes dan C.45. Algoritma ini bisa memberikan teknik prediksi dan menemukan formula baru dalam penentuan model kelulusan. Pengelolaan data ini nantinya akan mengklasifikasi faktor-faktor apa saja

yang menjadi penentu kelulusan murid SMA pada Perguruan Tinggi Negeri. Kriteria yang digunakan dalam tesis ini berdasarkan Tahun Lulus, Pendidikan Orang Tua, Penghasilan Orang Tua, Nilai Ujian Nasional, Nilai Tes meliputi Pengetahuan Umum, Bahasa dan Agama serta Nilai Wawancara dan Nilai Baca Tulis Huruf Qur'an (BTHQ).

B. Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah :

1. Bagaimana penerapan metode klasifikasi dengan Naive bayes dan C.45 dalam memprediksi tingkat kesuksesan kelulusan murid pada IAIN Bone?
2. Faktor-faktor apa saja yang menjadi pendukung kesuksesan kelulusan murid pada IAIN Bone?
3. Bagaimana menemukan pola atau basis aturan untuk menentukan tingkat kesuksesan kelulusan murid pada IAIN Bone?
4. Bagaimana membandingkan tingkat akurasi algoritma Naive Bayes dan C.45 dalam penentuan kelulusan murid pada IAIN Bone dengan parameter precision, recall dan akurasi?

C. Tujuan Penelitian

Tujuan dari penelitian ini adalah :

1. Memprediksi tingkat keberhasilan kelulusan murid pada Perguruan Tinggi Negeri dengan menggunakan algoritma Naive Bayes dan C.45.
2. Mengetahui kriteria yang menjadi penentu kesuksesan kelulusan murid pada IAIN Bone.

3. Menemukan pola atau basis aturan untuk penentuan kelulusan murid pada IAIN Bone.
4. Mengetahui tingkat akurasi algoritma Naive Bayes dan C.45 dalam penentuan kelulusan murid pada IAIN Bone.

D. Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Memberikan informasi kepada IAIN Bone mengenai klasifikasi tingkat kelulusan murid SMA yang diterima.
2. Memberikan informasi kepada murid SMA yang akan lulus dari sekolah untuk mengetahui faktor-faktor apa saja yang menjadi penentu tingkat kesuksesan kelulusan pada IAIN Bone.

E. Batasan Masalah

Perguruan Tinggi Negeri yang akan menjadi objek penelitian adalah terfokus pada Penerimaan Mahasiswa Baru Jalur Mandiri Tahun 2018 di IAIN Bone dengan mengambil data dari 3 (tiga) sekolah yakni SMAN 4 Watampone, MAN 1 Bone, dan SMKN 1 Watampone. Sejumlah 170 (seratus tujuh puluh) data untuk data training dan 10 (sepuluh) data untuk data testing/uji.

BAB II

TINJAUAN PUSTAKA

A. Perguruan Tinggi

Perguruan tinggi adalah satuan pendidikan penyelenggara pendidikan tinggi. Peserta didik perguruan tinggi disebut mahasiswa, sedangkan tenaga pendidik perguruan tinggi disebut dosen. Menurut jenisnya, perguruan tinggi dibagi menjadi dua [1] :

1. Perguruan tinggi negeri adalah perguruan tinggi yang diselenggarakan oleh pemerintah.
2. Perguruan tinggi swasta adalah perguruan tinggi yang diselenggarakan oleh pihak swasta.

Di Indonesia, Perguruan Tinggi Negeri dikelola oleh Kementerian Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia. Rektor perguruan tinggi negeri merupakan pejabat setingkat eselon 2 di bawah Menteri Riset, Teknologi, dan Pendidikan Tinggi Republik Indonesia ataupun kementerian lainnya [1].

Perguruan Tinggi Islam Negeri (PTAN) di Indonesia berada di bawah tanggung jawab Kementerian Agama. Ada tiga jenis perguruan tinggi yang termasuk ke dalam kategori ini, yaitu Universitas Islam Negeri (UIN), Institut Agama Islam Negeri (IAIN), dan Sekolah Tinggi Agama Islam Negeri (STAIN). Di setiap provinsi di Indonesia umumnya terdapat satu UIN, IAIN, atau STAIN.

Sesuai dengan Undang-Undang No. 12 Tahun 2012, amanat tentang Perguruan Tinggi diatur pada pasal 1 ayat (6) bahwa perguruan tinggi adalah satuan pendidikan yang menyelenggarakan pendidikan tinggi [3].

B. Data

Himpunan data (data-set) merupakan kumpulan dari objek dan atributnya. Atribut merupakan sifat atau karakteristik dari suatu objek. Atribut juga dikenal sebagai variabel, field, karakteristik atau fitur. Kumpulan dari atribut merupakan sebuah objek yang kemudian disebut dengan record, titik, kasus, sample, entitas atau instance (Fajar Astuti Herawati. 2013).

	Tid	Refund	Marital Status	Taxable Income	Cheat
1		Yes	Single	125K	No
2		No	Married	100K	No
3		No	Single	70K	No
4		Yes	Married	120K	No
5		No	Divorced	95K	Yes
6		No	Married	60K	No
7		Yes	Divorced	220K	No
8		No	Single	65K	Yes
9		No	Married	75K	No
10		No	Single	90K	Yes

Gambar 1. Contoh Atribut dan Object

Nilai-nilai atribut adalah angka atau simbol yang memberi nilai pada atribut tersebut. Perbedaan antara atribut dan nilai-nilai atribut, yaitu :

1. Atribut yang sama dapat dipetakan ke nilai atribut yang berbeda. Contoh : ketinggian dapat diukur dengan satuan kaki atau meter

2. Atribut yang berbeda dapat dipetakan ke himpunan nilai yang sama. Contoh : nilai atribut untuk ID dan umur adalah bilangan bulat, tetapi sifat dari nilai atribut dapat berbeda. ID tidak terbatas tetapi umur mempunyai nilai minimal dan maksimal.

C. Data Mining

Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar (Turban dkk. 2005). Terdapat beberapa istilah lain yang memiliki makna samadengan data mining, yaitu *Knowledge discovery in databases* (KDD), ekstraksi pengetahuan (*knowledge extraction*), Analisa data/pola (*data/pattern analysis*), kecerdasan bisnis (*business intelligence*) dan *data archaeology* dan *data dredging* (Larose, 2005)

Kemampuan Data mining untuk mencari informasi bisnis yang berharga dari basis data yang sangat besar, dapat dianalogikan dengan penambangan logam mulia dari lahan sumbernya, teknologi ini dipakai untuk :

1. Prediksi trend dan sifat-sifat bisnis, dimana data mining mengotomatisasi proses pencarian informasi memprediksi di dalam basis data yang besar.
2. Penemuan pola-pola yang tidak diketahui sebelumnya, dimana data mining menyapu basis data, kemudian mengidentifikasi pola-pola yang sebelumnya tersembunyi dalam satu sapuan.
3. Data mining berguna untuk membuat keputusan yang kritis, terutama dalam strategi.

Proses yang umumnya dilakukan oleh data mining antara lain: deskripsi, prediksi, estimasi, klasifikasi, clustering dan asosiasi. Secara rinci proses data mining dijelaskan sebagai berikut (Larose, 2005):

1. Deskripsi

Deskripsi bertujuan untuk mengidentifikasi pola yang muncul secara berulang pada suatu data dan mengubah pola tersebut menjadi aturan dan kriteria yang dapat mudah dimengerti oleh para ahli pada domain aplikasinya. Aturan yang dihasilkan harus mudah dimengerti agar dapat dengan efektif meningkatkan tingkat pengetahuan (knowledge) pada sistem. Tugas deskriptif merupakan tugas data mining yang sering dibutuhkan pada teknik postprocessing untuk melakukan validasi dan menjelaskan hasil dari proses data mining. Postprocessing merupakan proses yang digunakan untuk memastikan hanya hasil yang valid dan berguna yang dapat digunakan oleh pihak yang berkepentingan.

2. Prediksi

Prediksi memiliki kemiripan dengan klasifikasi, akan tetapi data diklasifikasikan berdasarkan perilaku atau nilai yang diperkirakan pada masa yang akan datang. Contoh dari tugas prediksi misalnya untuk memprediksikan adanya pengurangan jumlah pelanggan dalam waktu dekat dan prediksi harga saham dalam tiga bulan yang akan datang.

3. Estimasi

Estimasi hampir sama dengan prediksi, kecuali variabel target estimasi lebih ke arah numerik dari pada ke arah kategori. Model dibangun menggunakan record lengkap

yang menyediakan nilai dari variabel target sebagai nilai prediksi. Selanjutnya, pada peninjauan berikutnya estimasi nilai dari variabel target dibuat berdasarkan nilai variabel prediksi. Sebagai contoh, akan dilakukan estimasi tekanan darah sistolik pada pasien rumah sakit berdasarkan umur pasien, jenis kelamin, berat badan, dan level sodium darah. Hubungan antara tekanan darah sistolik dan nilai variabel prediksi dalam proses pembelajaran akan menghasilkan model estimasi.

4. Klasifikasi

Klasifikasi merupakan proses menemukan sebuah model atau fungsi yang mendeskripsikan dan membedakan data ke dalam kelas-kelas. Klasifikasi melibatkan proses pemeriksaan karakteristik dari objek dan memasukkan objek ke dalam salah satu kelas yang sudah didefinisikan sebelumnya.

5. Clustering

Clustering merupakan pengelompokan data tanpa berdasarkan kelas data tertentu ke dalam kelas objek yang sama. Sebuah kluster adalah kumpulan record yang memiliki kemiripan suatu dengan yang lainnya dan memiliki ketidakmiripan dengan record dalam kluster lain. Tujuannya adalah untuk menghasilkan pengelompokan objek yang mirip satu sama lain dalam kelompok-kelompok. Semakin besar kemiripan objek dalam suatu cluster dan semakin besar perbedaan tiap cluster maka kualitas analisis cluster semakin baik.

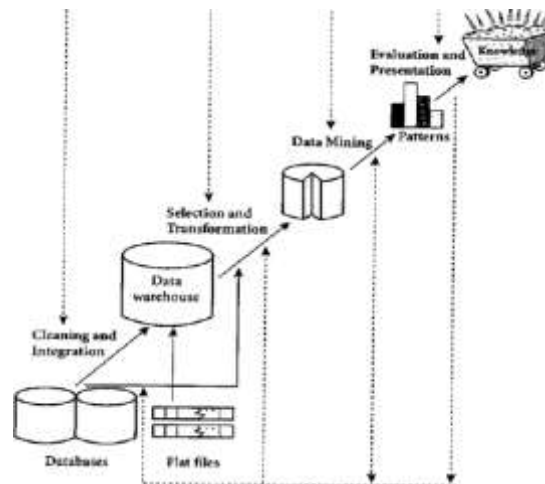
6. Asosiasi

Tugas asosiasi dalam data mining adalah menemukan atribut yang muncul dalam suatu waktu. Dalam dunia bisnis lebih umum disebut analisis keranjang belanja

(market basket analysis). Tugas asosiasi berusaha untuk mengungkap aturan untuk mengukur hubungan antara dua atau lebih atribut.

D. Proses Data Mining

Dalam penerapan data mining, diperlukan pemahaman terhadap data, proses diperolehnya data, alasan menerapkan data mining dan target yang ingin dicapai. Sehingga secara garis besar sudah ada hipotesa mengenai aksi- aksi yang dapat diterapkan dari hasil penerapan data mining. Pemahaman- pemahaman tersebut akan sangat membantu dalam mendesain proses data mining dan juga pemilihan teknik data mining yang akan diterapkan. Pada umumnya proses data mining berjalan interaktif karena tidak jarang hasil data mining pada awalnya tidak sesuai dengan harapan analisnya sehingga perlu dilakukan desain ulang prosesnya.



Gambar 2. Tahapan Proses Dalam Data Mining [5]

E. Tahap-Tahap Data Mining

Sebagai suatu rangkaian proses, data mining dapat dibagi menjadi beberapa tahap yang diilustrasikan di Gambar 2. Tahap-tahap tersebut bersifat interaktif di mana

pemakai terlibat langsung atau dengan perantara knowledge base.

1. Pembersihan data (untuk membuang data yang tidak konsisten dan noise)

Pada umumnya data yang diperoleh, baik dari database suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak relevan dengan hipotesa data mining yang kita miliki. Pembersihan data yang tidak relevan akan mempengaruhi performansi dari sistem data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

2. Integrasi data (penggabungan data dari beberapa sumber)

Integrasi data dilakukan pada atribut-atribut yang mengidentifikasi entitas-entitas yang unik. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Dalam integrasi data ini juga perlu dilakukan transformasi dan pembersihan data karena seringkali data dari dua database berbeda tidak sama cara penulisannya atau bahkan data yang ada di satu database ternyata tidak ada di database lainnya. Hasil integrasi data sering diwujudkan dalam sebuah data warehouse

3. Transformasi data (data diubah menjadi bentuk yang sesuai untuk di-mining)

Beberapa teknik data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Karenanya data berupa angka numerik yang berlanjut perlu dibagi-bagi menjadi beberapa interval. Proses ini sering disebut binning. Transformasi dan pemilihan data ini juga menentukan kualitas dari hasil data mining nantinya karena

ada beberapa karakteristik dari teknik-teknik data mining tertentu yang tergantung pada tahapan ini.

4. Aplikasi teknik data mining

Aplikasi teknik data mining sendiri hanya merupakan salah satu bagian dari proses data mining. Beberapa teknik data mining sudah umum dipakai. Ada kalanya teknik-teknik data mining umum yang tersedia di pasar tidak mencukupi untuk melaksanakan data mining di bidang tertentu atau untuk data tertentu.

5. Evaluasi pola yang ditemukan (untuk menemukan yang menarik/bernilai)

Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti : menjadikannya umpan balik untuk memperbaiki proses data mining, mencoba teknik data mining lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

6. Presentasi pola yang ditemukan untuk menghasilkan aksi

Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisa yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami data mining. Karenanya presentasi hasil data mining dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses data mining.

F. Algoritma Klasifikasi

Algoritma klasifikasi yang cenderung digunakan dalam pengembangan data mining, yaitu:

1. C.45, metode ini menjadi pilihan pertama yang sering digunakan dalam pengembangan data mining karena kecepatan dalam pengklasifikasian pohonkeputusan disamping dapat mengkonstruksi pengklasifikasian dengan aturan-aturan yang lain.
2. k -NN, beberapa hal yang menjadi perhatian dalam penggunaan algoritma ini adalah penggunaan pilihan k , jika k sangat kecil maka akan mengakibatkan noise. Sebaliknya jika terlalu besar dapat menyebabkan N dengan banyak kelas yang harus diklasifikasikan. Tetapi kesederhanaan metode menjadi nilai lebih sehingga menjadi pilihan banyak developer data mining selain itu, algoritma ini mudah untuk dipahami dan diimplementasikan dalam tekniknya. Penelitian ini menyatakan bahwa banyak peneliti berpendapat bahwa algoritma ini lebih baik dari SVM berdasarkan skema pengklasifikasiannya
3. Naive Bayes, penelitian tersebut menyimpulkan bahwa metode algoritma Naive Bayes memiliki keunggulan untuk pengembangan data mining, yaitu kemudahan konstruksinya dan tidak membutuhkan parameter skema pengulangan yang kompleks sehingga mudah dalam membaca data dalam jumlah yang besar. Hal ini terjadi karena desain rancangan penuntun klasifikasi terhadap data. Selain itu, metode ini dinyatakan sebagai algoritma yang mempunyai sifat *simplicity*, *elegance* dan *robustness*.

4. CART, penerapan metode algoritmaini banyak digunakan dalam berbagai bidang yang membutuhkan pengolahan data yangkomprehensif. Hanya saja mekanismenya terdiri dari beberapa tahap yang bertingkatmeliputi *automatic class balancing*, *automatic missing*, *value handling* dan *allows forcost-sensitive learning*, *dynamic feature construction* dan *probability tree estimation* sehingga tingkat kompleksitas menjadi pertimbangan para peneliti pemula. Hasil akhirnya adalah gambaran atribut berdasarkan prioritas kebutuhan proses.

G. Penelitian Terkait

Penelitian yang terkait dengan judul adalah sebagai berikut :

Judul Jurnal	Penulis Jurnal	Metode dan Hasil Penelitian	Keterkaitan
Klasifikasi Profil Siswa SMA/SMK yang Masuk PTN (Perguruan Tinggi Negeri) dengan k-Nearest Neighbor	YuandriTri saputra, Indriyani, Shellafuri MardikaBiru, Muhammad Ervan	Dari beberapa percobaan yang telah dilakukan diperoleh kesimpulan sebagai berikut: 1. Metode k-Nearest Neighbor sebagai teknik dalam data mining dapat digunakan untuk <i>classifier</i> pada data SNMPTN. 2. <i>Classifier</i> terbaik dihasilkan dari percobaan 4 dengan <i>dataset</i> 2440 <i>record</i> yang proporsi <i>record</i> pada setiap kelas target seimbang yang merupakan hasil <i>Oversampling</i> replikasi dengan metode uji <i>10-fold cross validation</i> . 3. Data yang seimbang atau <i>balance</i> lebih baik untuk klasifikasi daripada data yang <i>imbalance</i> . 4. Akurasi yang diperoleh dari <i>classifier</i> terbaik adalah 83.3607 %. 5. <i>Classifier</i> terbaik yang dihasilkan dapat digunakan untuk memprediksi kampus pilihan yang menerima siswa	Meneliti tentang PTN dengan obyek adalah murid/siswa SMA tapi dengan algoritma yang berbeda

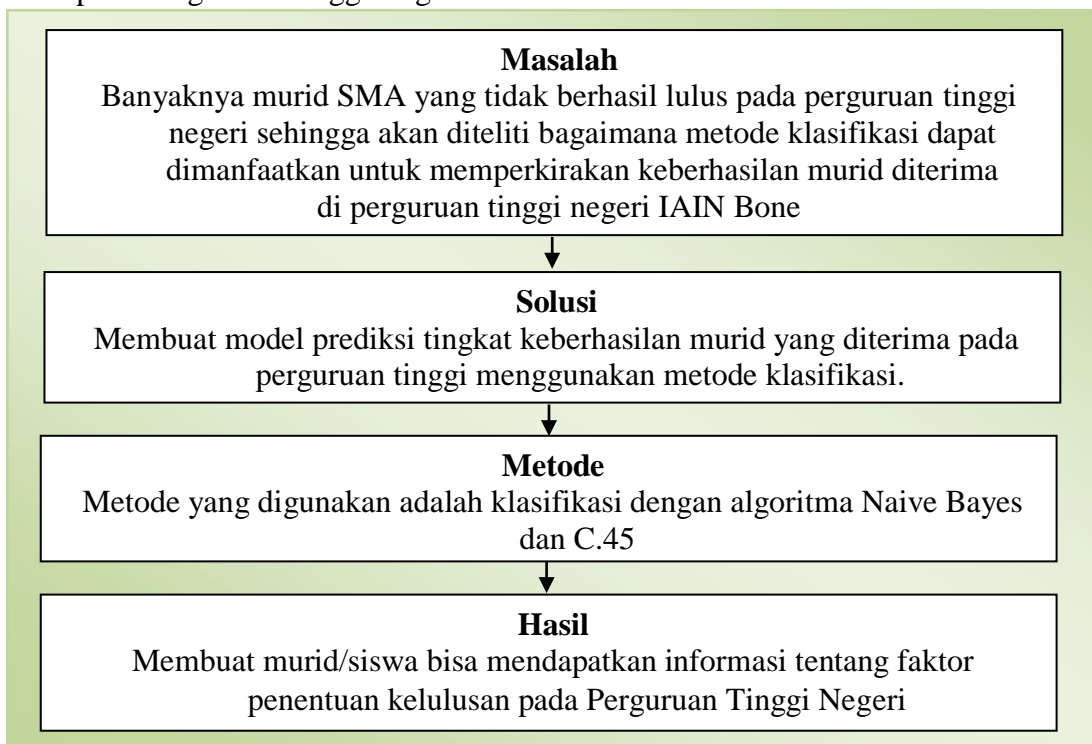
		dengan model terbaik yang diperoleh.	
Sistem Simulasi Prediksi Profil Kelulusan Mahasiswa Dengan Decison Tree	Bekti Amalia Arifiyani1), Ridha Sefina Samosir2)	Dari seluruh tahapan penelitian yang telah dilakukan serta hasil yang diperoleh maka peneliti menarik beberapa simpulan sebagai berikut:(1) Prediksi profil kelulusan mahasiswa dapat dilakukan berdasarkan data saat mahasiswa mengikuti sistem penerimaan mahasiswa baru dan data waktu kelulusan mahasiswa tersebut; (2) Metode pengembangan sistem yang digunakan yaitu prototype sesuai dengan kebutuhan pengembangan system karena ruang lingkup system tidak besar; dan (3) Hasil prediksi dapat digunakan oleh program studi untuk melakukan tindakan antisipasi melalui kegiatan pembimbingan akademik, pembimbingan tugas akhir dan kegiatan yang lainnya.	Sama-sama meneliti tentang kelulusan
Metode <i>naive bayes</i> Untuk Penentuan Penerima Beasiswa Bidikmisi Universitas mulawarman	Diasrina Dahri1), Fahrul Agus2), Dyna Marisa Khairina3)	Kurang lebih 900 pendaftar beasiswa bidikmisi Universitas Mulawatman setiap tahun, menyebabkan proses seleksinya berjalan lamban dan berpotensi tidak konsisten. Ketidakkonsistenan pada sistem penentuan penerima menyebabkan tujuan penyelenggaraan beasiswa menjadi kabur, tidak transparan dan tidak tepat sasaran. Penelitian ini bertujuan untuk membantu bagian proses seleksi dengan membuat aplikasi perangkat lunak sistem pendukung keputusan untuk penentuan penerima beasiswa bidikmisi Universita Mulawarman. Penentuan penerima beasiswa menggunakan beberapa kriteria antara lain: pekerjaan orang tua, penghasilan orang tua, jumlah tanggungan, daya listrik (<i>watt</i>), dan nilai ujian nasional. Kelayakan calon penerima beasiswa	Sama-sama meneliti tentang data mining menggunakan algoritma <i>naive bayes</i>

		bidikmisi ditentukan dengan menerapkan metode <i>Naive Bayes</i> . Metode ini dipilih karena mampu mempelajari data kasus sebelumnya yang digunakan sebagai data uji. Penelitian ini telah menghasilkan sebuah aplikasi sistem pendukung keputusan dengan tingkat akurasi sebesar 85.56%.	
--	--	---	--

Perbedaan antara penelitian sebelumnya dengan penelitian yang akan dikerjakan adalah dari segi algoritma yang digunakan. Kalau dari penelitian yang ada menggunakan algoritma k-neighborhood dan decision tree, untuk penelitian yang akan dilakukan menggunakan algoritma Naive Bayes dan C.45.

