

PENERAPAN DATA MINING DENGAN METODE *NAIVE BAYES CLASIFFIER* UNTUK Mendukung STRATEGI PROMOSI FAKULTAS SAINS DAN TEKNOLOGI UIN RADEN FATAH PALEMBANG

SKRIPSI

WAWAN (13540157)



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI RADEN FATAH
PALEMBANG 2018**

PENERAPAN DATA MINING DENGAN METODE *NAIVE BAYES CLASIFFIER* UNTUK Mendukung STRATEGI PROMOSI FAKULTAS SAINS DAN TEKNOLOGI UIN RADEN FATAH PALEMBANG

SKRIPSI

**Dianjukan Untuk Memenuhi Salah Satu Syarat
Guna Memperoleh Gelas Strata Satu Sarjana Komputer (S. Kom)
Pada Program Studi Sistem Informasi**

Oleh

WAWAN (13540157)



**PROGRAM STUDI SISTEM INFORMASI
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI RADEN FATAH
PALEMBANG 2018**

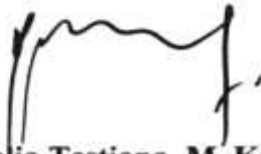
HALAMAN PENGESAHAN

PENERAPAN DATA MINING DENGAN METODE *NAIVE BAYES CLASIFFIER* UNTUK Mendukung STRATEGI PROMOSI FAKULTAS SAINS DAN TEKNOLOGI UIN RADEN FATAH PALEMBANG

Oleh:
WAWAN
13540157

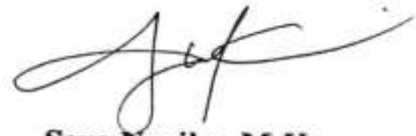
Telah dipertahankan didepan sidang pengujian skripsi
pada tanggal 12 SEPTEMBER 2018
dan dinyatakan memenuhi syarat untuk memperoleh gelar Sarjana
Komputer dalam bidang Sistem Informasi

Pembimbing I



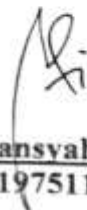
Gusmelia Testiana, M. Kom
NIP. 197508012009122001

Pembimbing II



Seva Novika, M. Kom
NIDN. 0218119101

Mengetahui,
Kepala Program Studi Sistem Informasi
Fakultas Sains dan Teknologi
UIN Raden Fatah Palembang



Ruliansyah, S.T, M.Kom
NIP.197511222006041003

PERSETUJUAN TIM PENGUJI SKRIPSI

Judul Skripsi : Penerapan *Data Mining* Dengan Metode *Naive Bayes Classifier* Untuk Mendukung Strategi Promosi Fakultas Sains dan Teknologi UIN Raden Fatah Palembang
NIM : 13540157
Program : Sarjana (S1) Fakultas Sains dan Teknologi

Telah disetujui oleh tim penguji sidang skripsi.

1. Ketua : Gusmelia Testiana, M.Kom
NIP. 197508012009122001
2. Sekertaris : Fathiyah Nopriani, ST., M.Kom
NIDN. 2017118205
3. Penguji I : M.Haviz Irfani, S.Si., M.T.I.
NIDN. 0209087903
4. Penguji II : Muhamad Kadafi, M.Kom
NIDN. 0223108404



(.....)

(.....)

(.....)

(.....)

Diuji di Palembang pada tanggal 12 September 2018
Waktu : 08.00-09.00 WIB
Hasil/IPK : B / 3.09
Predikat : Baik

Dekan,
Fakultas Sains dan Teknologi
UIN Raden Fatah



Dr. Dian Erlina, S.Pd, M.Hum
NIP. 197301021999032001

MOTTO DAN PERSEMBAHAN

Motto:

“Memulai dengan penuh keyakinan, Menjalankan dengan penuh keikhlasan dan Menyelesaikan dengan penuh kebahagiaan”

Skripsi ini kupersembahkan Kepada :