H. Kerangka Pikir

Berikut adalah kerangka pikir penelitian mengenai model tingkat kelulusan SMA pada Perguruan Tinggi Negeri :

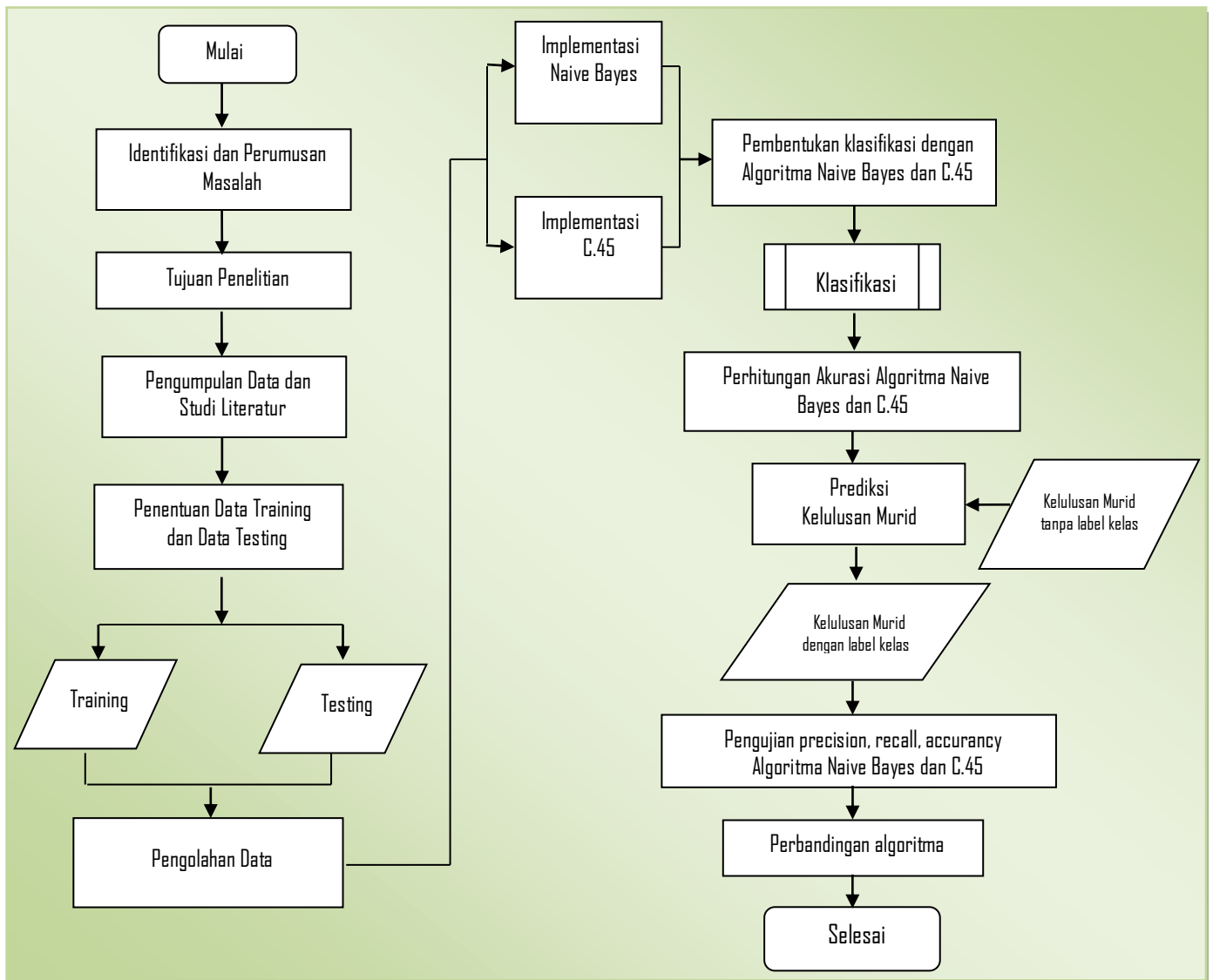


Gambar 3. Kerangka Pikir

BAB III METODE PENELITIAN

A. Tahapan Penelitian

Penelitian ini akan melalui beberapa tahap, dimana setiap tahap berdekatan dan saling mempengaruhi satu sama lain. Adapun tahapan penelitian tersebut dapat dilihat pada alur metode penelitian di bawah ini :



Gambar 4. Alur Metode Penelitian

Tahapan penelitian yang dilakukan pada gambar 4 dapat dijelaskan sebagai berikut :

1. Studi Literatur atau Kepustakaan

Pada tahap ini, dilakukan pencarian studi literatur tentang materi penelitian mengenai data mining, klasifikasi, algoritma Naive Bayes dan C.45 serta cara-cara pengujian menggunakan precision, recall dan accuracy. Pencarian ini didasarkan pada teori yang berkaitan dengan penelitian beserta teori yang berkembang saat ini.

2. Pengelompokan Kategori Variabel Penelitian

Dari data yang diperoleh, maka dilakukan proses pembersihan data yakni menyeleksi data yang akan digunakan dengan menghapus data yang tidak penting yang sudah diambil pada database sisfo akademik untuk penerimaan mahasiswa baru. Setelah itu dilakukan integrasi data dengan membuat pengelompokan kategori dari setiap data yang ada untuk lebih mempermudah dalam proses pengklasifikasian. Adapun kategori yang terbentuk, yaitu

a. Variabel Tahun Lulus

Variabel tahun lulus merupakan tahun lulus dari calon mahasiswa baru yang mendaftar pada IAIN Bone. Pengelompokan kategori variabel tahun lulus pada pendaftar Tahun 2018, yaitu :

Tabel 1. Kategori Tahun Lulus

No	Tahun Lulus	Kategori
1	2014	1
2	2015	2
3	2016	3
4	2017	4
5	2018	5

b. Variabel Pendidikan Orang Tua

Variabel pendidikan orang tua merupakan tingkat pendidikan orang tua dari semua calon mahasiswa baru yang mendaftar. Untuk variabel pendidikan orang tua pengelompokan kategori dilakukan dengan melihat aturan yang ada pada Undang-Undang No. 20 Tahun 2003 tentang pembahasan teori pendidikan yang terdiri sebagai berikut :

Tabel 2. Kategori Pendidikan Orang Tua

No	Sekolah	Kategori
1	SD	Pendidikan Dasar
2	SMP	Pendidikan Dasar
3	SMA	Pendidikan Menengah
4	Diploma	Pendidikan Tinggi
5	S1	Pendidikan Tinggi
6	S2	Pendidikan Tinggi
7	S3	Pendidikan Tinggi

c. Variabel Penghasilan Orang Tua

Variabel penghasilan orang tua merupakan jumlah pendapatan atau penghasilan orang tua per bulan dari calon mahasiswa baru yang mendaftar. Untuk variabel penghasilan orang tua dilakukan dengan memperhatikan data statistik Badan Pusat Statistik (BPS) Tahun 2011 yang digolongkan ke dalam 4 (empat) kelompok yaitu :

Tabel 3. Kategori Penghasilan Orang Tua

No	Rentang	Kategori
1	> 3.500.000	Pendapatan Sangat Tinggi
2	2.500.000 – 3.500.00	Pendapatan Tinggi
3	1.500.000 – 2.500.000	Pendapatan Sedang
4	< 1.500.000	Pendapatan Rendah

d. Variabel Nilai Ujian Nasional

Variabel ujian nasional merupakan nilai ujian nasional yang didapatkan calon mahasiswa baru yang mendaftar pada saat lulus SMA. Untuk variabel ujian nasional, merujuk pada aturan kategori Ujian Akhir Nasional yang digolongkan dalam 4 kategori, yaitu :

Tabel 4. Kategori Nilai Ujian Nasional

No	Nilai Ujian	Kategori
1	>85-100	Sangat Baik
2	>70-85	Baik
3	55-70	Cukup
4	< 55	Kurang

e. Variabel Nilai Tes

Variabel nilai tes merupakan nilai yang didapatkan semua calon mahasiswa baru yang mendaftar. Sumber nilai tes didapatkan dari hasil nilai tes untuk pengetahuan umum, agama dan pengetahuan bahasa. Untuk variabel nilai tes dilakukan berdasarkan aturan yang ada pada IAIN Bone, yang terdiri dari kategori :

Tabel 5. Kategori Nilai Tes

No	Nilai Tes	Kategori
1	85-100	Sangat Baik
2	75-84	Baik
3	60-74	Cukup
4	< 60	Kurang

f. Variabel Nilai Wawancara

Variabel nilai wawancara merupakan nilai dari hasil wawancara yang dilakukan calon mahasiswa baru dengan pewawancara yang sudah ditunjuk pada IAIN Bone. Untuk variabel nilai wawancara juga merujuk pada aturan yang ada pada IAIN Bone, yang terdiri dari kategori :

Tabel 6. Kategori Nilai Wawancara

No	Nilai Wawancara	Kategori
1	85-100	Sangat Baik
2	75-84	Baik
3	60-74	Cukup
4	< 60	Kurang

g. Variabel Nilai Baca Tulis Huruf Qur'an (BTHQ)

Variabel nilai BTHQ merupakan nilai dari hasil membaca dan menulis huruf qur'an. Untuk variabel nilai BTHQ juga merujuk pada aturan yang ada di IAIN Bone, yang terdiri dari kategori :

Tabel 7. Kategori Nilai BTHQ

No	Nilai BTHQ	Kategori
1	85-100	Lancar, Fasih, Tajwid Bagus
2	75-84	Lancar, Fasih, Tajwid Kurang Bagus
3	60-74	Tidak Lancar
4	<60	Tidak Bisa

3. Proses Pengolahan Data

Setelah dilakukan studi literatur kemudian dilakukan pengolahan data dengan menentukan data latih dan data uji. Data latih ini diimplementasikan dengan menerapkan persamaan algoritma Naive Bayes dan C.45. Kemudian melakukan

perhitungan akurasi dari kedua algoritma tersebut dan dilanjutkan dengan melakukan perhitungan data uji untuk menentukan prediksi kelulusan dari yang berlabel kelas maupun yang tidak berlabel kelas. Selanjutnya dilakukan perhitungan nilai precision, recall dan akurasi. Dari perhitungan tersebut kemudian dilakukan perbandingan algoritma yang digunakan.

4. Implementasi Naive Bayes

Setelah perancangan dan pemodelan sistem, kemudian menerjemahkan desain yang telah terbentuk menjadi sebuah sistem dengan menerapkan algoritma Naive Bayes. Algoritma Naive Bayes ini merupakan sebuah metode klasifikasi menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes. Algoritma Naive Bayes memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya sehingga dikenal sebagai Teorema Bayes. Ciri utama dari Naive Bayes Classifier ini adalah asumsi yang sangat kuat (naif) akan independensi darimasing-masing kondisi/kejadian.

Secara garis besar Jedhav et all (2016), algortima Naive Bayes dapat dijelaskan seperti persamaan (1).

$$P(R/S) = \frac{P(R)P(S/R)}{P(S)} \quad (1)$$

Keterangan :

R : Data yang belum diketahui kelasnya

S : Hipotesis pada data R yang merupakan class khusus

P(R/S) : Nilai probabilitas pada hipotesis R yang berdasarkan kondisi S

P(R) : Nilai probabilitas pada hipotesis R

P(S/R) : Nilai probabilitas S yang berdasarkan dengan kondisi hipotesis R

$P(S)$: Nilai probabilitas S

Dengan menggunakan persamaan di atas, data yang telah diperoleh dapat diproses dengan algoritman Naive Bayes untuk penilaian data yang akan diklasifikasikan.

5. Implementasi C.45

Pada tahap ini, menerjemahkan desain yang telah terbentuk menjadi sebuah sistem dengan menerapkan algoritma C.45. Purushottam, et al (2016) menyatakan algoritma C.45 merupakan algoritma yang dipergunakan dalam membentuk decision tree (pengambilan keputusan). Algoritma C.45 adalah salah satu algoritma dalam induksi decision tree yaitu ID3 (Iterative Dichotomiser 3) yang dikembangkan oleh J.Ross Quinlan. Dalam prosedur algoritma ID3, input berupa sampel training, label training dan atribut. Algoritma C.45 ini, merupakan pengembangan dari ID3. Dasar algoritma ini adalah pembuatan pohon keputusan berdasarkan pemilihan atribut yang memiliki prioritas tertinggi atau dapat disebut memiliki nilai gain tertinggi berdasarkan nilai entropy atribut tersebut sebagai poros atribut klasifikasi. Kemudian secara rekursif cabang-cabang pohon diperluas sehingga seluruh pohon terbentuk.

Terdapat empat langkah dalam proses pembuatan pohon keputusan pada algoritma C.45, yaitu :

- a. Memilih atribut sebagai akar
- b. Membuat cabang untuk masing-masing nilai
- c. Membagi setiap kasus dalam cabang

d. Mengulangi proses dalam setiap cabang sehingga semua kasus dalam cabang memiliki kelas yang sama

Kemudian dilakukan perhitungan untuk mencari nilai entropy dan gain.

Berikut rumus untuk mencari nilai entropy dan gain, sebagai berikut :

$$Entropy(S) = \sum_{j=1}^k -p_j \log_2 p_j \quad (2)$$

Keterangan :

S : Himpunan kasus

k : Jumlah partisi S

p_j : Jumlah kasus pada partisi ke-j

Persamaan (2) adalah persamaan yang digunakan dalam perhitungan entropy untuk menentukan heterogenity dari sebuah kumpulan data sample (Amin et al, 2015).

Persamaan (3) merupakan rumus yang digunakan dalam perhitungan gain setelah melakukan perhitungan entropy. Berikut rumus mencari nilai gain :

$$Gain(A) = Entropy(S) - \sum_{j=1}^k \frac{S_i}{S} * Entropy(S_i) \quad (3)$$

Keterangan :

A : Atribut dari dataset

k : Jumlah partisi S

S : Himpunan Kasus

Dari rumus yang ada di atas data yang diperoleh dapat dimasukkan dan diproses dengan algoritma C.45 untuk proses pembuatan decision tree.

6. Pengujian Precision, Recall dan Accuracy Naive Bayes dan C.45

Dalam dunia pengenalan pola (pattern recognition) dan temu kembali informasi (information retrieval), precision dan recall adalah dua perhitungan yang banyak digunakan kinerja dari sistem/metode yang digunakan. Precision adalah tingkat ketepatan antara informasi yang diminta oleh pengguna dengan jawaban yang diberikan oleh sistem. Sedangkan recall adalah tingkat keberhasilan sistem dalam menemukan kembali sebuah informasi. Sedangkan di dunia lain seperti dunia statistika dikenal juga istilah accuracy yang didefinisikan sebagai tingkat kedekatan antara nilai prediksi dengan nilai aktual.

Kurniawan, et al (2018) menyatakan pengujian sebuah algoritma membutuhkan standar dan alat uji. Membandingkan dua buah algoritma harus memiliki standar yang sama sehingga dapat diketahui algoritma yang terbaik dari perbandingan tersebut. Oleh karena itu, pada tahap ini dilakukan pengujian precision, recall dan akurasi agar standar dari dua algoritma yang digunakan memiliki kesamaan.

Pengujian dengan menghitung nilai precision, recal dan accuracy dari algoritma Naive Bayes dan C.45, dilakukan dengan membagi data menjadi dua yaitu data training atau data latihan dan data testing atau data uji. Data training digunakan untuk sebagai data rujukan dalam perhitungan setiap algoritma, sedangkan data testing digunakan untuk menilai prediksi maupun penentuan yang dilakukan oleh setiap algoritma sudah tepat atau tidak. Setelah data training dan

data testing sudah dilakukan pengujian, kemudian dihitung nilai precision, recall dan accuracy pada masing-masing algoritma.

Precision merupakan perhitungan terhadap perkiraan proporsi kasus positif yang benar dan dirumuskan dalam persamaan (4) :[8]

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4)$$

Recall merupakan perhitungan terhadap perkiraan proporsi kasus positif yang diidentifikasi benar dan dirumuskan dalam persamaan (5)

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

Accuracy merupakan perhitungan terhadap proporsi dari jumlah total prediksi yang benar dan dirumuskan dalam persamaan (6)

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

Keterangan :

TP : True Positive

TN : True Negative

FP : False Positive

FN : False Negative

Persamaan di atas dapat digambarkan seperti pada tabel 9 di bawah ini :

Tabel 8. Tabel Persamaan Precision, Recall dan Akurasi

		Nilai Sebenarnya	
		TP (True Positive)	FP (False Positive)
Nilai Prediksi	TRUE	TP (True Positive)	FP (False Positive)
	FALSE	FN (False Negative)	TN (True Negative)

7. Pengujian Sistem

Pengujian sistem pada tahap terakhir dilakukan pengujian secara keseluruhan, dimana model pengujian dilakukan dengan cara manual ke dalam microsoft excel dan blackbox untuk melihat fungsionalitas aplikasi secara langsung.

8. Perbandingan Algoritma

Pada tahap ini, dilakukan perbandingan nilai precision, recall dan accuracy pada masing-masing algoritma kemudian dilakukan rekapitulasi hasil dari masing-masing algoritma untuk membandingkan kedua algoritma tersebut dan dapat dilihat mana hasil dari algoritma yang terbaik.

9. Pencarian Pola Sequensial

Pencarian pola sequensial dilakukan dengan mencari sejumlah event yang secara umum terjadi bersama-sama. Jika diberikan sekumpulan obyek, dengan masing-masing obyek dihubungkan dengan waktu kejadiannya, maka dapatkan pola yang memprediksi ketergantungan sekuensial (sequential dependences) yang kuat diantara kejadian-kejadian yang berbeda.

Pola-pola sekuensial pertama, pada dasarnya dibentuk dengan cara mencari semua kemungkinan pola yang ada. Nilai-nilai kejadian dalam pola diatur berdasarkan urutan waktu kejadian.

B. Waktu dan Lokasi Penelitian

Waktu pelaksanaan penelitian dilaksanakan selama 2 (dua) bulan dimulai pertengahan bulan April 2019 sampai pertengahan bulan Juni 2019 yang dilakukan pada Institut Agama Islam Negeri (IAIN) Bone.

C. Jenis Penelitian

Jenis penelitian ini adalah penelitian eksperimental dimana ruang lingkup masalah dapat dilakukan dengan metode studi pustaka (library research), metode pengumpulan data lapangan (field research) dan pengujian sistem untuk algoritma yang digunakan.

D. Sumber Data

Pengumpulan data terdiri dari :

1. Data Primer

Data yang digunakan merupakan data yang bersumber dari data Sistem Informasi (SISFO) Akademik IAIN Bone, dimana terdapat data dari semua siswa SMA yang dinyatakan lulus dan tidak lulus dengan indikator Tahun Lulus, Pendidikan Orang Tua, Penghasilan Orang Tua, Nilai Ujian Nasional, Nilai Tes, Nilai Wawancara dan Nilai BTHQ pada Jalur Mandiri IAIN Bone Tahun 2018.

2. Data Sekunder

Data sekunder dilakukan dengan melakukan pencarian literatur, baik dari buku, jurnal maupun internet. Selain itu mempelajari literatur yang berkaitan dengan data mining, klasifikasi algoritma Naive Bayes dan C.45.

E. Instrumen Penelitian

Instrumen yang digunakan pada penelitian ini adalah :

1. Software

- a. PHP dan Database MySql
- b. Windows 10 Pro

c. Microsoft Excel 2010

2. Hardware

a. Processor Intel(R) Core(TM) i5-8250U

b. RAM 4 GB

c. 64-bit OS, X64 based processor

BAB IV

HASIL DAN PEMBAHASAN

A. Implementasi Sistem

Data yang telah dikumpulkan sebanyak 180, yang terdiri dari data lulus dan tidak lulus akan diimplementasikan pada proses perhitungan data mining dengan menggunakan algoritma Naive Bayes dan algoritma C.45. Data yang ada dikumpulkan berdasarkan kategori, kemudian dilakukan proses perhitungan sesuai dengan formula yang berlaku pada kedua algoritma tersebut.

Variabel yang digunakan yaitu asal sekolah, tahun lulus, pendidikan orang tua, penghasilan orang tua, nilai ujian nasional, nilai tes meliputi tes pengetahuan umum, bahasa dan agama, nilai tes wawancara dan nilai tes baca tulis huruf qur'an.

B. Pengolahan Data

Pada proses ini, akan dilakukan pengolahan data berdasarkan algoritma yang digunakan yaitu Naive Bayes dan C.45.

1. Algoritma Naive Bayes

Pada algoritma naive bayes pengolahan data dilakukan dengan menggunakan persamaan sebagai berikut :

$$P(R/S) = \frac{P(R)P(S/R)}{P(S)}$$

Keterangan :

R : Data yang belum diketahui kelasnya

S : Hipotesis pada data R yang merupakan class khusus

P(R/S) : Nilai probabilitas pada hipotesis R yang berdasarkan kondisi S

P(R) : Nilai probabilitas pada hipotesis R

P(S/R) : Nilai probabilitas S yang berdasarkan dengan kondisi hipotesis R

P(S) : Nilai probabilitas S

Untuk perhitungan data yang ada, yaitu :

a. Penentuan nilai persentase lulus dan tidak lulus

Lulus = $115/170 = 0,63889 = 64\%$

Tidak Lulus = $65/170 = 0,36111 = 36\%$

b. Penentuan nilai persentase dari variabel yang ada

Tabel 9. Perhitungan Algoritma Naive Bayes

No	Variabel	Kategori	Lulus			Tidak Lulus		
			Total Data	Hitung	Nilai	Total Data	Hitung	Nilai
1	Tahun Lulus	2014	1	$1/109=0,00917$	1%	0	$0/61=0$	0%
		2015	0	$0/109=0$	0%	2	$2/61=0,03279$	3%
		2016	4	$4/109=0,0367$	4%	7	$7/61=0,11475$	11%
		2017	25	$25/109=0,22936$	23%	6	$6/61=0,09836$	10%
		2018	79	$79/109=0,72477$	72%	46	$46/61=0,7541$	75%
Jumlah					100%			100%
2	Pendidikan Orang Tua	Dasar	69	$69/109=0,63303$	63%	33	$33/61=0,54098$	54%
		Menengah	22	$22/109=0,20183$	20%	12	$12/61=0,19672$	20%
		Tinggi	18	$18/109=0,16514$	17%	16	$16/61=0,2623$	26%
Jumlah					100%			100%
3	Penghasilan Orang Tua	Rendah	75	$75/109=0,68807$	69%	41	$41/61=0,67213$	67%
		Sedang	21	$21/109=0,19266$	19%	8	$8/61=0,13115$	13%
		Tinggi	3	$3/109=0,02752$	3%	11	$11/61=0,18033$	18%
		Sangat Tinggi	10	$10/109=0,09174$	9%	1	$1/61=0,01639$	2%
Jumlah					100%			100%
4	Nilai Ujian Nasional	Baik	64	$64/109=0,58716$	59%	40	$40/61=0,65574$	66%
		Sangat Baik	45	$46/109=0,42202$	41%	21	$21/61=0,34426$	34%
Jumlah					100%			100%
5	Nilai Tes	Cukup	7	$7/109=0,06422$	6%	53	$53/61=0,86885$	87%
		Baik	20	$20/109=0,18349$	18%	8	$8/61=0,13115$	13%
		Sangat Baik	82	$82/109=0,75229$	75%	0	$0/61=0$	0%
Jumlah					100%			100%
6	Nilai Wawancara	Kurang	0	$0/109=0$	0%	4	$4/61=0,06557$	7%
		Cukup	0	$0/109=0$	0%	9	$9/61=0,14754$	15%
		Baik	13	$13/109=0,11927$	12%	41	$41/61=0,67213$	67%
		Sangat Baik	96	$96/109=0,88073$	88%	7	$7/61=0,11475$	11%
Jumlah					100%			100%
7	Nilai Baca Tulis Huruf Qur'an	Lancar, Fasih, Tajwid Bagus	107	$107/109=0,98165$	98%	0	$0/61=0$	0%
		Lancar, Fasih, Tajwid Kurang Bagus	2	$2/109=0,01835$	2%	0	$0/61=0$	0%
		Tidak Lancar	0	$0/109=0$	0%	61	$61/61=1$	100%
Jumlah					100%			100%

c. Mengalikan semua hasil variabel Lulus dan Tidak Lulus

Pada proses ini mengalikan semua hasil variabel lulus dan tidak lulus dengan memasukkan data testing yang telah dibuat (lihat Tabel 15. Data Testing).

Untuk perhitungan data testing secara keseluruhan dapat dilihat pada halaman lampiran.

- *Variabel Lulus*

P (Tahun Lulus=2018), (Pendidikan Orang Tua=Tinggi), (Penghasilan Orang Tua=Rendah), (Nilai Ujian Nasional=Baik), (Nilai Tes=Baik), (Nilai Wawancara=Sangat Baik), (Nilai BTHQ=Lancar,Fasih,Tajwid Bagus)|Lulus
 $=0,7247*0,1651*0,6880*0,5871*0,1834*0,8807*0,98165*0,6411$
 $=0,492\%$

- *Variabel Tidak Lulus*

P(Tahun Lulus=2018), (Pendidikan Orang Tua=Tinggi), (Penghasilan Orang Tua=Rendah), (Nilai Ujian Nasional=Baik), (Nilai Tes=Baik), (Nilai Wawancara=Sangat Baik), (Nilai BTHQ= Lancar,Fasih,Tajwid Bagus)|Tidak Lulus
 $=0,7247*0,1651*0,6880*0,5871*0,1183*0*0,8807*0,9816*0,3588$
 $=0,275\%$

d. *Membandingkan hasil class Lulus dan Tidak Lulus*

Hasil perhitungan antara class Lulus dengan Tidak Lulus, maka dihasilkan (P|Lulus) lebih besar dari (P|Tidak Lulus), maka keputusannya dari uji coba diatas menghasilkan “LULUS”

2. Algoritma C.45

Pada algoritma C.45 akan dilakukan perhitungan untuk mencari nilai entropy dan gain. Entropy merupakan jumlah bit yang diperkirakan dibutuhkan untuk dapat mengekstrak suatu kelas (+ atau -) dari sejumlah data acak pada ruang sampel S. Entropy dapat dikatakan sebagai kebutuhan bit untuk menyatakan suatu kelas dan digunakan untuk mengukur ketidakkaslian S. Berikut rumus untuk mencari nilai entropy, sebagai berikut :

$$Entropy(S) = \sum_{j=1}^k -p_j \log_2 p_j$$

Keterangan :

S : Himpunan kasus

k : Jumlah partisi S

p_j : Jumlah kasus pada partisi ke-j

sedangkan gain merupakan perolehan informasi dari atribut A relative terhadap output data S. Perolehan informasi didapat dari output data atau variabel dependen S yang dikelompokkan berdasarkan atribut A, dinotasikan dengan gain (S,A). Rumus untuk mencari nilai gain, sebagai berikut :

$$Gain(A) = Entropy(S) - \sum_{j=1}^k \frac{S_j}{S} * Entropy(S_j)$$

Keterangan :

A : Atribut dari dataset

k : Jumlah partisi S

S : Himpunan Kasus

Untuk perhitungan data yang ada, yaitu :

- a. Berdasarkan data yang ada terlebih dahulu akan dibuat calon percabangan yang mungkin terjadi

Tabel 10. Tabel Penentuan Calon Cabang

Penentuan Calon Cabang						
1	Tahun Lulus	2014	2015	2016	2017	2018
2	Pendidikan Orang Tua	Dasar	Menengah	Tinggi		
3	Penghasilan Orang Tua	Rendah	Sedang	Tinggi	Sangat Tinggi	
4	Nilai Ujian Nasional	Baik	Sangat Baik			

5	Nilai Tes	Cukup	Baik	Sangat Baik		
6	Nilai Wawancara	Kurang	Cukup	Baik	Sangat Baik	
7	Nilai Baca Tulis Huruf Qur'an	Lancar, Fasih, Tajwid Bagus	Lancar, Fasih, Tajwid Krng Bagus	Tidak Lancar		

Dari tabel di atas menunjukkan, bahwa ada delapan calon percabang yang mungkin terjadi, namun percabangan akan terjadi dengan melihat nilai gain masing-masing.

- b. Menghitung frekuensi yang muncul serta nilai entropy dari data yang ada

Proses selanjutnya melakukan perhitungan nilai frekuensi dan entropy dari data yang ada.

Tabel 11. Tabel Hasil Perhitungan Nilai Entrophy

Hasil	Frekuensi	Pj	log₂pj	-pj.log₂pj
Lulus	109	0.641176471	-0.641206611	0.411126592
Tidak Lulus	61	0.358823529	-1.478653599	0.530575703
Total Data	170	Total Entrophy (S)		0,941702295

Keterangan :

- Lulus

$$P_j = 109/170 = 0,64118$$

$$\log_2 p_j = \log_2(0,64118) = -0,641206611$$

$$-p_j \cdot \log_2 p_j = (-0,64118) \cdot -0,641206611 = \mathbf{0,411126592}$$

- Tidak Lulus

$$P_j = 61/170 = 0.358823529$$

$$\log_2 p_j = \log_2(0.358823529) = -1.478653599$$

$$-p_j \cdot \log_2 p_j = (-0.358823529) \cdot (-1.478653599) = \mathbf{0.530575703}$$

Nilai Entrophy

$$= (-p_j \cdot \log_2 p_j | \text{lulus}) + (-p_j \cdot \log_2 p_j | \text{tidak lulus})$$

$$= \mathbf{0,411126592 + 0.530575703}$$

$$= \mathbf{0,941702295}$$

- c. Menghitung gain dari masing-masing variabel yang ada

Setelah mendapatkan nilai entrophy, kemudian dilakukan perhitungan dari masing-masing variabel untuk mencari nilai gain masing-masing dan selanjutnya menentukan nilai gain tertinggi.

1. Variabel Tahun Lulus

- 2014

$$= \frac{-1}{1} \log_2 \frac{1}{1} + \left(\frac{-0}{1} \right) * \text{Log}_2 \frac{0}{1}$$

$$= 0 + 0 = 0$$

$$= 0 * 1/170$$

$$= 0$$

- 2015

$$= \frac{-0}{2} \log_2 \frac{0}{2} + \left(\frac{-2}{2} \right) * \text{Log}_2 \frac{2}{2}$$

$$= 0 + 0 = 0$$

$$= 0 * 2/170$$

$$= 0$$

- 2016

$$\begin{aligned}
 &= \frac{-4}{11} \log_2 \frac{4}{11} + \left(-\frac{7}{11} * \text{Log}_2 \frac{7}{11} \right) \\
 &= 0,530702407 + 0,414957898 = 0.94566 \\
 &= 0,94566 * 11/170 \\
 &= 0,0611897647
 \end{aligned}$$

- 2017

$$\begin{aligned}
 &= \frac{-25}{31} \log_2 \frac{25}{31} + \left(-\frac{6}{31} * \text{Log}_2 \frac{6}{31} \right) \\
 &= 0,250274291 + 0,458561383 = 0,684038436 \\
 &= 0.708836 * 33/170 \\
 &= 0.129258
 \end{aligned}$$

- 2018

$$\begin{aligned}
 &= \frac{-79}{125} \log_2 \frac{79}{125} + \left(-\frac{46}{125} * \text{Log}_2 \frac{46}{125} \right) \\
 &= 0.418386235 + 0.530737817 = 0.949124 \\
 &= 0.949124 * 125/170 \\
 &= 0.697885
 \end{aligned}$$

Nilai Gain untuk variabel Tahun Lulus, yaitu :

$$\begin{aligned}
 &= \mathbf{0,941702295} - \left(\frac{1}{170} * (0) + \frac{2}{170} * (0) + \frac{11}{170} * (0.94566) \right. \\
 &\quad \left. + \frac{31}{170} * (0,684038436) + \frac{125}{170} * (0,949124) \right) \\
 &= \mathbf{0,941702295} - (0 + 0 + 0,06118 + 0,29258 + 0,697885) \\
 &= \mathbf{0,941702295} - \mathbf{0.888333} \\
 &= \mathbf{0.053369}
 \end{aligned}$$

2. Variabel Pendidikan Orang Tua

- Dasar

$$\begin{aligned}
 &= \frac{-69}{102} \log_2 \frac{69}{102} + \left(-\frac{33}{102} * \text{Log}_2 \frac{33}{102} \right) \\
 &= 0.381462364 + 0.526715984 = 0.908178 \\
 &= 0.908178 * 102/170
 \end{aligned}$$

$$= 0,544907$$

- Menengah

$$= \frac{-22}{34} \log_2 \frac{22}{34} + \left(\frac{-12}{34} * \text{Log}_2 \frac{12}{34} \right)$$

$$= 0,406373144 + 0,530294238 = 0,936667$$

$$= 0,936667 * 34/170$$

$$= 0,187333$$

- Tinggi

$$= \frac{-18}{34} \log_2 \frac{18}{34} + \left(\frac{-16}{34} * \text{Log}_2 \frac{16}{34} \right)$$

$$= 0,485755327 + 0,511747219 = 0,997503$$

$$= 0,997503 * 34/170$$

$$= 0,199501$$

Nilai Gain untuk variabel Pendidikan Orang Tua, yaitu :

$$= \mathbf{0,941702295} - \left(\frac{102}{170} * (0) + \frac{34}{170} * (0) + \frac{34}{170} * (0,94566) \right)$$

$$= \mathbf{0,941702295} - (0,544907 + 0,187333 + 0,199501)$$

$$= \mathbf{0,941702295} - \mathbf{0,931741}$$

$$= \mathbf{0,009961}$$

3. Variabel Penghasilan Orang Tua

- Rendah

$$= \frac{-75}{116} \log_2 \frac{75}{116} + \left(\frac{-41}{116} * \text{Log}_2 \frac{41}{116} \right)$$

$$= 0,406785973 + 0,53032404 = 0,93711$$

$$= 0,93711 * 116/170$$

$$= 0,63944$$

- Sedang

$$= \frac{-21}{29} \log_2 \frac{21}{29} + \left(\frac{-8}{29} * \text{Log}_2 \frac{8}{29} \right)$$

$$= 0,337204656 + 0,512546481 = 0,849751$$

$$= 0.849751 * 29/170$$

$$= 0.144958$$

- Tinggi

$$= \frac{-3}{14} \log_2 \frac{3}{14} + \left(-\frac{-11}{14} * \text{Log}_2 \frac{11}{14}\right)$$

$$= 0,476226947 + 0,27336831 = 0,749595257$$

$$= 0,749595257 * 14/170$$

$$= 0.061731$$

- Sangat Tinggi

$$= \frac{-10}{11} \log_2 \frac{10}{11} + \left(-\frac{-1}{11} * \text{Log}_2 \frac{1}{11}\right)$$

$$= 0.125003203 + 0.314493784 = 0.439497$$

$$= 0.439497 * 11/170$$

$$= 0.028438$$

Nilai Gain untuk variabel Penghasilan Orang Tua, yaitu :

$$= \mathbf{0, 941702295} - \left(\frac{116}{170} * (0.93711) + \frac{29}{170} * (0.849751) + \frac{14}{170} * (0,749595257) + \frac{11}{170} * (0.439497)\right)$$

$$= \mathbf{0, 941702295} - (0.63944 + 0.144958 + 0.061731 + 0.028438)$$

$$= \mathbf{0, 941702295} - \mathbf{0.874567}$$

$$= \mathbf{0.067136}$$

4. Variabel Nilai Ujian Nasional

- Baik

$$= \frac{-64}{104} \log_2 \frac{64}{104} + \left(-\frac{-40}{104} * \text{Log}_2 \frac{40}{104}\right)$$

$$= 0.431039827 + 0.530196778 = 0.961237$$

$$= 0.961237 * 104/170$$

$$= 0.588051$$

- Sangat Baik

$$= \frac{-45}{66} \log_2 \frac{45}{66} + \left(-\frac{-21}{66} * \text{Log}_2 \frac{21}{66}\right)$$

$$\begin{aligned}
&= 0.376732516 + 0.525660767 = 0.902393 \\
&= 0.902393 * 66/170 \\
&= 0.350341
\end{aligned}$$

Nilai Gain untuk variabel Penghasilan Orang Tua, yaitu :

$$\begin{aligned}
&= \mathbf{0,941702295} - \left(\frac{104}{170} * (0.961237) + \frac{66}{170} * (0.902393) \right) \\
&= \mathbf{0,941702295} - (0.588051 + 0.350341) \\
&= \mathbf{0,941702295} - \mathbf{0.938392} \\
&= \mathbf{0.003311}
\end{aligned}$$

5. Variabel Nilai Tes

- Cukup

$$\begin{aligned}
&= \frac{-7}{60} \log_2 \frac{7}{60} + \left(-\frac{53}{60} * \text{Log}_2 \frac{53}{60} \right) \\
&= 0.361612495 + 0.158090291 = 0.519703 \\
&= 0.519703 * 60/170 \\
&= 0.183425
\end{aligned}$$

- Baik

$$\begin{aligned}
&= \frac{-20}{28} \log_2 \frac{20}{28} + \left(-\frac{8}{28} * \text{Log}_2 \frac{8}{28} \right) \\
&= 0.346733448 + 0.516387121 = 0.863121 \\
&= 0.863121 * 28/170 \\
&= 0.142161
\end{aligned}$$

- Sangat Baik

$$\begin{aligned}
&= \frac{-82}{82} \log_2 \frac{82}{82} + \left(-\frac{0}{82} * \text{Log}_2 \frac{0}{82} \right) \\
&= 0 + 0 = 0 \\
&= 0 * 82/170 \\
&= 0
\end{aligned}$$

Nilai Gain untuk variabel Nilai Tes, yaitu :

$$= \mathbf{0,941702295} - \left(\frac{60}{170} * (0.519703) + \frac{28}{170} * (0.863121) + \frac{82}{170} * (0) \right)$$

$$\begin{aligned}
&= \mathbf{0,941702295} - (0.183425 + 0.142161 + 0) \\
&= \mathbf{0,941702295} - \mathbf{0.325586} \\
&= \mathbf{0.616117}
\end{aligned}$$

6. Variabel Nilai Wawancara

- Kurang

$$\begin{aligned}
&= \frac{-0}{4} \log_2 \frac{0}{4} + \left(\frac{-4}{4} * \text{Log}_2 \frac{4}{4} \right) \\
&= 0 + 0 = 0 \\
&= 0 * 4/170 \\
&= 0
\end{aligned}$$

- Cukup

$$\begin{aligned}
&= \frac{-0}{9} \log_2 \frac{0}{9} + \left(\frac{-9}{9} * \text{Log}_2 \frac{9}{9} \right) \\
&= 0 + 0 = 0 \\
&= 0 * 9/170 \\
&= 0
\end{aligned}$$