- * Kedua orang tua ku tercinta Ayahanda Talib (Alm) dan Ibunda tercinta Samsiah yang selalu memberikan kasih sayangnya yang tak terhingga dan selalu mendoakanku serta menanti keberhasilanku.
- * Saudara tercinta Hernina dan Marlina serta keluarga besarku terima kasih buat dukungannya.
- * Ibuk Gusmelia Testiani, M.Kom dan Ibuk Seva Novika, M.kom yang telah membimbing dan memotivasiku untuk menyelesaikan skripsi ini.
- * Dosen–dosen Sistem Informasi yang telah mengajarkan banyak ilmu kepadaku terima kasih semuanya.
- * Teman seperjuangan saya Syarif dan Sulaiman yang dari awal sampai saat ini tetap selalu bersama dalam memberikan semangat dan dukungan untuk satu tujuan.
- * Teman-teman prodi sistem informasi angkatan 2013 khususnya anak kelas SIE terima kasih dukungannya.
- * Dan semua orang yang telah membantu terselesaikannya skripsi ini terima kasih banyak.
- * Almamater yang saya banggakan.

HALAMAN PERNYATAAN

Saya yang bertanda-tangan dibawah ini:

Nama : WAWAN
Tempat dan tanggal lahir : Mesuji, 22 Mei 1994
Program Studi : Sistem Informasi
NIM : 13540157

Menyatakan dengan sesungguhnya bahwa:

1. Seluruh data informasi, interpretasi serta pernyataan dalam pembahasan dan kesimpulan yang disajikan dalam skripsi ini, kecuali yang disebutkan sumbernya ditulis dalam daftar pustaka adalah merupakan hasil pengamatan, penelitian, pengolahan, serta pemikiran saya dengan pengarahan dari pembimbing yang diterapkan.
 2. Skripsi yang saya tulis ini adalah asli, bukan jiplakan dan belum pernah diajukan untuk mendapatkan gelar akademik, baik di UIN Raden Fatah maupun perguruan tinggi lainnya.
 3. Apabila dikemudian hari ditemukan adanya bukti ketidakbenaran dalam pernyataan tersebut diatas, maka saya bersedia menerima sanksi akademis berupa pembatalan gelar yang saya peroleh melalui pengajuan karya ilmiah ini.
- Demikian pernyataan ini dibuat dengan penuh kesadaran dan dapat dipertanggung jawabkan

Palembang, 25 Oktober 2018

Yang membuat pernyataan,



WAWAN

NIM. 13540157

**IMPLEMENTASI DATA MINING WITG NAIVE BAYES
CLASSIFIER METHOD TO SUPPORT THE PROMOTION
STRATEGY OF SCIENCE AND TECNOLOGY FACULTY UIN
RADEN FATAH PALEMBANG**

ABSTRACT

Data mining is the process of finding new patterns in the data by filtering large quantities. search pattern data mining uses pattern recognition technology that is similar to the statistical techniques and mathematical techniques. Found pattern is expected to provide useful information to produce economic benefits, effectiveness and efficiency. One method is the classification of data mining is data mining techniques have the capability classification which aims to make decisions predict a case, based on the classification results obtained. Naive Bayes classifier algorithm can be used to predict interest in studies based on the evidence provided. Naive Bayes classifier algorithm is one of data mining methods that can be used to support the promotion of effective strategies and efficient. The results of this research is the application of data mining algorithms are built using a Naive Bayes classifier that can provide vital information such as the results predicted interest in the study of students that can be used to help promotion science and technology faculty UIN Raden Fatah Palembang. The data used is enrollment the previous year. Prediction results obtained, is expected to help to support the promotion strategies that have an impact on the effectiveness and efficiency of the promotion and increase the number of new students who enroll.

Keywords: Data Mining, Naive Bayes, Classification

PENERAPAN DATA MINING DENGAN METODE *NAIVE BAYES CLASSIFIER* UNTUK Mendukung STRATEGI PROMOSI FAKULTAS SAINS DAN TEKNOLOGI UIN RADEN FATAH PALEMBANG