- Baik

$$\begin{aligned}
&= \frac{-13}{54} \log_2 \frac{13}{54} + \left(\frac{-41}{54} * \text{Log}_2 \frac{41}{54} \right) \\
&= 0.494589281 + 0.301680656 = 0.79627 \\
&= 0.79627 * 54/170 \\
&= 0.252933
\end{aligned}$$

- Sangat baik

$$\begin{aligned}
&= \frac{-96}{103} \log_2 \frac{96}{103} + \left(\frac{-7}{103} * \text{Log}_2 \frac{7}{103} \right) \\
&= 0.094637384 + 0.263631255 = 0.358269 \\
&= 0.358269 * 103/170 \\
&= 0.217069
\end{aligned}$$

Nilai Gain untuk variabel Nilai Wawancara, yaitu :

$$\begin{aligned} &= \mathbf{0,941702295} - \left(\frac{4}{170} * (0) + \frac{9}{170} * (0) + \frac{54}{170} * (0.79627) + \frac{103}{170} * (0.358269) \right) \\ &= \mathbf{0,941702295} - (0 + 0 + 0.252933 + 0.217069) \\ &= \mathbf{0,941702295} - \mathbf{0.470001} \\ &= \mathbf{0.471701} \end{aligned}$$

7. Variabel Nilai BTHQ

- Lancar, Fasih, Tajwid Bagus

$$\begin{aligned} &= \frac{-107}{107} \log_2 \frac{107}{107} + \left(\frac{-0}{107} * \text{Log}_2 \frac{0}{107} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 107/170 \\ &= 0 \end{aligned}$$

- Lancar, Fasih, Tajwid Kurang Bagus

$$\begin{aligned} &= \frac{-2}{2} \log_2 \frac{2}{2} + \left(\frac{-0}{2} * \text{Log}_2 \frac{0}{2} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 2/170 \\ &= 0 \end{aligned}$$

- Tidak Lancar

$$\begin{aligned} &= \frac{-0}{61} \log_2 \frac{0}{61} + \left(\frac{-61}{61} * \text{Log}_2 \frac{61}{61} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 61/170 \\ &= 0 \end{aligned}$$

Nilai Gain untuk variabel Nilai BTHQ, yaitu :

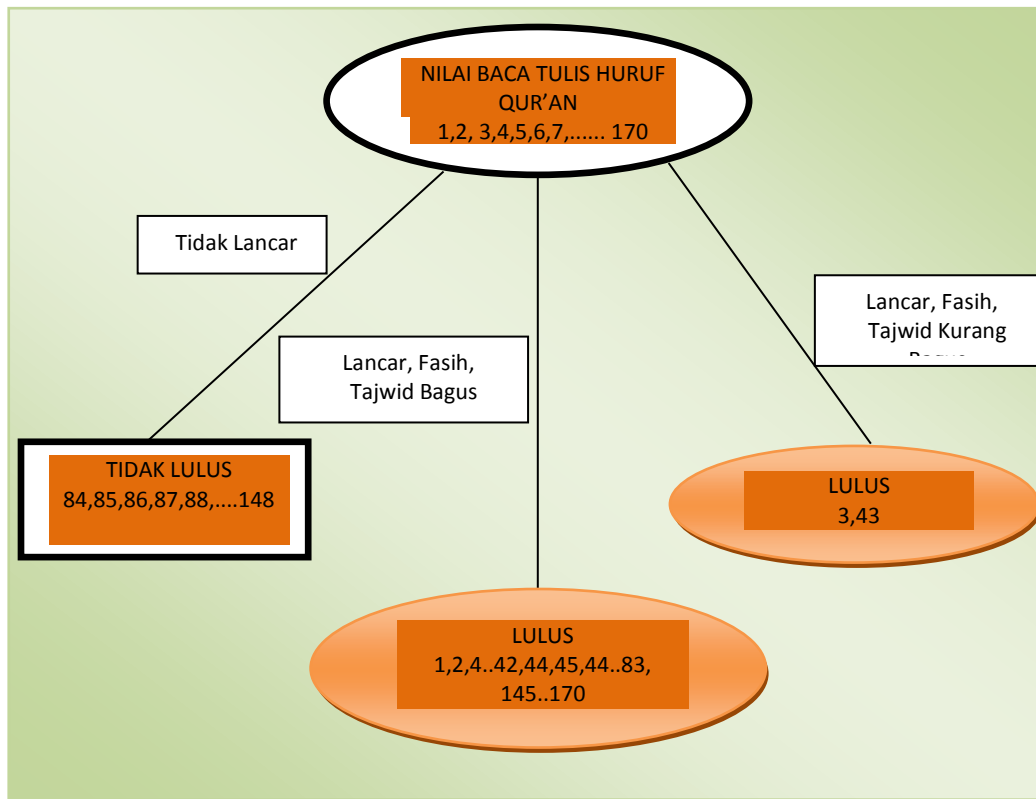
$$\begin{aligned} &= \mathbf{0,941702295} - \left(\frac{107}{170} * (0) + \frac{2}{170} * (0) + \frac{61}{170} * (0) \right) \\ &= \mathbf{0,941702295} - (0 + 0 + 0) \\ &= \mathbf{0,941702295} - \mathbf{0} \\ &= \mathbf{0,9417} \end{aligned}$$

Tabel 12. Perhitungan Algoritma C.45

CABANG			LULUS	TIDAK LULUS	P(LULUS)	P(TIDAK LULUS)	- P(LULUS)*LOG2 P(LULUS)	P(TIDAK LULUS)*LOG2 P(TIDAK LULUS)	TOTAL	P(TAHUN LULUS=...)	TOTAL X P(TAHUN LULUS=...)	E(TOTAL * P(TAHUN LULUS=...))	GAIN
1	Tahun Lulus	2014	1	0	1.000000	0	0	0	0	0.005882	0	0.888333	0.0533
		2015	0	2	0	1	0	0	0	0.011765	0		
		2016	4	7	0.363636	0.63636	0.53070	0.414957	0.94566	0.064706	0.06119		
		2017	25	6	0.806451	0.19354	0.25027	0.458561	0.708836	0.182353	0.129258		
		2018	79	46		0.368	0.41838	0.530737	0.949124	0.735294	0.697885		
2	Pendidikan Orang Tua	Dasar	69	33	0.676470	0.32352	0.38146	0.526715	0.908178	0.6	0.544907	0.931741	0.0099
		Menengah	22	12	0.647058	0.35294	0.40637	0.530294	0.936667	0.2	0.187333		
		Tinggi	18	16	0.529411	0.47058	0.48575	0.511747	0.997503	0.2	0.199501		
3	Pendapatan Orang Tua	Rendah	75	41	0.646551	0.35344	0.40678	0.530324	0.93711	0.682353	0.63944	0.874567	0.0671
		Sedang	21	8	0.724137	0.27586	0.33720	0.512546	0.849751	0.170588	0.144958		
		Tinggi	3	11	0.214285	0.78571	0.47622	0.273368	0.749595	0.082353	0.061731		
		Sangat Tinggi	10	1	0.909090	0.09090	0.12500	0.314493	0.439497	0.064706	0.028438		
4	Nilai Ujian Nasional	Baik	64	40	0.615384	0.38461	0.43103	0.530196	0.961237	0.611765	0.588051	0.938392	0.0033
		Sangat Baik	45	21	0.681818	0.31818	0.37673	0.525660	0.902393	0.388235	0.350341		
5	Nilai Tes	Cukup	7	53	0.116666	0.88333	0.36161	0.158090	0.519703	0.352941	0.183425	0.325586	0.6161
		Baik	20	8	0.714285	0.28571	0.34673	0.516387	0.863121	0.164706	0.142161		
		Sangat Baik	82	0	1	0	0	0	0	0.482353	0		
6	Nilai Wawancara	Kurang	0	4	0	1	0	0	0	0.023529	0	0.470001	0.4717
		Cukup	0	9	0	1	0	0	0	0.052941	0		
		Baik	13	41	0.240740	0.75925	0.49458	0.301680	0.79627	0.317647	0.252933		
		Sangat Baik	96	7	0.932038	0.06796	0.09463	0.263631	0.358269	0.605882	0.217069		
7	Nilai BTHQ	Tidak Lancar	0	61	0	1	0	0	0	0.358824	0	0	0.9417
		Lancar,Fasih, Tajwid Kurang Bagus	2	0	1	0	0	0	0	0.011765	0		
		Lancar,Fasih, Tajwid Bagus	107	0	1	0	0	0	0	0.629412	0		

d. Membuat percabangan atau pohon keputusan dari gain yang tertinggi

Dari perhitungan yang telah dilakukan untuk mencari nilai gain dari variabel yang ada, maka didapatkan nilai gain tertinggi terdapat pada variabel Nilai Baca Tulis Huruf Qur'an dengan gain sebesar 0,9417 kemudian akan dibuat percabangan dari gain tersebut. Adapun cabang yang terbentuk, sebagai berikut :



Gambar 6. Gambar Pohon Keputusan

Percabangan diatas terlihat dua simbol yaitu, simbol kotak dan simbol elips. Simbol kotak menandakan keputusan final, artinya jika Nilai Baca Tulis Huruf Qur'an tidak lancar, maka hasil yang didapatkan pasti tidak lulus tapi jika Nilai Baca Tulis Huruf Qur'an Lancar, Fasih, Tajwid Bagus maka hasilnya bisa lulus

atau tidak begitupun jika Nilai Baca Tulis Huruf Qur'an Lancar, Fasih, Tajwid Kurang Bagus maka hasil yang didapatkan juga bisa lulus atau tidak dengan melihat variabel percabangan yang lain.

e. Melakukan Perhitungan Untuk Nilai Gain Tertinggi

Setelah terbentuk percabangan pertama, maka dilakukan kembali perhitungan nilai gain tertinggi dengan menghilangkan data yang sudah ada dalam simbol final. Dari 170 data yang ada, variabel Nilai Baca Tulis Huruf Qur'an untuk kategori tidak lancar dihilangkan sehingga data menjadi 109.

Adapun prosesnya sebagai berikut :

Hitung Frekuensi dan Nilai Entrophy

Tabel 13. Tabel Hasil Perhitungan Entrophy Gain Tertinggi

Hasil	Frekuensi	Pj	log ₂ pj	-pj.log ₂ pj
Lulus	109	1	0	0
Total Data	109	Total Entrophy (S)		0

Keterangan :

- Lulus

$$P_j = 109/109 = 1$$

$$\log_2 p_j = \log_2(1) = 0$$

$$-p_j \cdot \log_2 p_j = (-1) \cdot 0 = \mathbf{0}$$

- Nilai Entrophy

$$= (-p_j \cdot \log_2 p_j | \text{lulus}) + (-p_j \cdot \log_2 p_j | \text{tidak lulus})$$

$$= \mathbf{0}$$

Menghitung Nilai Gain

Variabel Tahun Lulus

- 2014

$$\begin{aligned} &= \frac{-1}{1} \log_2 \frac{1}{1} + \left(\frac{-0}{1} * \text{Log}_2 \frac{0}{1} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 1/109 \\ &= 0 \end{aligned}$$

- 2015

$$\begin{aligned} &= \frac{-0}{0} \log_2 \frac{0}{0} + \left(\frac{-0}{0} * \text{Log}_2 \frac{0}{0} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 0/109 \\ &= 0 \end{aligned}$$

- 2016

$$\begin{aligned} &= \frac{-4}{4} \log_2 \frac{4}{4} + \left(\frac{-0}{4} * \text{Log}_2 \frac{0}{4} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 4/109 \\ &= 0 \end{aligned}$$

- 2017

$$\begin{aligned} &= \frac{-25}{25} \log_2 \frac{25}{25} + \left(\frac{-0}{25} * \text{Log}_2 \frac{0}{25} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 25/109 \\ &= 0 \end{aligned}$$

- 2018

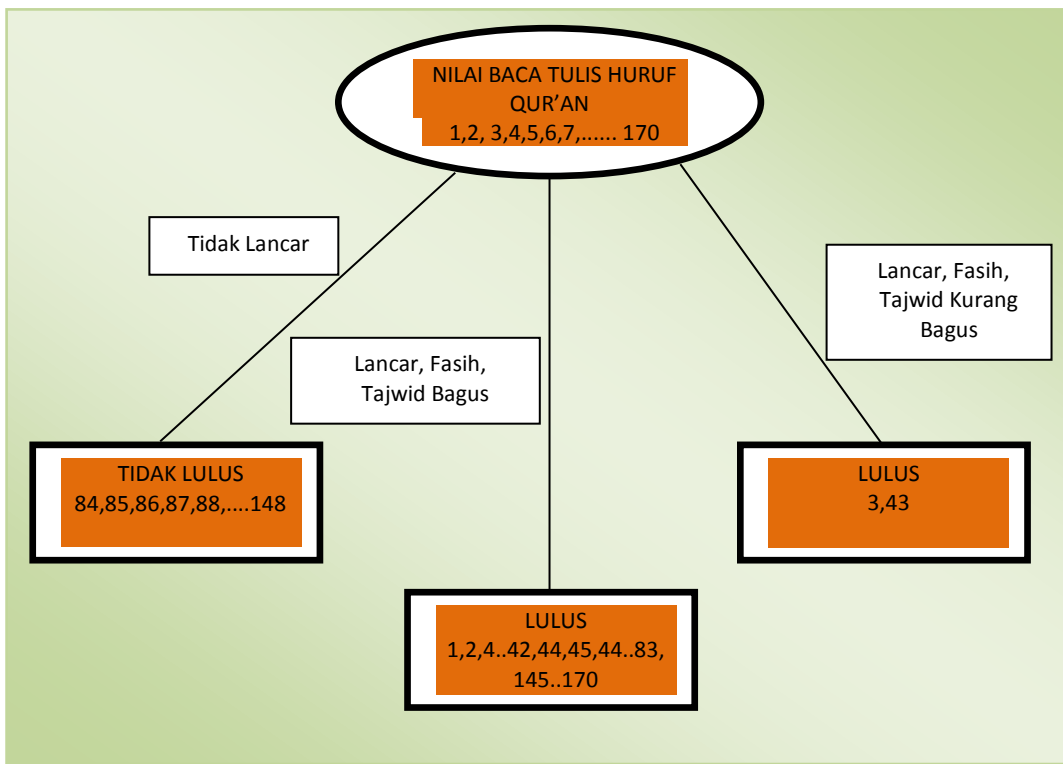
$$\begin{aligned} &= \frac{-79}{79} \log_2 \frac{79}{79} + \left(\frac{-0}{79} * \text{Log}_2 \frac{0}{79} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 79/109 \end{aligned}$$

$$= 0$$

Nilai Gain untuk variabel Tahun Lulus, yaitu :

$$\begin{aligned}
 &= 0 - \left(\frac{1}{109} * (0) + \frac{0}{109} * (0) + \frac{4}{109} * (0) \right. \\
 &\quad \left. + \frac{25}{109} * (0) + \frac{79}{109} * (0) \right) \\
 &= 0 - (0 + 0 + 0 + 0 + 0) \\
 &= 0 - 0 \\
 &= 0
 \end{aligned}$$

Dari perhitungan variabel Tahun Lulus, untuk semua kategori mengandung nilai 0 karena data untuk hasil Tidak Lulus sudah dihilangkan. Begitupun untuk nilai gain hasilnya juga mengandung Nilai 0. Jadi, hasil yang didapatkan untuk semua variabel memiliki nilai gain sama, yaitu 0. Sehingga hasil percabangan yang terbentuk hanya terdiri dari satu cabang. Berikut gambar pohon keputusan terakhir dan tabel hasil perhitungan dari gain tertinggi



Gambar 7. Pohon Keputusan Terakhir

Tabel 14. Perhitungan Algoritma C.45 Gain Tertinggi

CABANG			LULUS	TIDAK LULUS	P(LULUS)	P(TIDAK LULUS)	- P(LULUS) * LOG2 P(LULUS)	P(TIDAK LULUS) * LOG2 P(TIDAK LULUS)	TOTAL	P(ASAL SEKOLAH=...)	TOTAL X P(ASAL SEKOLAH=...)	E(TOTAL * P(ASAL SEKOLAH=...))	GAIN
2	Tahun Lulus	2014	1	0	1.000000	0	0	0	0	0.009174	0	0	0
		2015	0	0	0	0	0	0	0	0	0		
		2016	4	0	1	0	0	0	0	0.036697	0		
		2017	25	0	1	0	0	0	0	0.229358	0		
		2018	79	0	1	0	0	0	0	0.724771	0		
3	Pendidikan Orang Tua	Dasar	69	0	1	0	0	0	0	0.633028	0	0	0
		Menengah	22	0	1	0	0	0	0	0.201835	0		
		Tinggi	18	0	1	0	0	0	0	0.165138	0		
4	Pendapatan Orang Tua	Rendah	75	0	1	0	0	0	0	0.688073	0	0	0
		Sedang	21	0	1	0	0	0	0	0.192661	0		
		Tinggi	3	0	1	0	0	0	0	0.027523	0		
		Sangat Tinggi	10	0	1	0	0	0	0	0.091743	0		
5	Nilai Ujian Nasional	Baik	64	0	1	0	0	0	0	0.587156	0	0	0
		Sangat Baik	45	0	1	0	0	0	0	0.412844	0		
6	Nilai Tes	Cukup	7	0	1	0	0	0	0	0.06422	0	0	0
		Baik	20	0	1	0	0	0	0	0.183486	0		
		Sangat Baik	82	0	1	0	0	0	0	0.752294	0		
7	Nilai Wawancara	Kurang	0	0	0	0	0	0	0	0	0	0	0
		Cukup	0	0	0	0	0	0	0	0	0		
		Baik	13	0	1	0	0	0	0	0.119266	0		
		Sangat Baik	96	0	1	0	0	0	0	0.880734	0		
8	Nilai BTHQ	Lancar,Fasih, Tajwid Kurang Bagus	2	0	1	0	0	0	0	0.018349	0	0	0
		Lancar,Fasih, Tajwid Bagus	107	0	1	0	0	0	0	0.981651	0		

C. Hasil Pengujian Precision, Recall dan Akurasi

Pengujian untuk algoritma Naive Bayes dan C.45 dilakukan dengan menggunakan data testing. Untuk data ini dilakukan pengujian pada 10 data yang berbeda dengan data training, seperti yang terlihat pada tabel 15.

Tabel 15. Data Testing

NO	Nama Lengkap	Tahun Lulus	Pendidikan Orang Tua	Penghasilan Orang Tua	Nilai Akhir	Nilai Tes	Nilai Wawancara	Nilai BTHQ	KET
						Umum, Agama, Bahasa			
1	MUH SULAIMAN	2018	DASAR	SEDANG	SANGAT BAIK	CUKUP	BAIK	TIDAK LANCAR	TIDAK LULUS
2	KARTINI	2018	DASAR	SEDANG	SANGAT BAIK	CUKUP	BAIK	TIDAK LANCAR	TIDAK LULUS
3	DENI JLSAIN	2018	MENENGAH	SEDANG	BAIK	CUKUP	BAIK	TIDAK LANCAR	TIDAK LULUS
4	M, ABI SARWAN	2018	DASAR	SEDANG	BAIK	CUKUP	SANGAT BAIK	TIDAK LANCAR	TIDAK LULUS
5	AMRULLAH SIRIH	2018	TINGGI	RENDAH	BAIK	BAIK	SANGAT BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
6	AAN HALIM ARAS	2018	DASAR	RENDAH	BAIK	SANGAT BAIK	SANGAT BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
7	MUSYAKIRAH	2017	MENENGAH	RENDAH	BAIK	SANGAT BAIK	BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
8	MUH TAUFUQL RAFLI	2018	MENENGAH	RENDAH	BAIK	SANGAT BAIK	SANGAT BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
9	DEVI SEPTISARI	2018	TINGGI	SANGAT TINGGI	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
10	ASNIA NANDA SUKMAN	2017	DASAR	SANGAT TINGGI	BAIK	BAIK	BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS

1. Algoritma Naive Bayes

Pada algoritma Naive Bayes, pengujian dilakukan dengan menggunakan data testing yang ada pada tabel 15. Prosesnya dengan membuat tabel prediksi dengan menghitung persentase nilai lulus dan tidak lulus, yang dapat dilihat pada tabel di bawah ini.

Tabel 16. Tabel Prediksi

CLASS PREDICTION	LULUS	TIDAK LULUS
TIDAK LULUS	0.000%	0.000%
TIDAK LULUS	0.000%	0.000%
TIDAK LULUS	0.000%	0.000%
TIDAK LULUS	0.000%	0.000%
LULUS	0.492%	0.275%
LULUS	7.730%	4.326%
LULUS	0.106%	0.059%
LULUS	2.465%	1.379%
LULUS	0.004%	0.002%
LULUS	0.000%	0.000%

Dari tabel 16 terlihat hasil tabel data testing dan hasil tabel prediksi memiliki nilai yang sama. Oleh karena itu, kemudian dibuat lagi tabel perbandingan dan tabel persamaan untuk menghitung precision, recall dan accuracy dari proses perhitungan algoritma Naive Bayes.

Tabel 17. Tabel Perbandingan

PREDICTED	CLASS	
	LULUS	TIDAK LULUS
LULUS	6	0
TIDAK LULUS	0	4

Tabel 18. Tabel Persamaan Precision, Recall dan Akurasi

		Nilai Sebenarnya	
		6 (True Positive)	0 (False Positive)
Nilai Prediksi	TRUE	6 (True Positive)	0 (False Positive)
	FALSE	0 (False Negative)	4 (True Negative)

- Nilai Precision

$$\begin{aligned}
 \text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}} \\
 &= 6 / (6+0) \\
 &= 6 / 6 \\
 &= 100\%
 \end{aligned}$$

- Nilai Recall

$$\begin{aligned}
 \text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}} \\
 &= 6 / (6+0) \\
 &= 6 / 6 \\
 &= 100\%
 \end{aligned}$$

- Nilai Accuracy

$$\begin{aligned}
 \text{Accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \\
 &= (6+4)/(6+0+4+0) \\
 &= 10/10 \\
 &= 100\%
 \end{aligned}$$

Hasil perhitungan di atas menunjukkan nilai precision sebesar 100%, nilai recall sebesar 100% dan nilai accuracy sebesar 100%

2. Algoritma C.45

Seperti pada pengujian algoritma Naive Bayes, pengujian juga dilakukan dengan menggunakan data testing pada tabel 15, dengan melihat hasil perhitungan entropy

dan gain pada data training serta hasil percabangan atau pohon keputusan yang telah dibuat. Proses pengujian juga dilakukan dengan menghitung nilai precision, recall dan accuracy.

- Nilai Precision

$$\begin{aligned}\text{Precision} &= \frac{\text{TP}}{\text{TP} + \text{FP}} \\ &= 6 / (6+0) \\ &= 6 / 6 \\ &= 100\%\end{aligned}$$

- Nilai Recall

$$\begin{aligned}\text{Recall} &= \frac{\text{TP}}{\text{TP} + \text{FN}} \\ &= 6 / (6+0) \\ &= 6 / 6 \\ &= 100\%\end{aligned}$$

- Nilai Accuracy

$$\begin{aligned}\text{Accuracy} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \\ &= (6+4)/(6+0+4+0) \\ &= 10/10 \\ &= 100\%\end{aligned}$$

Dari perhitungan di atas nilai precision sebesar 100%, nilai recall 100% dan nilai accuracy sebesar 100%

D. Perbandingan Algoritma

Hasil pengujian yang telah dilakukan, terlihat nilai *precision*, *recall* dan *accuracy* menghasilkan data yang sama dari kedua algoritma. Hal ini menunjukkan bahwa implementasi algoritma Naive Bayes dan C.45 sudah benar dan valid dengan melihat hasil prediksi dari kedua algoritma tersebut.

E. Pola Sekuensial Yang Terbentuk

Berdasarkan pengujian yang telah dilakukan, dibentuk pola berdasarkan urutan kejadian dari atribut kemudian dilakukan pengujian. Atribut yang digunakan yaitu Nilai Ujian Nasional, Nilai Tes, Nilai Wawancara dan Nilai BTHQ. Dari hasil pengujian akan terbentuk pola baru yang menjadi tolak ukur bagi murid yang akan mengikuti seleksi pada IAIN Bone. Pola sekuensial yang telah terbentuk terdapat 24 pola untuk kategori LULUS dan 48 pola kategori TIDAK LULUS yang dapat dilihat pada tabel 20 dibawah ini :

Tabel 20. Tabel Pola Sekuensial Yang Terbentuk

No.	Nilai Akhir	Nilai Tes	Nilai Wawancara	Nilai BTHQ	KET
		Umum, Agama, Bahasa			
1	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
2	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
3	SANGAT BAIK	SANGAT BAIK	BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
4	SANGAT BAIK	SANGAT BAIK	BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
5	SANGAT BAIK	BAIK	SANGAT BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
6	SANGAT BAIK	BAIK	SANGAT BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
7	SANGAT BAIK	BAIK	BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
8	SANGAT BAIK	BAIK	BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
9	SANGAT BAIK	CUKUP	SANGAT BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
10	SANGAT BAIK	CUKUP	SANGAT BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
11	SANGAT BAIK	CUKUP	BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
12	SANGAT BAIK	CUKUP	BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
13	BAIK	SANGAT BAIK	SANGAT BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
14	BAIK	SANGAT	SANGAT	LANCAR, FASIH, TAJWID	LULUS

		BAIK	BAIK	KRNG BAGUS	
15	BAIK	SANGAT BAIK	BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
16	BAIK	SANGAT BAIK	BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
17	BAIK	BAIK	SANGAT BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
18	BAIK	BAIK	SANGAT BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
19	BAIK	BAIK	BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
20	BAIK	BAIK	BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
21	BAIK	CUKUP	SANGAT BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
22	BAIK	CUKUP	SANGAT BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
23	BAIK	CUKUP	BAIK	LANCAR, FASIH, TAJWID BAGUS	LULUS
24	BAIK	CUKUP	BAIK	LANCAR, FASIH, TAJWID KRNG BAGUS	LULUS
25	SANGAT BAIK	SANGAT BAIK	SANGAT BAIK	TIDAK LANCAR	TIDAK LULUS
26	SANGAT BAIK	SANGAT BAIK	BAIK	TIDAK LANCAR	TIDAK LULUS
27	SANGAT BAIK	SANGAT BAIK	CUKUP	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS
28	SANGAT BAIK	SANGAT BAIK	CUKUP	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS
29	SANGAT BAIK	SANGAT BAIK	CUKUP	TIDAK LANCAR	TIDAK LULUS
30	SANGAT BAIK	SANGAT BAIK	KURANG	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS
31	SANGAT BAIK	SANGAT BAIK	KURANG	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS
32	SANGAT BAIK	SANGAT BAIK	KURANG	TIDAK LANCAR	TIDAK LULUS
33	SANGAT BAIK	BAIK	SANGAT BAIK	TIDAK LANCAR	TIDAK LULUS
34	SANGAT BAIK	BAIK	BAIK	TIDAK LANCAR	TIDAK LULUS
35	SANGAT BAIK	BAIK	CUKUP	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS
36	SANGAT BAIK	BAIK	CUKUP	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS

37	SANGAT BAIK	BAIK	CUKUP	TIDAK LANCAR	TIDAK LULUS
38	SANGAT BAIK	BAIK	KURANG	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS
39	SANGAT BAIK	BAIK	KURANG	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS
40	SANGAT BAIK	BAIK	KURANG	TIDAK LANCAR	TIDAK LULUS
41	SANGAT BAIK	CUKUP	SANGAT BAIK	TIDAK LANCAR	TIDAK LULUS
42	SANGAT BAIK	CUKUP	BAIK	TIDAK LANCAR	TIDAK LULUS
43	SANGAT BAIK	CUKUP	CUKUP	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS
44	SANGAT BAIK	CUKUP	CUKUP	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS
45	SANGAT BAIK	CUKUP	CUKUP	TIDAK LANCAR	TIDAK LULUS
46	SANGAT BAIK	CUKUP	KURANG	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS
47	SANGAT BAIK	CUKUP	KURANG	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS
48	SANGAT BAIK	CUKUP	KURANG	TIDAK LANCAR	TIDAK LULUS
49	BAIK	SANGAT BAIK	SANGAT BAIK	TIDAK LANCAR	TIDAK LULUS
50	BAIK	SANGAT BAIK	BAIK	TIDAK LANCAR	TIDAK LULUS
51	BAIK	SANGAT BAIK	CUKUP	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS
52	BAIK	SANGAT BAIK	CUKUP	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS
53	BAIK	SANGAT BAIK	CUKUP	TIDAK LANCAR	TIDAK LULUS
54	BAIK	SANGAT BAIK	KURANG	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS
55	BAIK	SANGAT BAIK	KURANG	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS
56	BAIK	SANGAT BAIK	KURANG	TIDAK LANCAR	TIDAK LULUS
57	BAIK	BAIK	SANGAT BAIK	TIDAK LANCAR	TIDAK LULUS
58	BAIK	BAIK	BAIK	TIDAK LANCAR	TIDAK LULUS
59	BAIK	BAIK	CUKUP	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS

60	BAIK	BAIK	CUKUP	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS
61	BAIK	BAIK	CUKUP	TIDAK LANCAR	TIDAK LULUS
62	BAIK	BAIK	KURANG	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS
63	BAIK	BAIK	KURANG	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS
64	BAIK	BAIK	KURANG	TIDAK LANCAR	TIDAK LULUS
65	BAIK	CUKUP	SANGAT BAIK	TIDAK LANCAR	TIDAK LULUS
66	BAIK	CUKUP	BAIK	TIDAK LANCAR	TIDAK LULUS
67	BAIK	CUKUP	CUKUP	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS
68	BAIK	CUKUP	CUKUP	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS
69	BAIK	CUKUP	CUKUP	TIDAK LANCAR	TIDAK LULUS
70	BAIK	CUKUP	KURANG	LANCAR, FASIH, TAJWID BAGUS	TIDAK LULUS
71	BAIK	CUKUP	KURANG	LANCAR, FASIH, TAJWID KRNG BAGUS	TIDAK LULUS
72	BAIK	CUKUP	KURANG	TIDAK LANCAR	TIDAK LULUS

Tabel sekuensial di atas, menghasilkan 72 pola baru dengan tabel prediksi dapat dilihat pada tabel di bawah ini :

Tabel 21. Tabel Prediksi Pola Sekuensial

CLASS PREDICTION	LULUS	TIDAK LULUS
LULUS	17.217%	9.635%
LULUS	0.322%	0.180%
LULUS	2.331%	1.305%
LULUS	0.044%	0.024%
LULUS	4.199%	2.350%
LULUS	0.078%	0.044%
LULUS	0.569%	0.318%

TIDAK LULUS	0.000%	0.000%
TIDAK LULUS	0.000%	0.000%
TIDAK LULUS	0.000%	0.000%
TIDAK LULUS	0.000%	0.000%
TIDAK LULUS	0.000%	0.000%

F. Implementasi Algoritma Dalam Aplikasi

Pada tahap ini, kedua algoritma Naive Bayes dan C.45 di implementasikan kedalam sebuah aplikasi dengan menggunakan PHP MySql.

1. Form Login

Form login terdiri dari username dan password beserta tombol login yang digunakan untuk mengakses aplikasi yang telah dibuat. Dapat dilihat pada gambar 7 di bawah ini :



Gambar 7. Form Login

2. Halaman Depan

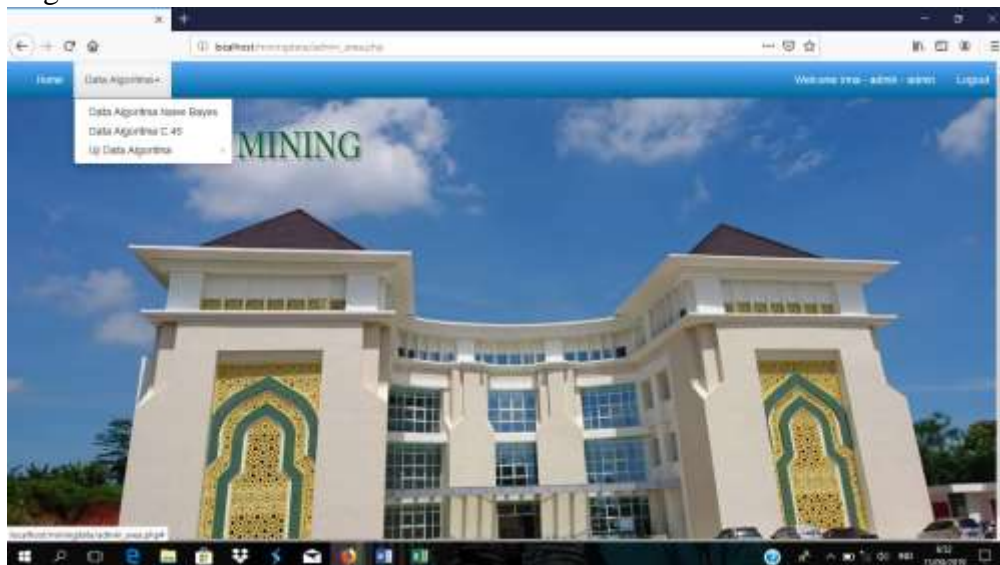
Tampilan halaman depan terdiri dari menu home, data algoritma dan logout. Dapat dilihat pada gambar 8 di bawah ini.



Gambar 8. Tampilan Halaman Depan

3. Tampilan Algoritma Naive Bayes

Halaman ini menampilkan hasil dari proses algoritma Naive Bayes, dapat dilihat pada gambar 9 dan 10 di bawah ini.



Gambar 9. Menu Algoritma Naive Bayes

Tahun	Jumlah Lulus	All Data Lulus	Ruang Lulus	Nilai Lulus	Jumlah Tidak Lulus	All Tidak Lulus	Ruang Tidak Lulus	Nilai Tidak Lulus
2014	1	100	$1/100 = 0.0101$	1%	0	01	$0/01 = 0$	0%
2015	0	100	$0/100 = 0$	0%	2	01	$2/01 = 0.0202$	2%
2016	4	100	$4/100 = 0.0404$	4%	0	01	$0/01 = 0.0101$	1%
2017	25	100	$25/100 = 0.2505$	25%	0	01	$0/01 = 0.0101$	1%
2018	70	100	$70/100 = 0.7007$	70%	0	01	$0/01 = 0.0101$	1%

Penelitian Orang Tua	Jumlah Lulus	All Data Lulus	Ruang Lulus	Nilai Lulus	Jumlah Tidak Lulus	All Tidak Lulus	Ruang Tidak Lulus	Nilai Tidak Lulus
DASAR	60	100	$60/100 = 0.6002$	60%	33	01	$33/01 = 0.5403$	54%
MENENGGAH	22	100	$22/100 = 0.2203$	22%	12	01	$12/01 = 0.0902$	9%
TINGGI	18	100	$18/100 = 0.1804$	18%	10	01	$10/01 = 0.2503$	25%

Pengkabaran Orang Tua	Jumlah Lulus	All Data Lulus	Ruang Lulus	Nilai Lulus	Jumlah Tidak Lulus	All Tidak Lulus	Ruang Tidak Lulus	Nilai Tidak Lulus
RENDAH	75	100	$75/100 = 0.6607$	66%	41	01	$41/01 = 0.6703$	67%
SANGAT RENDAH	10	100	$10/100 = 0.0917$	9%	1	01	$1/01 = 0.0108$	1%
SEDANG	21	100	$21/100 = 0.1928$	19%	0	01	$0/01 = 0.0109$	1%

Gambar 10. Tampilan Proses Algoritma Naive Bayes

4. Tampilan Algoritma C.45

Halaman ini menampilkan menu algoritma C.45 dan hasil dari proses algoritma C.45, dapat dilihat pada gambar 11 dan 12 di bawah ini.



Gambar 11. Menu Algoritma C.45

Tahun	Jumlah Lulus	Jumlah Tidak Lulus	P/Total Lulus	P/Total Tidak Lulus	P/Total Lulus + Tidak Lulus	Total	P/Total Lulus + Tidak Lulus	TOTAL X P/Total Lulus + Tidak Lulus	TOTAL X P/Total Lulus + Tidak Lulus	
2014	5	0	1	0	0	0	0	0.000000000000	0	
2015	0	2	0	1	0	0	0	0.011347528824	0	
2016	4	7	0.3636363636	0.6363636364	0.5000000000	0.4140578121	0.04906034691	0.3647056022520	0.901189744103	
2017	20	8	0.8046875000	0.1953125000	0.2500000000	0.482891742719	0.188234873332	0.16232541176	0.4262828843	
2018	79	46	0.632	0.368	0.4900000000	0.5307378887	0.04813461889	0.75824171647	0.68788022342	
									0.88533	0.88533

Provinsi	Jumlah Lulus	Jumlah Tidak Lulus	P/Total Lulus	P/Total Tidak Lulus	P/Total Lulus + Tidak Lulus	Total	P/Total Lulus + Tidak Lulus	TOTAL X P/Total Lulus + Tidak Lulus	TOTAL X P/Total Lulus + Tidak Lulus	
DAKAR	80	23	0.7747088235	0.2252911765	0.28446232615	0.520115003187	0.0001700472	0.0	0.5448070028	
HECINGIH	22	12	0.6460905263	0.3539094737	0.48027144844	0.53324237534	0.0000181870	0.0	0.10733347878	
TIVIGI	16	18	0.5294117647	0.4705882353	0.48575520937	0.511747218452	0.00756294869	0.2	0.1088688274	
									0.931761	0.88888

Gambar 12. Tampilan Proses Algoritma C.45

5. Form Pengujian

Halaman ini menampilkan form untuk data testing atau data uji yang telah dibuat, dengan memasukkan data pada entry data uji, kemudian hasil inputan dan proses uji dapat dilihat pada view data uji. Seperti yang terlihat pada gambar 13 dan 14 di bawah ini.

Gambar 13. Tampilan Form Entry Data Uji

Uji Hasil	No	Nama	Tahun	Agama	Alamat	Umur	Gender	Agama	Agama	Agama	Agama	Agama	Agama	Agama	Agama
Uji Hasil 1	1	REN SUAMBY	2018	ISLAM	DEWANG	SANGAT BENE	CUKUP	BNK	TEKNIK LANCAR	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	0.000000000000
Uji Hasil 2	2	KARTO	2018	ISLAM	DEWANG	SANGAT BENE	CUKUP	BNK	TEKNIK LANCAR	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	0.000000000000
Uji Hasil 3	3	DEWALDAN	2018	ISLAM	DEWANG	BNK	CUKUP	BNK	TEKNIK LANCAR	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	0.000000000000
Uji Hasil 4	4	NUBI SARAH	2018	ISLAM	DEWANG	BNK	CUKUP	SANGAT BENE	TEKNIK LANCAR	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	0.000000000000
Uji Hasil 5	5	MARULLAH SIDI	2018	ISLAM	DEWANG	BNK	BNK	SANGAT BENE	LANCARKAN	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	0.000000000000
Uji Hasil 6	6	MARULLAH SIDI	2018	ISLAM	DEWANG	BNK	SANGAT BENE	SANGAT BENE	LANCARKAN	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	0.000000000000
Uji Hasil 7	7	ALYAMINAH	2017	ISLAM	DEWANG	BNK	SANGAT BENE	BNK	LANCARKAN	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	0.000000000000
Uji Hasil 8	8	SEI TALYUOL RIFLI	2018	ISLAM	DEWANG	BNK	SANGAT BENE	SANGAT BENE	LANCARKAN	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	0.000000000000
Uji Hasil 9	9	DEA SOTTODAN	2018	ISLAM	DEWANG	SANGAT BENE	SANGAT BENE	SANGAT BENE	LANCARKAN	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	TEKNIK LULUS	0.000000000000

Gambar 14. Tampilan Hasil Olah Data Uji

G. Pengujian Blackbox

Pengujian blackbox merupakan pengujian terhadap fitur-fitur serta fungsionalitas yang terdapat pada aplikasi dengan memberikan sejumlah testcase pada aplikasi tersebut. Daftar testcase yang diberikan ke aplikasi untuk pengujian blackbox dapat dilihat pada tabel 19.