ABSTRAK

Data mining adalah proses untuk menemukan pola-pola baru dalam data dengan menyaring jumlah besar. pola pencarian data mining menggunakan teknologi pengenalan pola yang mirip dengan teknik statistik dan teknik matematika. Pola yang ditemukan diharapkan dapat memberikan informasi yang berguna untuk menghasilkan manfaat ekonomi, efektivitas dan efisiensi. Salah satu metode *data mining* adalah klasifikasi adalah teknik *Data Mining* yang memiliki kemampuan klasifikasi yang bertujuan mengambil keputusan dengan memprediksikan suatu kasus, berdasarkan hasil klasifikasi yang diperoleh. Algoritma *Naive Bayes Classifier* dapat digunakan untuk memprediksi minat studi berdasarkan bukti yang diberikan.. Hasil dari penelitian ini adalah penerapan algoritma *Data mining* yang dibangun menggunakan *Naive Bayes Classifier* yang dapat memberikan informasi penting seperti hasil prediksi yang menarik dalam studi mahasiswa yang dapat digunakan. Data yang digunakan adalah pendaftaran siswa tahun sebelumnya. Prediksi hasil yang diperoleh, diharapkan dapat membantu untuk mendukung strategi promosi pada tahun yang akan datang promosi dan meningkatkan jumlah mahasiswa baru yang mendaftar.

Kata Kunci : *Data Mining, Naive Bayes, Klasifikasi*

KATA PENGANTAR

Assalamu'alaikum, Wr. Wb.

Alhamdulillah, Segala puji kehadiran Allah *Subhanahu Wa Ta'ala* karena atas berkat rahmat dan hidayah-Nya sehingga laporan skripsi ini dapat terselesaikan sebagai salah satu syarat untuk menyelesaikan studi Strata Satu (S-1) pada Program Studi Sistem Informasi Fakultas Sains dan Teknologi Universitas Islam Negeri (UIN) Raden Fatah Palembang. Shalawat beserta salam semoga senantiasa tercurah kepada junjungan kita Baginda Rasulullah *Shalallahu 'Alaihi Wassalam* beserta para keluarga, sahabat, dan para pengikut Beliau hingga akhir zaman.

Setelah melakukan kegiatan penelitian, akhirnya skripsi yang berjudul "Penerapan Data Mining Dengan Metode *Naive Bayes Classifier* Untuk Mendukung Strategi Promosi Fakultas Sains dan Teknologi UIN Raden Fatah Palembang" telah selesai dibuat. Dalam pembuatan skripsi ini, tentu mendapatkan banyak bantuan dan bimbingan dari berbagai pihak dengan memberikan banyak masukan dan nasehat, serta mendukung dan menjadi motivasi tersendiri. Maka dari itu, ucapan terimakasih dihaturkan kepada:

1. Bapak Prof. Drs. H. M. Sirozi, MA. Ph.D. selaku Rektor UIN Raden Fatah Palembang.
2. Ibu Dr. Dian Erlina, S.Pd., M.Hum selaku Dekan Fakultas Sains dan Teknologi UIN Raden Fatah Palembang.
3. Bapak Ruliansyah, M.Kom selaku Ketua Program Studi Sistem Informasi Fakultas Sains dan Teknologi UIN Raden Fatah Palembang.
4. Ibu Rusmala Santi, M.Kom selaku Sekretaris Program Studi Sistem Informasi Fakultas Sains dan Teknologi UIN Raden Fatah Palembang.
5. Ibuk Gusmelia Testiani, M.Kom selaku Pembimbing I yang telah membimbing penulis sehingga penulis dapat menyelesaikan penulisan skripsi ini.

6. Ibu Seva Novika, M.Kom selaku Pembimbing II yang telah membimbing penulis sehingga penulis dapat menyelesaikan penulisan skripsi ini.
7. Untuk sahabatku Syarif Hidayatullah ,Sulaiman teman-temanku yang lainnya yang tdak bisa disebutkan satu per satu.
8. Rekan Mahasiswa/i Program Studi Sistem Informasi Angkatan 2013 terkhusus kelasi SI-E.

Semoga Allah SWT senantiasa melimpahkan rahmat dan hidayah-Nya kepada kita semua, Amin Yaa Rabbal Alamin.

Wassalamu`alaikum, Wr. Wb.