Tabel 21. Test Case Pengujian Blackbox

Modul	Test Case	Output Yang Diharapkan	Hasil di Aplikasi
Login	Input username dan password benar	Masuk ke halaman admin	Valid
Manage Data	Memasukkan data dan semua data telah terisi	Data masuk ke dalam database	Valid
Manage Data	Mengubah data	Data di dalam database telah terupdate	Valid
Manage Data	Menghapus 1 buah data	Data di dalam database terhapus	Valid
Manage Data	Memasukkan beberapa buah data sekaligus kedalam database	File excel masuk ke dalam database	Valid

	menggunakan file excel		
Pengujian	Menekan tombol pengujian	Aplikasi menghitung hasil prediksi dari algoritma Naive Bayes	Valid
Pengujian	Menekan tombol pengujian	Aplikasi menghitung hasil prediksi dari algoritma C.45	Valid
Home	Keluar dari proses dan Kembali ke halaman depan	Masuk ke halaman depan	Valid
Logout	Keluar dari aplikasi	Keluar dari halaman admin	Valid

H. Listing Program Untuk Perhitungan Algoritma

1. Algoritma Naive Bayes

```

<?php

//delete
$id = abs((int)$_GET['id']);

$id=$_GET[id];
$aksi=$_GET[aksi];

function anti_sql_injection_2($id) {

$id = stripslashes($id);

$id = strip_tags($id);

$id = mysql_real_escape_string($id);

return $id;

}
$aksi = mysql_real_escape_string($aksi);
if($aksi=="delete"){

```

```

$sqldel="delete from sapto_upps where id=$_GET[id]";
$querydel=mysql_query($sqldel,$conn);

}

//TAMPIL

$sqltampil="select  asalsekolah, sum(IF(ket='LULUS',1,0)), sum(IF(ket='TIDAK
LULUS',1,0))
                                from kelulusan
                                group      by
asalsekolah                                order by id";

$sqlquery=mysql_query($sqltampil,$conn);
$sqlbaris=mysql_fetch_row($sqlquery);

echo "<table class='table table-striped'>";

        //TAMPIL berdasarkan tahun

        $sqltampil="select  tahunlulus, sum(IF(ket='LULUS',1,0)), sum(IF(ket='TIDAK
        LULUS',1,0))
                                from kelulusan
                                group      by
        tahunlulus                                order  by  id
        desc";
        $sqlquery=mysql_query($sqltampil,$conn);
        $sqlbaris=mysql_fetch_row($sqlquery);

        echo "<tr>

                <td colspan=9>&nbsp;  </td>

        </tr>";

        echo "<tr>

                <th>Tahun</th>

                <th>Jum Lulus</th>
                <th>All Data Lulus</th>
                <th>Hitung Lulus</th>
                <th>Nilai Lulus</th>

```

```

                <th>Jum Tidak Lulus</td>
                <th>All Tidak Lulus</td>
                <th>Hitung Tidak Lulus</td>
                <th>Nilai Tidak Lulus</td>

                </tr>";
do{
    list( $AsalSekolah,
          $Lulus,
          $TidakLulus
          )=$sqlbaris;

    $QueJumlahLulusAll=mysql_query("select count(ket) as JumLulus from
kelulusan where ket='LULUS' ",$conn);
    $jl=mysql_fetch_array($QueJumlahLulusAll);

    $QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
from kelulusan where ket='TIDAK LULUS' ",$conn);
    $jln=mysql_fetch_array($QueJumlahNoLulusAll);

    $JumLulus=$jl[JumLulus];
    $JumTidakLulus=$jln[JumNLulus];

    $HitungLulus=round($Lulus/$JumLulus, 5);
    $HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

    $NilaiLulus=round($HitungLulus * 100 );
    $NilaiNoLulus=round($HitungNoLulus * 100 );

    echo "<tr>
                <td>$AsalSekolah</td>
                <td>$Lulus</td>
                <td>$jl[JumLulus]</td>
                <td>$Lulus / $jl[JumLulus] = $HitungLulus</td>
                <td>$NilaiLulus %</td>

                <td>$TidakLulus</td>
                <td>$jln[JumNLulus]</td>
                <td>$TidakLulus / $jln[JumNLulus] = $HitungNoLulus</td>
                <td>$NilaiNoLulus %</td>

                </tr>";
    }while($sqlbaris=mysql_fetch_row($sqlquery));

```



```

        </tr>";

echo "<tr>
        <th>Penghasilan Orang Tua</th>

        <th>Jum Lulus</th>
        <th>All Data Lulus</th>
        <th>Hitung Lulus</th>
        <th>Nilai Lulus</th>

        <th>Jum Tidak Lulus</th>
        <th>All Tidak Lulus</th>
        <th>Hitung Tidak Lulus</th>
        <th>Nilai Tidak Lulus</th>

        </tr>";
do{
    list( $AsalSekolah,
          $Lulus,
          $TidakLulus
          )=$sqlbaris;

    $QueJumlahLulusAll=mysql_query("select count(ket) as JumLulus from
kelulusan where ket='LULUS' ",$conn);
    $jl=mysql_fetch_array($QueJumlahLulusAll);

    $QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
from kelulusan where ket='TIDAK LULUS' ",$conn);
    $jln=mysql_fetch_array($QueJumlahNoLulusAll);

    $JumLulus=$jl[JumLulus];
    $JumTidakLulus=$jln[JumNLulus];

    $HitungLulus=round($Lulus/$JumLulus, 5);
    $HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

    $NilaiLulus=round($HitungLulus * 100 );
    $NilaiNoLulus=round($HitungNoLulus * 100 );

echo "<tr>
        <td>$AsalSekolah</td>
        <td>$Lulus</td>
        <td>$jl[JumLulus]</td>

```



```

                </tr>";
do{
    list( $AsalSekolah,
          $Lulus,
          $TidakLulus
        )=$sqlbaris;

    $QueJumlahLulusAll=mysql_query("select count(ket) as JumLulus from
kelulusan where ket='LULUS' ",$conn);
    $jl=mysql_fetch_array($QueJumlahLulusAll);

    $QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
from kelulusan where ket='TIDAK LULUS' ",$conn);
    $jln=mysql_fetch_array($QueJumlahNoLulusAll);

    $JumLulus=$jl[JumLulus];
    $JumTidakLulus=$jln[JumNLulus];

    $HitungLulus=round($Lulus/$JumLulus, 5);
    $HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

    $NilaiLulus=round($HitungLulus * 100 );
    $NilaiNoLulus=round($HitungNoLulus * 100 );

    echo "<tr>
        <td>$AsalSekolah</td>
        <td>$Lulus</td>
        <td>$jl[JumLulus]</td>
        <td>$Lulus / $jl[JumLulus] = $HitungLulus</td>
        <td>$NilaiLulus %</td>

        <td>$TidakLulus</td>
        <td>$jln[JumNLulus]</td>
        <td>$TidakLulus / $jln[JumNLulus] = $HitungNoLulus</td>
        <td>$NilaiNoLulus %</td>

    </tr>";

}while($sqlbaris=mysql_fetch_row($sqlquery));

```

```

//TAMPIL berdasarkan Nilai Tes

$sqltampil="select  nilaites,  sum(IF(ket='LULUS',1,0)),  sum(IF(ket='TIDAK
LULUS',1,0))

                                from kelulusan
                                group      by

nilaites                                order  by  id

DESC";
$sqlquery=mysql_query($sqltampil,$conn);
$sqlbaris=mysql_fetch_row($sqlquery);

echo "<tr>

                                <td colspan=9>&nbsp;   </td>

                                </tr>";

echo "<tr>

                                <th>Nilai Tes</th>

                                <th>Jum Lulus</th>
                                <th>All Data Lulus</th>
                                <th>Hitung Lulus</th>
                                <th>Nilai Lulus</th>

                                <th>Jum Tidak Lulus</th>
                                <th>All Tidak Lulus</th>
                                <th>Hitung Tidak Lulus</th>
                                <th>Nilai Tidak Lulus</th>

                                </tr>";
do{
    list( $AsalSekolah,
          $Lulus,
          $TidakLulus
          )=$sqlbaris;

    $QueJumlahLulusAll=mysql_query("select count(ket) as JumLulus from
kelulusan where ket='LULUS' ",$conn);
    $jl=mysql_fetch_array($QueJumlahLulusAll);

```

```

$QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
from kelulusan where ket='TIDAK LULUS' ",$conn);
$jln=mysql_fetch_array($QueJumlahNoLulusAll);

$JumLulus=$jl[JumLulus];
$JumTidakLulus=$jln[JumNLulus];

$HitungLulus=round($Lulus/$JumLulus, 5);
$HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

$NilaiLulus=round($HitungLulus * 100 );
$NilaiNoLulus=round($HitungNoLulus * 100 );

echo "<tr>
        <td>$AsalSekolah</td>
        <td>$Lulus</td>
        <td>$jl[JumLulus]</td>
        <td>$Lulus / $jl[JumLulus] = $HitungLulus</td>
        <td>$NilaiLulus %</td>

        <td>$TidakLulus</td>
        <td>$jln[JumNLulus]</td>
        <td>$TidakLulus / $jln[JumNLulus] = $HitungNoLulus</td>
        <td>$NilaiNoLulus %</td>

</tr>";

}while($sqlbaris=mysql_fetch_row($sqlquery));

//TAMPIL berdasarkan Nilai Wawancara

$sqltampil="select          nilaiwawancara,          sum(IF(ket='LULUS',1,0)),
sum(IF(ket='TIDAK LULUS',1,0))
                                from kelulusan
                                group      by
nilaiwawancara
                                order  by  id
DESC";
$sqlquery=mysql_query($sqltampil,$conn);
$sqlbaris=mysql_fetch_row($sqlquery);

echo "<tr>

```



```

        </tr>";
do{
    list( $AsalSekolah,
          $Lulus,
          $TidakLulus
        )=$sqlbaris;

    $QueJumlahLulusAll=mysql_query("select count(ket) as JumLulus from
kelulusan where ket='LULUS' ",$conn);
    $jl=mysql_fetch_array($QueJumlahLulusAll);

    $QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
from kelulusan where ket='TIDAK LULUS' ",$conn);
    $jln=mysql_fetch_array($QueJumlahNoLulusAll);

    $JumLulus=$jl[JumLulus];
    $JumTidakLulus=$jln[JumNLulus];

    $HitungLulus=round($Lulus/$JumLulus, 5);
    $HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

    $NilaiLulus=round($HitungLulus * 100 );
    $NilaiNoLulus=round($HitungNoLulus * 100 );

    echo "<tr>
        <td>$AsalSekolah</td>
        <td>$Lulus</td>
        <td>$jl[JumLulus]</td>
        <td>$Lulus / $jl[JumLulus] = $HitungLulus</td>
        <td>$NilaiLulus %</td>

        <td>$TidakLulus</td>
        <td>$jln[JumNLulus]</td>
        <td>$TidakLulus / $jln[JumNLulus] = $HitungNoLulus</td>
        <td>$NilaiNoLulus %</td>

    </tr>";

}while($sqlbaris=mysql_fetch_row($sqlquery));

echo "</table>";

?>

```

2. Algoritma C.45

```
<?php

//delete
$Sid = abs((int)$_GET['id']);

$Sid=$_GET[id];
$aksi=$_GET[aksi];

function anti_sql_injection_2($Sid) {

$Sid = stripslashes($Sid);

$Sid = strip_tags($Sid);

$Sid = mysql_real_escape_string($Sid);

return $Sid;
}
$aksi = mysql_real_escape_string($aksi);
if($aksi=="delete"){

$sqldel="delete from sapto_upps where id=$_GET[id]";
$querydel=mysql_query($sqldel,$conn);
}
//TAMPIL asal sekolah

$sqltampil="select  asalsekolah, sum(IF(ket='LULUS',1,0)), sum(IF(ket='TIDAK
LULUS',1,0))
                                from kelulusan
                                group      by
asalsekolah
                                order by id";

$sqlquery=mysql_query($sqltampil,$conn);
$sqlbaris=mysql_fetch_row($sqlquery);
echo "<table class='table table-striped'>";

//TAMPIL berdasarkan tahun
$GTotTotalxPAsalSekolah=0;
$sqltampil="select  tahunlulus, sum(IF(ket='LULUS',1,0)), sum(IF(ket='TIDAK
LULUS',1,0))
                                from kelulusan
```



```

$minuslog2pjNLulus=(-$log2pjNLulus) * $pjNLulus;

$Entropy=$minuslog2pjLulus+$minuslog2pjNLulus;

do{
    list( $AsalSekolah,
          $Lulus,
          $TidakLulus
        )=$sqlbaris;

    $QueJumlahLulusAll=mysql_query("select count(ket) as JumLulus from
kelulusan where ket='LULUS' ",$conn);
    $jl=mysql_fetch_array($QueJumlahLulusAll);

    $QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
from kelulusan where ket='TIDAK LULUS' ",$conn);
    $jln=mysql_fetch_array($QueJumlahNoLulusAll);

    $JumLulus=$jl[JumLulus];
    $JumTidakLulus=$jln[JumNLulus];

    $HitungLulus=round($Lulus/$JumLulus, 5);
    $HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

    $NilaiLulus=round($HitungLulus * 100 );
    $NilaiNoLulus=round($HitungNoLulus * 100 );

    $PLulus=$Lulus/($Lulus+$TidakLulus);
    $PNLulus=$TidakLulus/($Lulus+$TidakLulus);

    $PlogLulus=(-$PLulus) * log($PLulus,2);
    if($PlogLulus=="NAN" or $PlogLulus==0){
        $PlogLulus=0;
    }
    $PlogNLulus=(-$PNLulus) * log($PNLulus,2);
    if($PlogNLulus=="NAN" or $PlogNLulus==0){
        $PlogNLulus=0;
    }

    $Total=$PlogLulus+$PlogNLulus;

    $QueJumSekolah=mysql_query("select count(tahunlulus) as JumSekolah
from kelulusan where tahunlulus='$AsalSekolah' ",$conn);

```

```

$JumSekolah=mysql_fetch_array($QueJumSekolah);

$QueJumSekolah2=mysql_query("select count(tahunlulus) as JumSekolah2
from kelulusan ",$conn);
$JumSekolah2=mysql_fetch_array($QueJumSekolah2);
$PAsalSekolah=$JumSekolah[JumSekolah]/$JumSekolah2[JumSekolah2];
$TotalxPAsalSekolah=$Total*$PAsalSekolah;

$GTotTotalxPAsalSekolah=$GTotTotalxPAsalSekolah+$TotalxPAsalSekol
ah;

$Gain=$Entropy - $GTotTotalxPAsalSekolah;

$GTotTotalxPAsalSekolah=round($GTotTotalxPAsalSekolah,6);
$Gain=round($Gain,5);

echo "<tr>
        <td>$AsalSekolah</td>
        <td>$Lulus</td>
        <td>$TidakLulus</td>
        <td>$PLulus</td>
        <td>$PNLulus</td>

        <td>$PlogLulus</td>
        <td>$PlogNLulus</td>
        <td>$Total</td>
        <td>$PAsalSekolah</td>
        <td>$TotalxPAsalSekolah</td>
        <td></td>
        <td></td>

    </tr>";

}while($sqlbaris=mysql_fetch_row($sqlquery));

echo "<tr>
        <td colspan=10>&nbsp;</td>
        <td>$GTotTotalxPAsalSekolah</td>
        <td><b><font
color=red>$Gain</font></b></td>
    </tr>
    ";

```



```

$TotAllLulusNLulus=$TotLulus+$TotNLulus;

$pjLulus=$TotLulus/$TotAllLulusNLulus;
$log2pjLulus=log($pjLulus,2);
$minuslog2pjLulus=(-$log2pjLulus)*$pjLulus;

$pjNLulus=$TotNLulus/$TotAllLulusNLulus;
$log2pjNLulus=log($pjNLulus,2);
$minuslog2pjNLulus=(-$log2pjNLulus) * $pjNLulus;
$Entropy=$minuslog2pjLulus+$minuslog2pjNLulus;

do{
    list( $AsalSekolah,
          $Lulus,
          $TidakLulus
        )=$sqlbaris;

    $QueJumlahLulusAll=mysql_query("select count(ket) as JumLulus from
kelulusan where ket='LULUS' ",$conn);
    $jl=mysql_fetch_array($QueJumlahLulusAll);

    $QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
from kelulusan where ket='TIDAK LULUS' ",$conn);
    $jln=mysql_fetch_array($QueJumlahNoLulusAll);

    $JumLulus=$jl[JumLulus];
    $JumTidakLulus=$jln[JumNLulus];

    $HitungLulus=round($Lulus/$JumLulus, 5);
    $HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

    $NilaiLulus=round($HitungLulus * 100 );
    $NilaiNoLulus=round($HitungNoLulus * 100 );

    $PLulus=$Lulus/($Lulus+$TidakLulus);
    $PNLulus=$TidakLulus/($Lulus+$TidakLulus);

    $PlogLulus=(-$PLulus) * log($PLulus,2);
    if($PlogLulus=="NAN" or $PlogLulus==0){
        $PlogLulus=0;
    }
    $PlogNLulus=(-$PNLulus) * log($PNLulus,2);

```

```

        if($PlogNLulus=="NAN" or $PlogNLulus==-0){
            $PlogNLulus=0;
        }

    $Total=$PlogLulus+$PlogNLulus;

    $QueJumSekolah=mysql_query("select count(pendortu) as JumSekolah
from kelulusan where pendortu='$AsalSekolah' ",$conn);
    $JumSekolah=mysql_fetch_array($QueJumSekolah);

    $QueJumSekolah2=mysql_query("select count(pendortu) as JumSekolah2
from kelulusan ",$conn);
    $JumSekolah2=mysql_fetch_array($QueJumSekolah2);
    $PAsalSekolah=$JumSekolah[JumSekolah]/$JumSekolah2[JumSekolah2];
    $TotalxPAsalSekolah=$Total*$PAsalSekolah;

    $GTotTotalxPAsalSekolah=$GTotTotalxPAsalSekolah+$TotalxPAsalSekol
ah;

    $Gain=$Entropy - $GTotTotalxPAsalSekolah;

    $GTotTotalxPAsalSekolah=round($GTotTotalxPAsalSekolah,6);
    $Gain=round($Gain,5);

    echo "<tr>
        <td>$AsalSekolah</td>
        <td>$Lulus</td>
        <td>$TidakLulus</td>
        <td>$PLulus</td>
        <td>$PNLulus</td>

        <td>$PlogLulus</td>
        <td>$PlogNLulus</td>
        <td>$Total</td>
        <td>$PAsalSekolah</td>
        <td>$TotalxPAsalSekolah</td>
        <td></td>
        <td></td>

    </tr>";

}while($sqlbaris=mysql_fetch_row($sqlquery));

```

```

echo "<tr>
                                <td colspan=10>&nbsp;</td>
                                <td>$GTotTotalxPAsalSekolah</td>
                                <td><b><font
color=red>$Gain</font></b></td>
                                </tr>
                                ";

//TAMPIL berdasarkan Penghasilan OrangTua
$GTotTotalxPAsalSekolah=0;
$sqltampil="select  penghortu,  sum(IF(ket='LULUS',1,0)),  sum(IF(ket='TIDAK
LULUS',1,0))
                                from kelulusan
                                group      by
penghortu
                                order      by
penghortu";
$sqlquery=mysql_query($sqltampil,$conn);
$sqlbaris=mysql_fetch_row($sqlquery);

                                echo "<tr>
                                <td colspan=9>&nbsp;</td>
                                </tr>";
echo "<tr>
                                <th>Penghasilan OrangTua</td>
                                <th>Jum Lulus</td>
                                <th>Jum Tidak Lulus</td>
                                <th>P(Lulus)</td>
                                <th>P(Tidak Lulus)</td>
                                <th>P(LULUS)*LOG2 P(LULUS)</td>
                                <th>P(TIDAK LULUS)*LOG2 P(TIDAK LULUS)</td>
                                <th>Total</td>
                                <td>P(ASAL SEKOLAH=...)</td>
                                <th>TOTAL X P(ASAL SEKOLAH=...)</td>
                                <th>E(TOTAL*P(ASAL SEKOLAH=..))</td>
                                <th bgcolor=red>GAIN</td>
                                </tr>";

```

```

//entropy H(t)
$queTotLulus=mysql_query("select PLulus,PTidakLulus from asalsekolah
",$conn);
$TL=mysql_fetch_array($queTotLulus);

$TotLulus=$TL[PLulus];
$TotNLulus=$TL[PTidakLulus];
$TotAllLulusNLulus=$TotLulus+$TotNLulus;

$pjLulus=$TotLulus/$TotAllLulusNLulus;
$log2pjLulus=log($pjLulus,2);
$minuslog2pjLulus=(-$log2pjLulus)*$pjLulus;

$pjNLulus=$TotNLulus/$TotAllLulusNLulus;
$log2pjNLulus=log($pjNLulus,2);
$minuslog2pjNLulus=(-$log2pjNLulus) * $pjNLulus;

$Entropy=$minuslog2pjLulus+$minuslog2pjNLulus;
do{
list( $AsalSekolah,
      $Lulus,
      $TidakLulus
      )=$sqlbaris;
$QueJumlahLulusAll=mysql_query("select count(ket) as
JumLulus from kelulusan where ket='LULUS' ",$conn);
$j1=mysql_fetch_array($QueJumlahLulusAll);

$QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
from kelulusan where ket='TIDAK LULUS' ",$conn);
$jln=mysql_fetch_array($QueJumlahNoLulusAll);

$JumLulus=$j1[JumLulus];
$JumTidakLulus=$jln[JumNLulus];
$HitungLulus=round($Lulus/$JumLulus, 5);
$HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

$NilaiLulus=round($HitungLulus * 100 );
$NilaiNoLulus=round($HitungNoLulus * 100 );

$PLulus=$Lulus/($Lulus+$TidakLulus);
$PNLulus=$TidakLulus/($Lulus+$TidakLulus);

```

```

$PlogLulus=(-$PLulus) * log($PLulus,2);
    if($PlogLulus=="NAN" or $PlogLulus==-0){
        $PlogLulus=0;
    }
$PlogNLulus=(-$PNLulus) * log($PNLulus,2);
    if($PlogNLulus=="NAN" or $PlogNLulus==-0){
        $PlogNLulus=0;
    }

$Total=$PlogLulus+$PlogNLulus;

$QueJumSekolah=mysql_query("select count(pendortu) as JumSekolah
from kelulusan where penghortu='$AsalSekolah' ",$conn);
$JumSekolah=mysql_fetch_array($QueJumSekolah);

$QueJumSekolah2=mysql_query("select count(pendortu) as JumSekolah2
from kelulusan ",$conn);
$JumSekolah2=mysql_fetch_array($QueJumSekolah2);
$PAsalSekolah=$JumSekolah[JumSekolah]/$JumSekolah2[JumSekolah2];
$TotalxPAsalSekolah=$Total*$PAsalSekolah;

$GTotTotalxPAsalSekolah=$GTotTotalxPAsalSekolah+$TotalxPAsalSekol
ah;

$Gain=$Entropy - $GTotTotalxPAsalSekolah;

$GTotTotalxPAsalSekolah=round($GTotTotalxPAsalSekolah,6);
$Gain=round($Gain,5);

echo "<tr>
    <td>$AsalSekolah</td>
    <td>$Lulus</td>
    <td>$TidakLulus</td>
    <td>$PLulus</td>
    <td>$PNLulus</td>

    <td>$PlogLulus</td>
    <td>$PlogNLulus</td>
    <td>$Total</td>
    <td>$PAsalSekolah</td>
    <td>$TotalxPAsalSekolah</td>
    <td></td>
    <td></td>

```

```

        </tr>";

    }while($sqlbaris=mysql_fetch_row($sqlquery));

    echo "<tr>
        <td colspan=10>&nbsp;</td>
        <td>$GTotTotalxPAsalSekolah</td>
        <td><b><font
color=red>$Gain</font></b></td>
        </tr>
        ";

//TAMPIL berdasarkan Nilai Ujian Nasional
$GTotTotalxPAsalSekolah=0;
$sqltampil="select nilaiuan, sum(IF(ket='LULUS',1,0)), sum(IF(ket='TIDAK
LULUS',1,0))
                                from kelulusan
                                group by
                                nilaiuan
                                order by
                                nilaiuan";
$sqlquery=mysql_query($sqltampil,$conn);
$sqlbaris=mysql_fetch_row($sqlquery);

    echo "<tr>
        <td colspan=9>&nbsp;</td>
        </tr>";

    echo "<tr>
        <th>Nilai Ujian Nasional</td>
        <th>Jum Lulus</td>
        <th>Jum Tidak Lulus</td>
        <th>P(Lulus)</td>
        <th>P(Tidak Lulus)</td>
        <th>P(LULUS)*LOG2 P(LULUS)</td>
        <th>P(TIDAK LULUS)*LOG2 P(TIDAK LULUS)</td>
        <th>Total</td>
        <td>P(ASAL SEKOLAH=...)</td>
        <th>TOTAL X P(ASAL SEKOLAH=....)</td>

```

```

                <th>E(TOTAL*P(ASAL SEKOLAH=..))</td>
                <th bgcolor=red>GAIN</td>

        </tr>";

        //entropy H(t)
        $queTotLulus=mysql_query("select PLulus,PTidakLulus from asalsekolah
        ",$conn);
        $TL=mysql_fetch_array($queTotLulus);
        $TotLulus=$TL[PLulus];
        $TotNLulus=$TL[PTidakLulus];
        $TotAllLulusNLulus=$TotLulus+$TotNLulus;

        $pjLulus=$TotLulus/$TotAllLulusNLulus;
        $log2pjLulus=log($pjLulus,2);
        $minuslog2pjLulus=(-$log2pjLulus)*$pjLulus;

        $pjNLulus=$TotNLulus/$TotAllLulusNLulus;
        $log2pjNLulus=log($pjNLulus,2);
        $minuslog2pjNLulus=(-$log2pjNLulus) * $pjNLulus;

        $Entropy=$minuslog2pjLulus+$minuslog2pjNLulus;

do{
        list( $AsalSekolah,
                $Lulus,
                $TidakLulus
                )=$sqlbaris;

        $QueJumlahLulusAll=mysql_query("select count(ket) as JumLulus from
        kelulusan where ket='LULUS' ",$conn);
        $jl=mysql_fetch_array($QueJumlahLulusAll);

        $QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
        from kelulusan where ket='TIDAK LULUS' ",$conn);
        $jln=mysql_fetch_array($QueJumlahNoLulusAll);

        $JumLulus=$jl[JumLulus];
        $JumTidakLulus=$jln[JumNLulus];

        $HitungLulus=round($Lulus/$JumLulus, 5);
        $HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

```

```

$NilaiLulus=round($HitungLulus * 100 );
$NilaiNoLulus=round($HitungNoLulus * 100 );

$PLulus=$Lulus/($Lulus+$TidakLulus);
$PNLulus=$TidakLulus/($Lulus+$TidakLulus);

$PlogLulus=(-$PLulus) * log($PLulus,2);
    if($PlogLulus=="NAN" or $PlogLulus==-0){
        $PlogLulus=0;
    }
$PlogNLulus=(-$PNLulus) * log($PNLulus,2);
    if($PlogNLulus=="NAN" or $PlogNLulus==-0){
        $PlogNLulus=0;
    }

$Total=$PlogLulus+$PlogNLulus;

$QueJumSekolah=mysql_query("select count(pendortu) as JumSekolah
from kelulusan where nilaiuan='$AsalSekolah' ",$conn);
$JumSekolah=mysql_fetch_array($QueJumSekolah);

$QueJumSekolah2=mysql_query("select count(pendortu) as JumSekolah2
from kelulusan ",$conn);
$JumSekolah2=mysql_fetch_array($QueJumSekolah2);
$PAsalSekolah=$JumSekolah[$JumSekolah]/$JumSekolah2[$JumSekolah2];
$TotalxPAsalSekolah=$Total*$PAsalSekolah;

$GTotTotalxPAsalSekolah=$GTotTotalxPAsalSekolah+$TotalxPAsalSekol
ah;

$Gain=$Entropy - $GTotTotalxPAsalSekolah;

$GTotTotalxPAsalSekolah=round($GTotTotalxPAsalSekolah,6);
$Gain=round($Gain,5);

echo "<tr>
    <td>$AsalSekolah</td>
    <td>$Lulus</td>
    <td>$TidakLulus</td>
    <td>$PLulus</td>
    <td>$PNLulus</td>

    <td>$PlogLulus</td>

```

```

        <td>$PlogNLulus</td>
        <td>$Total</td>
        <td>$PAsalSekolah</td>
        <td>$TotalxPAsalSekolah</td>
        <td></td>
        <td></td>
    </tr>";
    }while($sqlbaris=mysql_fetch_row($sqlquery));

    echo "<tr>
        <td colspan=10>&nbsp;</td>
        <td>$GTotTotalxPAsalSekolah</td>
        <td><b><font
color=red>$Gain</font></b></td>
    </tr>
    ";
    //TAMPIL berdasarkan Nilai Tes
    $GTotTotalxPAsalSekolah=0;
    $sqltampil="select  nilaites,  sum(IF(ket='LULUS',1,0)),  sum(IF(ket='TIDAK
LULUS',1,0))
                                from kelulusan
                                group      by
                                order  by  id
DESC";
    $sqlquery=mysql_query($sqltampil,$conn);
    $sqlbaris=mysql_fetch_row($sqlquery);

    echo "<tr>
        <td colspan=9>&nbsp;</td>
    </tr>";

    echo "<tr>
        <th>Nilai Tes</td>
        <th>Jum Lulus</td>
        <th>Jum Tidak Lulus</td>
        <th>P(Lulus)</td>
        <th>P(Tidak Lulus)</td>
        <th>P(LULUS)*LOG2 P(LULUS)</td>
    </tr>";

```

```

        <th>P(TIDAK LULUS)*LOG2 P(TIDAK LULUS)</th>
        <th>Total</th>
        <td>P(ASAL SEKOLAH=...)</td>
        <th>TOTAL X P(ASAL SEKOLAH=....)</th>
        <th>E(TOTAL*P(ASAL SEKOLAH=..))</th>
        <th bgcolor=red>GAIN</th>
    </tr>";

    //entropy H(t)
    $queTotLulus=mysql_query("select PLulus,PTidakLulus from asalsekolah
",$conn);
    $TL=mysql_fetch_array($queTotLulus);

    $TotLulus=$TL[PLulus];
    $TotNLulus=$TL[PTidakLulus];
    $TotAllLulusNLulus=$TotLulus+$TotNLulus;
    $pjLulus=$TotLulus/$TotAllLulusNLulus;
    $log2pjLulus=log($pjLulus,2);
    $minuslog2pjLulus=(-$log2pjLulus)*$pjLulus;

    $pjNLulus=$TotNLulus/$TotAllLulusNLulus;
    $log2pjNLulus=log($pjNLulus,2);
    $minuslog2pjNLulus=(-$log2pjNLulus) * $pjNLulus;

    $Entropy=$minuslog2pjLulus+$minuslog2pjNLulus;

do{
    list( $AsalSekolah,
        $Lulus,
        $TidakLulus
        )=$sqlbaris;

    $QueJumlahLulusAll=mysql_query("select count(ket) as JumLulus from
kelulusan where ket='LULUS' ",$conn);
    $jl=mysql_fetch_array($QueJumlahLulusAll);

    $QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
from kelulusan where ket='TIDAK LULUS' ",$conn);
    $jln=mysql_fetch_array($QueJumlahNoLulusAll);

    $JumLulus=$jl[JumLulus];
    $JumTidakLulus=$jln[JumNLulus];

```

```

$HitungLulus=round($Lulus/$JumLulus, 5);
$HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

$NilaiLulus=round($HitungLulus * 100 );
$NilaiNoLulus=round($HitungNoLulus * 100 );
$PLulus=$Lulus/($Lulus+$TidakLulus);
$PNLulus=$TidakLulus/($Lulus+$TidakLulus);

$PlogLulus=(-$PLulus) * log($PLulus,2);
    if($PlogLulus=="NAN" or $PlogLulus==0){
        $PlogLulus=0;
    }
$PlogNLulus=(-$PNLulus) * log($PNLulus,2);
    if($PlogNLulus=="NAN" or $PlogNLulus==0){
        $PlogNLulus=0;
    }
$Total=$PlogLulus+$PlogNLulus;

$QueJumSekolah=mysql_query("select count(pendortu) as JumSekolah
from kelulusan where nilaites='$AsalSekolah' ",$conn);
$JumSekolah=mysql_fetch_array($QueJumSekolah);

$QueJumSekolah2=mysql_query("select count(pendortu) as JumSekolah2
from kelulusan ",$conn);
$JumSekolah2=mysql_fetch_array($QueJumSekolah2);
$PAsalSekolah=$JumSekolah[JumSekolah]/$JumSekolah2[JumSekolah2];
$TotalxPAsalSekolah=$Total*$PAsalSekolah;

$GTotTotalxPAsalSekolah=$GTotTotalxPAsalSekolah+$TotalxPAsalSekol
ah;

$Gain=$Entropy - $GTotTotalxPAsalSekolah;

$GTotTotalxPAsalSekolah=round($GTotTotalxPAsalSekolah,6);
$Gain=round($Gain,5);

echo "<tr>
    <td>$AsalSekolah</td>
    <td>$Lulus</td>
    <td>$TidakLulus</td>
    <td>$PLulus</td>
    <td>$PNLulus</td>

    <td>$PlogLulus</td>

```

```

        <td>$PlogNLulus</td>
        <td>$Total</td>
        <td>$PAsalSekolah</td>
        <td>$TotalxPAsalSekolah</td>
        <td></td>
        <td></td>
    </tr>";

}while($sqlbaris=mysql_fetch_row($sqlquery));

echo "<tr>
        <td colspan=10>&nbsp;</td>
        <td>$GTotTotalxPAsalSekolah</td>
        <td><b><font
color=red>$Gain</font></b></td>
        </tr>
        ";
//TAMPIL berdasarkan Nilai Wawancara
$GTotTotalxPAsalSekolah=0;
$sqltampil="select          nilaiwawancara,          sum(IF(ket='LULUS',1,0)),
sum(IF(ket='TIDAK LULUS',1,0))
                                from kelulusan
                                group      by
                                nilaiwawancara
                                order  by  id
DESC";
$sqlquery=mysql_query($sqltampil,$conn);
$sqlbaris=mysql_fetch_row($sqlquery);
echo "<tr>
        <td colspan=9>&nbsp;</td>
        </tr>";

echo "<tr>
        <th>Nilai Wawancara</th>
        <th>Jum Lulus</th>
        <th>Jum Tidak Lulus</th>
        <th>P(Lulus)</th>
        <th>P(Tidak Lulus)</th>

```

```

        <th>P(LULUS)*LOG2 P(LULUS)</th>
        <th>P(TIDAK LULUS)*LOG2 P(TIDAK LULUS)</th>
        <th>Total</th>
        <td>P(ASAL SEKOLAH=...)</td>
        <th>TOTAL X P(ASAL SEKOLAH=...)</th>
        <th>E(TOTAL*P(ASAL SEKOLAH=..))</th>
        <th bgcolor=red>GAIN</th>

    </tr>";

    //entropy H(t)
    $queTotLulus=mysql_query("select PLulus,PTidakLulus from asalsekolah
",$conn);
    $TL=mysql_fetch_array($queTotLulus);

    $TotLulus=$TL[PLulus];
    $TotNLulus=$TL[PTidakLulus];
    $TotAllLulusNLulus=$TotLulus+$TotNLulus;

    $pjLulus=$TotLulus/$TotAllLulusNLulus;
    $log2pjLulus=log($pjLulus,2);
    $minuslog2pjLulus=(-$log2pjLulus)*$pjLulus;

    $pjNLulus=$TotNLulus/$TotAllLulusNLulus;
    $log2pjNLulus=log($pjNLulus,2);
    $minuslog2pjNLulus=(-$log2pjNLulus) * $pjNLulus;

    $Entropy=$minuslog2pjLulus+$minuslog2pjNLulus;

do{
    list( $AsalSekolah,
          $Lulus,
          $TidakLulus
        )=$sqlbaris;

    $QueJumlahLulusAll=mysql_query("select count(ket) as JumLulus from
kelulusan where ket='LULUS' ",$conn);
    $jl=mysql_fetch_array($QueJumlahLulusAll);

    $QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
from kelulusan where ket='TIDAK LULUS' ",$conn);
    $jln=mysql_fetch_array($QueJumlahNoLulusAll);

```

```

$JumLulus=$jl[JumLulus];
$JumTidakLulus=$jln[JumNLulus];

$HitungLulus=round($Lulus/$JumLulus, 5);
$HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

$NilaiLulus=round($HitungLulus * 100 );
$NilaiNoLulus=round($HitungNoLulus * 100 );

$PLulus=$Lulus/($Lulus+$TidakLulus);
$PNLulus=$TidakLulus/($Lulus+$TidakLulus);

$PlogLulus=(-$PLulus) * log($PLulus,2);
    if($PlogLulus=="NAN" or $PlogLulus==0){
        $PlogLulus=0;
    }
$PlogNLulus=(-$PNLulus) * log($PNLulus,2);
    if($PlogNLulus=="NAN" or $PlogNLulus==0){
        $PlogNLulus=0;
    }

$Total=$PlogLulus+$PlogNLulus;

$QueJumSekolah=mysql_query("select count(pendortu) as JumSekolah
from kelulusan where nilaiwawancara='$AsalSekolah' ",$conn);
$JumSekolah=mysql_fetch_array($QueJumSekolah);
$QueJumSekolah2=mysql_query("select count(pendortu) as JumSekolah2
from kelulusan ",$conn);
$JumSekolah2=mysql_fetch_array($QueJumSekolah2);
$PAsalSekolah=$JumSekolah[JumSekolah]/$JumSekolah2[JumSekolah2];
$TotalxPAsalSekolah=$Total*$PAsalSekolah;

$GTotTotalxPAsalSekolah=$GTotTotalxPAsalSekolah+$TotalxPAsalSekol
ah;

$Gain=$Entropy - $GTotTotalxPAsalSekolah;

$GTotTotalxPAsalSekolah=round($GTotTotalxPAsalSekolah,6);
$Gain=round($Gain,5);

echo "<tr>
        <td>$AsalSekolah</td>
        <td>$Lulus</td>

```

```

        <td>$TidakLulus</td>
        <td>$PLulus</td>
        <td>$PNLulus</td>

        <td>$PlogLulus</td>
        <td>$PlogNLulus</td>
        <td>$Total</td>
        <td>$PAsalSekolah</td>
        <td>$TotalxPAsalSekolah</td>
        <td></td>
        <td></td>

</tr>";

}while($sqlbaris=mysql_fetch_row($sqlquery));

echo "<tr>
        <td colspan=10>&nbsp;</td>
        <td>$GTotTotalxPAsalSekolah</td>
        <td><b><font
color=red>$Gain</font></b></td>
        </tr>
        ";

//TAMPIL berdasarkan Nilai BTHQ
$GTotTotalxPAsalSekolah=0;
$sqltampil="select  nilaibthq,  sum(IF(ket='LULUS',1,0)),  sum(IF(ket='TIDAK
LULUS',1,0))
                                from kelulusan
                                group      by
nilaibthq
                                order by id ";

$sqlquery=mysql_query($sqltampil,$conn);
$sqlbaris=mysql_fetch_row($sqlquery);

echo "<tr>
        <td colspan=9>&nbsp;</td>
        </tr>";

echo "<tr>
        <th>Nilai Wawancara</td>

```

```

        <th>Jum Lulus</td>
        <th>Jum Tidak Lulus</td>
        <th>P(Lulus)</td>
        <th>P(Tidak Lulus)</td>

        <th>P(LULUS)*LOG2 P(LULUS)</td>
        <th>P(TIDAK LULUS)*LOG2 P(TIDAK LULUS)</td>
        <th>Total</td>
        <td>P(ASAL SEKOLAH=...)</td>
        <th>TOTAL X P(ASAL SEKOLAH=....)</td>
        <th>E(TOTAL*P(ASAL SEKOLAH=..))</td>
        <th bgcolor=red>GAIN</td>

    </tr>";

    //entropy H(t)
    $queTotLulus=mysql_query("select PLulus,PTidakLulus from asalsekolah
", $conn);
    $TL=mysql_fetch_array($queTotLulus);

    $TotLulus=$TL[PLulus];
    $TotNLulus=$TL[PTidakLulus];
    $TotAllLulusNLulus=$TotLulus+$TotNLulus;

    $pjLulus=$TotLulus/$TotAllLulusNLulus;
    $log2pjLulus=log($pjLulus,2);
    $minuslog2pjLulus=(-$log2pjLulus)*$pjLulus;

    $pjNLulus=$TotNLulus/$TotAllLulusNLulus;
    $log2pjNLulus=log($pjNLulus,2);
    $minuslog2pjNLulus=(-$log2pjNLulus) * $pjNLulus;

    $Entropy=$minuslog2pjLulus+$minuslog2pjNLulus;

do{
    list( $AsalSekolah,
        $Lulus,
        $TidakLulus
        )=$sqlbaris;
        $QueJumlahLulusAll=mysql_query("select count(ket) as JumLulus
from kelulusan where ket='LULUS' ", $conn);
        $jl=mysql_fetch_array($QueJumlahLulusAll);