Palembang, 31 November 2018

Wawan

DAFTAR ISI

Halaman

Halaman Judul	i
Halaman Pengesahan	iii
Halaman Persetujuan	iv
Halaman Persembahan	v
Halaman Pernyataan	vi
<i>Abstract</i>	vii
Abstrak	viii
Kata Pengantar	ix
Daftar Isi	xi
Daftar Tabel	xv
Daftar Gambar	xvi
BAB I. PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 identifikasi Masalah	3
1.2.1 Rumusan Masalah	3
1.2.2 Batasan Masalah	3
1.3 Tujuan dan Manfaat Penelitian	3
1.3.1 Tujuan Penelitian	3
1.3.2 Manfaat Penelitian	3
1.4 Metodologi Penelitian	4
1.4.1 Metode Pengumpulan Data	4
1.5 Sistematika Penulisan	5
BAB II. LANDASAN TEORI	7
2.1 Ayat Al-Quran Yang berkenaan Dengan Penelitian	7
2.2 Teori Yang berhubungan dengan penelitian	7
2.2.1 Pengertian <i>Data Mining</i>	7
2.2.2 Konsep <i>Naive Bayes Classifier</i>	10
2.2.3 Klasifikasi	11
2.2.4 <i>Microsoft Office Exel</i>	12
2.2.5 <i>Promosi</i>	12
2.2.6 <i>Rapidminer</i>	13
2.3 Tinjauan Pustaka	16
BAB III. METODOLOGI PENELITIAN	16
3.1 Gambaran Umum Lokasi Penelitian	16
3.1.1 Sejarah Fakultas Sains dan Teknologi	16
3.1.2 Visi dan Misi Fakultas Sains dan Teknologi	18

3.1.3 Visi dan Misi DPM-PTSP Kota Palembang	18
3.1.3 Struktur Organisasi Fakultas Sains dan Teknologi	19
3.2 Kerangka Berfikir.....	22
3.2.1 Data Selection	22
3.2.2 Data <i>Pre processing dan Cleaning</i>	22
3.2.3 Data <i>Integration</i>	23
3.2.4 Data <i>Transformasi</i>	24
3.2.4 Data <i>Data Mining</i>	24
3.2.5 <i>Interpretation</i> atau Evaluasi.....	25
BAB IV. HASIL DAN PEMBAHASAN.....	25
4.1 Hasil	25
4.2 Pembahasan.....	25
4.2.1 <i>Naive Bayes</i>	25
4.2.2 <i>RapidMiner</i>	27
BAB V. PENUTUP.....	37
5.1 Kesimpulan	37
5.2 Saran.....	37
DAFTAR PUSTAKA	38
LAMPIRAN	

DAFTAR TABEL

	Halaman
Tabel 3.1 Contoh Data Penerimaan Mahasiswa Baru	3
Tabel 3.2 Contoh Proses Data <i>Cleaning</i>	55
Tabel 4.1 Data Pendaftar Calon Mahasiswa.....	135
Tabel 4.2 Data Pendaftar Calon Mahasiswa setelah <i>Cleaning</i>	136
Tabel 4.3 Status	138
Tabel 4.4 Data Klasifikasi Penerimaan Mahasiswa Baru	140

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Tahapan KDD	11
Gambar 3.1 Struktur Organisasi Fakultas Sains dan Teknologi.....	11
Gambar 3.2 Kerangka Berfikir	11
Gambar 4.1 Model Sub Proses Pada <i>Cross Validation Rapid Miner</i>	12
Gambar 4.2 Hasil <i>Performance Vector</i>	13
Gambar 4.3 Hasil <i>Simple Distribution</i>	14
Gambar 4.4 Grafik Berdasarkan Alamat	18
Gambar 4.5. Grafik Berdasarkan Asal Sekolah.....	19
Gambar 4.6. Grafik Berdasarkan Program Studi	22

BAB I

PENDAHULUAN

1.1 Latar Belakang

Seiring dengan perkembangan di era globalisasi dan kemajuan dibidang teknologi informasi yang memberikan pengaruh yang signifikan baik dalam bidang industri maupun bidang pendidikan di Indonesia. Pengaruh ini membawa suatu perubahan besar dalam tingkat persaingan antara instansi dan perusahaan. Banyak perusahaan dan instansi pemerintah yang menggunakan *data mining* untuk menggali informasi.

Data mining dapat diartikan sebagai suatu proses yang menggunakan teknik statistik untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai basis data. Penerapan *data mining* selain dipergunakan oleh suatu instansi maupun perusahaan penerapannya juga diterapkan di dunia pendidikan terutama pada perguruan tinggi negeri maupun swasta. Hal ini karena banyak perguruan tinggi yang berupaya untuk mendapatkan *competitive intelligence*.

Salah satu penerapan *data mining* dengan *Naive Bayes Classifier* yang diterapkan oleh suatu perguruan tinggi bertujuan untuk perkiraan mahasiswa baru yang akan mendaftar terhadap program studi yang tersedia dengan memprediksi probabilitas keanggotaan suatu *class*. *Naive Bayes Classifier* adalah pendekatan statistik yang fundamental dalam pengenalan pola (*pattern recognition*).

Penggunaan teknik *data mining* diharapkan dapat membantu mempercepat proses pengambilan keputusan, memungkinkan perusahaan untuk mengelola

informasi yang terkandung didalam data transaksi menjadi sebuah pengetahuan (*knowledge*) yang baru.

Dalam suatu hal kegiatan promosi guna meningkatkan calon mahasiswa baru untuk mendaftar di Fakultas Sains dan Teknologi UIN Raden Fatah Palembang, maka dari suatu Fakultas perlu masukan untuk meningkatkan calon mahasiswa baik lulusan Sekolah Menengah Atas (SMA) dan sederajat baik di Kota Palembang ataupun diluar kota Palembang..

Berdasarkan dari hasil observasi,masih banyak siswa atau calon pendaftar kurang mengetahui tentang Fakultas Sains dan Teknologi khususnya didaerah pedesaan atau pedalaman karena kurang promosi yang dilakukan di suatu Fakultas baik dalam pemasangan iklan maupun brosur. Sehingga masyarakat atau calon pendaftar kurang mengetahui tentang Fakultas Sains dan Teknologi. Dengan demikian, Fakultas Sains dan Teknologi perlu untuk menggali atau mencari informasi dari data pendaftaran tahun-tahun sebelumnya sebagai bahan evaluasi dan analisa untuk mendukung strategi promosi tahun berikutnya.

Berdasarkan uraian diatas maka peneliti tertarik untuk menerapkan *Data Mining* dengan *Naive Bayes* untuk untuk mengetahui minat calon Mahasiswa mendaftar di Fakultas Sains dan Teknologi data pendaftaran tahun-tahun sebelumnya sebagai bahan evaluasi dan analisa untuk mendukung strategi promosi tahun berikutnya, maka peneliti memilih judul “**Penerapan Data Mining Dengan Metode *Naive Bayes Clasiffier* Untuk Mendukung Strategi Promosi Fakultas Sains dan Teknologi UIN Raden Fatah Palembang**”.

1.2 Identifikasi Masalah

1.2.1 Rumusan Masalah

Rumusan masalah dalam penelitian ini adalah “Bagaimana penerapan *data mining* dengan metode *Naïve Bayes Classifier* menggunakan *Rapidminer* untuk mendukung strategi promosi guna meningkatkan penerimaan mahasiswa baru?”

1.2.2. Batasan Masalah

Adapun ruang lingkup batasan masalah yang dibahas agar lebih terarah dan tidak keluar dari pokok permasalahan adalah cara pemanfaatan data pendaftar calon mahasiswa Fakultas Sains dan Teknologi dengan status lulus dan tidak lulus yaitu :

1. Program studi Sistem informasi 2013 sampai dengan 2017
2. Kimia 2015 sampai dengan 2017
3. Biologi 2015 sampai dengan 2017

1.3 Tujuan dan Manfaat Penelitian

1.3.1 Tujuan Penelitian

Tujuan penelitian ini adalah untuk mendeskripsikan penerapan *data mining* dengan metode *Naïve Bayes Classifier* menggunakan *Rapidminer* untuk mendukung strategi promosi guna meningkatkan penerimaan mahasiswa baru di Fakultas Sains dan Teknologi UIN Raden Fatah Palembang.

1.3.2 Manfaat Penelitian

Manfaat dari penelitian ini adalah :

1. Dapat memberikan informasi bagaimana cara pemanfaatan basis data pendaftar calon mahasiswa sebelumnya yang besar, sehingga menjadi *data mining* yang bisa mencari solusi untuk meningkatkan strategi promosi.
2. Sebagai bahan masukan untuk mendukung pengambilan keputusan promosi yang tepat agar meningkatkan minat pendaftar di Fakultas Sains dan Teknologi UIN Raden Fatah Palembang.

1.4 Metodologi Penelitian

Metode penelitian yang digunakan adalah metode analisis deskriptif dengan pendekatan kuantitatif. Metode deskriptif merupakan metode yang digunakan untuk menggambarkan analisis suatu hasil penelitian tetapi tidak digunakan untuk membuat kesimpulan yang lebih luas.