```

```

$QueJumlahNoLulusAll=mysql_query("select count(ket) as JumNLulus
from kelulusan where ket='TIDAK LULUS' ",$conn);
$jln=mysql_fetch_array($QueJumlahNoLulusAll);

$JumLulus=$jln[JumLulus];
$JumTidakLulus=$jln[JumNLulus];

$HitungLulus=round($Lulus/$JumLulus, 5);
$HitungNoLulus=round($TidakLulus/$JumTidakLulus,5);

$NilaiLulus=round($HitungLulus * 100 );
$NilaiNoLulus=round($HitungNoLulus * 100 );

$PLulus=$Lulus/($Lulus+$TidakLulus);
$PNLulus=$TidakLulus/($Lulus+$TidakLulus);

$PlogLulus=(-$PLulus) * log($PLulus,2);
if($PlogLulus=="NAN" or $PlogLulus===-0){
    $PlogLulus=0;
}
$PlogNLulus=(-$PNLulus) * log($PNLulus,2);
if($PlogNLulus=="NAN" or $PlogNLulus===-0){
    $PlogNLulus=0;
}
$Total=$PlogLulus+$PlogNLulus;

$QueJumSekolah=mysql_query("select count(pendortu) as JumSekolah
from kelulusan where nilaibthq='$AsalSekolah' ",$conn);
$JumSekolah=mysql_fetch_array($QueJumSekolah);

$QueJumSekolah2=mysql_query("select count(pendortu) as JumSekolah2
from kelulusan ",$conn);
$JumSekolah2=mysql_fetch_array($QueJumSekolah2);
$PAsalSekolah=$JumSekolah[JumSekolah]/$JumSekolah2[JumSekolah2];
$TotalxPAsalSekolah=$Total*$PAsalSekolah;

$GTotTotalxPAsalSekolah=$GTotTotalxPAsalSekolah+$TotalxPAsalSekolah;

$Gain=$Entropy - $GTotTotalxPAsalSekolah;

$GTotTotalxPAsalSekolah=round($GTotTotalxPAsalSekolah,6);

```

```

$Gain=round($Gain,5);

echo "<tr>
    <td>$AsalSekolah</td>
    <td>$Lulus</td>
    <td>$TidakLulus</td>
    <td>$PLulus</td>
    <td>$PNLulus</td>

    <td>$PlogLulus</td>
    <td>$PlogNLulus</td>
    <td>$Total</td>
    <td>$PAsalSekolah</td>
    <td>$TotalxPAsalSekolah</td>
    <td></td>
    <td></td>

</tr>";

}while($sqlbaris=mysql_fetch_row($sqlquery));

echo "<tr>
    <td colspan=10>&nbsp;</td>
    <td>$GTotTotalxPAsalSekolah</td>
    <td><b><font
color=red>$Gain</font></b></td>
    </tr>
    ";

echo "</table>";

?>

```

BAB V

PENUTUP

A. Kesimpulan

Berdasarkan penelitian yang telah dilakukan, diperoleh kesimpulan sebagai berikut :

1. Penerapan metode klasifikasi dengan Naive bayes dan C.45 dalam memprediksi tingkat kesuksesan kelulusan murid pada IAIN Bone dilakukan dengan mengolah data kelulusan penerimaan Mahasiswa Baru Jalur Mandiri Tahun 2018 dengan mengambil data dari database Sistem Informasi (SISFO) Akademik.
2. Pengolahan data pada Algoritma Naive Bayes diperoleh Nilai Variabel Lulus sebesar 64% dan Nilai Variabel Tidak Lulus sebesar 36%. Jadi, ketika Nilai Variabel Lulus lebih besar dibandingkan dengan Nilai Variabel Tidak Lulus maka diperoleh Hasil Akhir “LULUS”.
3. Pengolahan data pada Algoritma C.45 diperoleh urutan gain yaitu Nilai Baca Tulis Huruf Qur’an (BTHQ) sebesar 0,9417, Nilai Tes sebesar 0,616, Nilai Wawancara sebesar 0,4717, Pendapatan Orang Tua sebesar 0,0671, Tahun Lulus sebesar 0,0534, Pendidikan Orang Tua sebesar 0,01 dan Nilai Ujian Nasional sebesar 0,0033.
4. Penelitian ini menghasilkan pola sekuensial untuk menjadi referensi bagi murid SMA yang akan mendaftar di Perguruan Tinggi Negeri IAIN Bone .
5. Pola sekuensial yang dihasilkan untuk kategori lulus sebanyak 24 pola dan tidak lulus sebanyak 48 pola yang dapat dilihat pada tabel 21.
6. Dari hasil pengujian ditemukan pada algoritma Naive Bayes menghasilkan Nilai Precision sebesar 100%, Recall 100%, Akurasi 100%. Sedangkan C.45 juga menghasilkan nilai yang sama yaitu Nilai Precision sebesar 100%, Recall 100%, Akurasi 100%. Ini artinya bahwa pengujian yang dilakukan pada kedua algoritma tersebut menghasilkan nilai prediksi yang sama.

B. Saran

Dari hasil pengujian dan kesimpulan, maka saran untuk penelitian selanjutnya yaitu:

1. Pengolahan data dapat dilakukan kembali dengan penambahan jumlah data yang digunakan sebagai sumber data agar hasil dari prediksi menjadi lebih baik.
2. Mencari algoritma lain dalam pengolahan data mining sebagai bahan perbandingan akurasi untuk penentuan tingkat kesuksesan kelulusan murid SMA.

DAFTAR PUSTAKA

- [1] https://id.wikipedia.org/wiki/Perguruan_tinggi [1]
- [2] Sitti Nasirotn. (2013). Pengaruh Kondisi Sosial Ekonomi Dan Pendidikan Orang Tua Terhadap Motivasi Melanjutkan Pendidikan Ke Perguruan Tinggi Pada Siswa. Jurnal Pendidikan Ekonomi IKIP Veteran Semarang. Volume 1 No.2 Oktober 2013. (<https://media.neliti.com/media/publications/37061>, diakses 27 Maret 2019)
- [3] Undang-Undang RI No. 12 Tahun 2012.(<http://sumberdaya.ristekdikti.go.id/wp-content/uploads/2016/02/uu-nomor-12-tahun-2012-ttg-pendidikan-tinggi.pdf>, diakses 25 Maret 2019)
- [4] Undang-Undang Dasar 1945.(<http://www.dpr.go.id/jdih/uu1945>, diakses 25 Maret 2019)
- [5] Marselina Silvia Suhartinah, Ernastuti.(2010).Graduation Prediction Of Gunadarma University Students Using Algorithm and Naive Bayes C4.5 Algorithm.Undergraduate Program, Faculty of Industrial Engineering Gunadarma University.(<http://www.gunadarma.ac.id>, diakses 22 Juni 2019).
- [6] <https://www.fandimedia.com/2015/11/entropi-dalam-algoritma-c45.html>
- [7]Yuandri Trisaputra, Indriyani, Shellafuri Mardika Biru, Muhammad Ervan.(2015).Klasifikasi ProfilSiswa SMA/SMK Yang Masuk PTN (Perguruan Tinggi Negeri) dengan k-Nearest Neighbor. Departemen Ilmu Komputer, FMIPA, Institut Pertanian Bogor. (<https://www.researchgate.net/publication/305917029>, diakses 27 Juni 2019)
- [8]Yogiek Indra Kurniawan.(2018).Perbandingan Algoritma Naïve Bayes dan C.45 Dalam Klasifikasi Data Mining.Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK).Vol.5 No. 4 September 2018. 455-464. (<https://www.researchgate.net/publication/328217641>, diakses 2 Juli 2019)
- [9] Bkti Amalia Arifiyani, Ridha Sefina Samosir.(2018).Sistem Simulasi Prediksi Profil Kelulusan Mahasiswa Dengan Decision Tree. Jurnal Sains dan Teknologi.Kalbisceentia, Volume 5 No. 2 Agustus 2018.ISSN 2356-4393(<http://research.kalbis.ac.id/Research/Files/Article/Full/MBLGUZ8JVOF9GV8CG85DTX6YN.pdf>, diakses 2 Juli 2019)

- [10] A'ang Subiyakto. Penggunaan Algoritma Klasifikasi Dalam Data Mining. Program Studi Sistem Informasi Fakultas Sains dan Teknologi UIN Jakarta. (<https://www.researchgate.net/publication/260025731>, diakses 5 Juli 2019)
- [12] Susi Mashlahah. (2013). Prediksi Kelulusan Mahasiswa Menggunakan Metode Decision Tree Dengan Penerapan Algoritma C4.5. Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Maulana Malik Ibrahim Malang. (<http://etheses.uin-malang.ac.id/7505/1/07650150.pdf>, diakses 7 Agustus 2019)
- [13] Fajar Astuti Hermawati. (2013). Data Mining. Penerbit Andi, Yogyakarta.

LAMPIRAN

Data Testing

1. Perhitungan Algoritma Naive Bayes

Penentuan nilai persentase lulus dan tidak lulus

$$\text{Lulus} = 6/10 = 0,6 = 60\%$$

$$\text{Tidak Lulus} = 4/10 = 0,4 = 40\%$$

Penentuan nilai persentase dari variabel yang ada

No	Variabel	Kategori	Lulus			Tidak Lulus		
			Total Data	Hitung	Nilai	Total Data	Hitung	Nilai
1	Tahun Lulus	2017	2	2/6=0,33333	33%	0	0/4=0	0%
		2018	4	4/6=0,66667	67%	4	4/4=1	100%
		Jumlah			100%			100%
2	Pendidikan Orang Tua	Dasar	2	2/6=0,33333	33%	3	3/4=0,75	75%
		Menengah	2	2/6=0,33333	33%	1	1/4=0,25	25%
		Tinggi	2	2/6=0,33333	33%	0	0/4=0	0%
	Jumlah			100%			100%	
3	Penghasilan Orang Tua	Rendah	4	4/6=0,66667	67%	0	0/4=0	0%
		Sedang	0	0/6=0	0%	4	4/4=1	100%
		Sangat Tinggi	2	2/6=0,33333	33%	0	0/4=0	0%
	Jumlah			100%			100%	
4	Nilai Ujian Nasional	Baik	5	5/6=0,83333	83%	2	2/4=0,5	50%
		Sangat Baik	1	1/6=0,16667	17%	2	2/4=0,5	50%
	Jumlah			100%			100%	
5	Nilai Tes	Cukup	0	0/6=0	0%	4	4/4=1	100%
		Baik	2	2/6=0,33333	33%	0	0/4=0	0%
		Sangat Baik	4	4/6=0,66667	67%	0	0/4=0	0%
	Jumlah			100%			100%	
6	Nilai Wawancara	Baik	2	2/6=0,33333	33%	3	3/4=0,75	75%
		Sangat Baik	4	4/6=0,66667	67%	1	1/4=0,25	25%
	Jumlah			100%			100%	
7	Nilai Baca Tulis Huruf Qur'an	Lancar, Fasih, Tajwid Bagus	4	4/6=0,66667	67%	0	0/4=0	0%
		Lancar, Fasih, Tajwid Kurang Bagus	2	2/6=0,33333	33%	0	0/4=0	0%
		Tidak Lancar	0	0/6=0	0%	4	4/4=1	100%
	Jumlah			100%			100%	

2. Perhitungan Algoritma C.45

Penentuan Nilai frekuensi dan nilai entropy

Hasil	Frekuensi	Pj	log ₂ pj	-pj.log ₂ pj
Lulus	6	0,6	-0.736965594	0.442179356
Tidak Lulus	4	0,4	-1.321928095	0.528771238
Total Data	170	Total Entropy (S)		0,970950594

Keterangan :

- Lulus

$$P_j = 6/10 = 0,6$$

$$\log_2 p_j = \log_2(0,6) = -0.736965594$$

$$-p_j \cdot \log_2 p_j = (-0,6) \cdot -0.736965594 = \mathbf{0,442179356}$$

- Tidak Lulus

$$P_j = 4/10 = 0,4$$

$$\log_2 p_j = \log_2(0,4) = -1.321928095$$

$$-p_j \cdot \log_2 p_j = (-0,4) \cdot -1.321928095 = \mathbf{0.528771238}$$

Nilai Entrophy

$$= (-p_j \cdot \log_2 p_j | \text{lulus}) + (-p_j \cdot \log_2 p_j | \text{tidak lulus})$$

$$= \mathbf{0,442179356 + 0,528771238}$$

$$= \mathbf{0,970950594}$$

Menghitung Gain Masing-Masing Variabel

1. Variabel Tahun Lulus

- 2017

$$= \frac{-2}{2} \log_2 \frac{2}{2} + \left(\frac{-0}{2} * \log_2 \frac{0}{2} \right)$$

$$= 0 + 0 = 0$$

$$= 0 * 2/10$$

$$= 0$$

- 2018

$$= \frac{-4}{8} \log_2 \frac{4}{8} + \left(\frac{-4}{8} * \log_2 \frac{4}{8} \right)$$

$$= 0,5 + 0,5 = 1$$

$$= 1 * 8/10$$

$$= 0,8$$

Nilai Gain untuk variabel Tahun Lulus, yaitu :

$$= \mathbf{0,970950594} - \left(\frac{2}{10} * (0) + \frac{8}{10} * (1) \right)$$

$$= \mathbf{0,970950594} - (0 + 0,8)$$

$$= \mathbf{0,970950594} - \mathbf{0,8}$$

$$= \mathbf{0,170951}$$

2. Variabel Pendidikan Orang Tua

- Dasar

$$= \frac{-2}{5} \log_2 \frac{2}{5} + \left(\frac{-3}{5} * \text{Log}_2 \frac{3}{5} \right)$$

$$= 0.528771 + 0.442179 = 0.970951$$

$$= 0.970951 * 5/10$$

$$= 0.485475$$

- Menengah

$$= \frac{-2}{3} \log_2 \frac{2}{3} + \left(\frac{-1}{3} * \text{Log}_2 \frac{1}{3} \right)$$

$$= 0.389975 + 0.528321 = 0.918296$$

$$= 0.918296 * 3/10$$

$$= 0.275489$$

- Tinggi

$$= \frac{-2}{2} \log_2 \frac{2}{2} + \left(\frac{-0}{2} * \text{Log}_2 \frac{0}{2} \right)$$

$$= 0 + 0 = 0$$

$$= 0 * 2/0$$

$$= 0$$

Nilai Gain untuk variabel Pendidikan Orang Tua, yaitu :

$$\begin{aligned} &= \mathbf{0,970950594} - \left(\frac{5}{10} * (0.970951) + \frac{3}{10} * (0.918296) + \frac{2}{10} * (0) \right) \\ &= \mathbf{0,970950594} - (0.485475 + 0.275489 + 0) \\ &= \mathbf{0,970950594} - \mathbf{0.760964} \\ &= \mathbf{0.209987} \end{aligned}$$

3. Variabel Penghasilan Orang Tua

- Rendah

$$\begin{aligned} &= \frac{-4}{4} \log_2 \frac{4}{4} + \left(\frac{-0}{4} * \text{Log}_2 \frac{0}{4} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 4/10 \\ &= 0 \end{aligned}$$

- Sedang

$$\begin{aligned} &= \frac{-0}{4} \log_2 \frac{0}{4} + \left(\frac{-4}{4} * \text{Log}_2 \frac{4}{4} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 4/10 \\ &= 0 \end{aligned}$$

- Sangat Tinggi

$$\begin{aligned} &= \frac{-2}{2} \log_2 \frac{2}{2} + \left(\frac{-0}{2} * \text{Log}_2 \frac{0}{2} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 2/10 \\ &= 0 \end{aligned}$$

Nilai Gain untuk variabel Penghasilan Orang Tua, yaitu :

$$\begin{aligned} &= \mathbf{0,970950594} - \left(\frac{4}{10} * (0) + \frac{4}{10} * (0) + \frac{2}{10} * (0) \right) \\ &= \mathbf{0,970950594} - (0 + 0 + 0) \\ &= \mathbf{0,970950594} - \mathbf{0} \\ &= \mathbf{0.970951} \end{aligned}$$

4. Variabel Nilai Ujian Nasional

- Baik

$$\begin{aligned} &= \frac{-5}{7} \log_2 \frac{5}{7} + \left(\frac{-2}{7} * \text{Log}_2 \frac{2}{7} \right) \\ &= 0.346733 + 0.516387 = 0.863121 \\ &= 0.863121 * 7/10 \\ &= 0.604184 \end{aligned}$$

- Sangat Baik

$$\begin{aligned} &= \frac{-1}{3} \log_2 \frac{1}{3} + \left(\frac{-2}{3} * \text{Log}_2 \frac{2}{3} \right) \\ &= 0.528321 + 0.389975 = 0.918296 \\ &= 0.918296 * 3/10 \\ &= 0.275489 \end{aligned}$$

Nilai Gain untuk variabel Penghasilan Orang Tua, yaitu :

$$\begin{aligned} &= \mathbf{0,970950594} - \left(\frac{7}{10} * (0.863121) + \frac{3}{10} * (0.902393) \right) \\ &= \mathbf{0,970950594} - (0.604184 + 0.275489) \\ &= \mathbf{0,970950594} - \mathbf{0.879673} \\ &= \mathbf{0.091277} \end{aligned}$$

5. Variabel Nilai Tes

- Cukup

$$\begin{aligned} &= \frac{-0}{4} \log_2 \frac{0}{4} + \left(\frac{-4}{4} * \text{Log}_2 \frac{4}{4} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 4/10 \\ &= 0 \end{aligned}$$

- Baik

$$\begin{aligned} &= \frac{-2}{2} \log_2 \frac{2}{2} + \left(\frac{-0}{2} * \text{Log}_2 \frac{0}{2} \right) \\ &= 0 + 0 = 0 \\ &= 0 * 2/2 \end{aligned}$$

$$= 0$$

- Sangat Baik

$$= \frac{-0}{4} \log_2 \frac{0}{4} + \left(\frac{-4}{4} * \text{Log}_2 \frac{4}{4} \right)$$

$$= 0 + 0 = 0$$

$$= 0 * 4/10$$

$$= 0$$

Nilai Gain untuk variabel Nilai Tes, yaitu :

$$= \mathbf{0,970950594} - \left(\frac{4}{10} * (0) + \frac{2}{10} * (0) + \frac{4}{10} * (0) \right)$$

$$= \mathbf{0,970950594} - (0 + 0 + 0)$$

$$= \mathbf{0,970950594} - \mathbf{0}$$

$$= \mathbf{0,970951}$$

6. Variabel Nilai Wawancara

- Baik

$$= \frac{-2}{5} \log_2 \frac{2}{5} + \left(\frac{-3}{5} * \text{Log}_2 \frac{3}{5} \right)$$

$$= 0.528771 + 0.442179 = 0.970951$$

$$= 0.970951 * 5/10$$

$$= 0.485475$$

- Sangat baik

$$= \frac{-4}{5} \log_2 \frac{4}{5} + \left(\frac{-1}{5} * \text{Log}_2 \frac{1}{5} \right)$$

$$= 0.257542 + 0.464386 = 0.721928$$

$$= 0.721928 * 5/10$$

$$= 0.360964$$

Nilai Gain untuk variabel Nilai Wawancara, yaitu :

$$= \mathbf{0,970950594} - \left(\frac{5}{10} * (0.970951) + \frac{5}{10} * (0.721928) \right)$$

$$= \mathbf{0,970950594} - (0.485475 + 0.360964)$$

$$= \mathbf{0,941702295} - \mathbf{0.846439}$$

$$= \mathbf{0.124511}$$

7. Variabel Nilai Baca Tulis Huruf Qur'an

- Lancar, Fasih, Tajwid Bagus

$$= \frac{-4}{4} \log_2 \frac{4}{4} + \left(\frac{-0}{4} * \text{Log}_2 \frac{0}{4} \right)$$

$$= 0 + 0 = 0$$

$$= 0 * 4/10$$

$$= 0$$

- Lancar, Fasih, Tajwid Kurang Bagus

$$= \frac{-2}{2} \log_2 \frac{2}{2} + \left(\frac{-0}{2} * \text{Log}_2 \frac{0}{2} \right)$$

$$= 0 + 0 = 0$$

$$= 0 * 2/10$$

$$= 0$$

- Tidak Lancar

$$= \frac{-0}{4} \log_2 \frac{0}{4} + \left(\frac{-4}{4} * \text{Log}_2 \frac{4}{4} \right)$$

$$= 0 + 0 = 0$$

$$= 0 * 4/10$$

$$= 0$$

Nilai Gain untuk variabel Nilai Baca Tulis Huruf Qur'an, yaitu :

$$= \mathbf{0,970950594} - \left(\frac{4}{10} * (0) + \frac{2}{10} * (0) + \frac{4}{10} * (0) \right)$$

$$= \mathbf{0,970950594} - (0 + 0 + 0)$$

$$= \mathbf{0,970950594} - \mathbf{0}$$

$$= \mathbf{0.970951}$$