Menurut Sugiyono (2017 : 8) dapat diartikan sebagai metode penelitian yang berlandaskan pada *filsafat positivisme*, digunakan untuk meneliti pada populasi atau sampel tertentu. Teknik pengambilan sampel pada umumnya dilakukan secara random, pengumpulan data menggunakan instrumen penelitian, analisis data bersifat kuantitatif atau statistik dengan tujuan untuk menguji hipotesis yang telah ditetapkan.

1.4.1 Metode Pengumpulan Data

Adapun teknik-teknik pengambilan data yang peneliti lakukan dalam penelitian adalah sebagai berikut:

1. Studi Pustaka

Yaitu mempelajari data dengan cara mengumpulkan dan mengkaji Promosi di Universitas Islam Negeri Raden Fatah Palembang selama periode analisis.

2. Wawancara

Merupakan teknik pengumpulan data dengan cara bertanya kepada Pusat Teknologi dan Pangkalan Data (PUSTPD) Universitas Islam Negeri Raden Fatah Palembang

1.5 Sistematika Penulisan

Sistematika penulisan skripsi ini dimaksudkan agar dapat menjadi pedoman atau garis besar penulisan laporan penelitian ini dan dapat menggambarkan secara jelas isi dari laporan penelitian sehingga terlihat hubungan antara bab awal hingga bab terakhir. Sistem penulisan laporan penelitian ini.

BAB I PENDAHULUAN

Pada bab ini berisi mengenai latar belakang, rumusan masalah, batasan masalah, tujuan dan manfaat, metode penelitian, dan sistematika penulisan.

BAB II LANDASAN TEORI

Pada Bab ini merupakan tinjauan pustaka terdiri dari dua bagian yaitu jurnal pendukung dan teori pendukung jurnal pendukung berisi berbagai teori yang didapatkan dari berbagai jurnal yang dipilih berdasarkan metode dan judul yang sama yang nantinya digunakan untuk menyusun skripsi. Teori pendukung memuat penjelasan sejarah penemu dan semua yang terkait dengan metode ini.

BAB III METODOLOGI PENELITIAN

Pada bab ini membahas tentang pengklasifikasian untuk pengujian dari analisis dan penerapan Algoritma *Naive Bayes Classifier* dalam data

mining untuk melihat tingkat rendah dan tingginya pendaftar pada Fakultas Sains dan Teknologi UIN Raden Fatah Palembang

BAB IV HASIL DAN PEMBAHASAN

Pada bab ini berisi hasil dan pembahasan mengenai penerapan *data mining* untuk menghitung tingkat rendah dan tingginya pendaftar dengan metode *Naive Bayes Classifier* menggunakan *software RapidMiner* yang dihasilkan oleh penelitian ini.

BAB V PENUTUP

Pada bab ini membahas kesimpulan dari keseluruhan penelitian yang dilakukan dan saran-saran yang diharapkan dapat berguna bagi pembaca.

BAB II

LANDASAN TEORI

2.1 Ayat Al-Qur'an Berkenaan dengan Penelitian

Kemajuan suatu teknologi di era zaman saat ini bukanlah hal yang asing lagi, kemajuan suatu teknologi saat ini di berbagai bidang ilmu mulai dari bisnis, kesehatan maupun pendidikan saat ini. Pemanfaatan sains dan teknologi dalam kajian Islam sebagaimana tertulis dalam Al-Quran dan Surat Al-Jatsiyah ayat 17 :

وَأَنبِئْهُمْ بِنَبَأِ مَنْ أَمَرَ ط فَمَا اخْتَلَفُوا إِلَّا مِنْ بَعْدِ مَا جَاءَهُمُ الْعِلْمُ بَغْيًا
بَيْنَهُمْ إِنَّ رَبَّكَ يَقْضِي بَيْنَهُمْ يَوْمَ الْقِيَامَةِ فِيمَا كَانُوا فِيهِ يَخْتَلِفُونَ ١٧

Artinya : *“Dan Kami berikan kepada mereka keterangan-keterangan yang nyata tentang urusan (agama) maka mereka tidak berselisih melainkan sesudah datang kepada mereka pengetahuan karena kedengkian yang ada di antara mereka. Sesungguhnya Tuhanmu akan memutuskan antara mereka pada hari kiamat terhadap apa yang mereka selalu berselisih padanya.”* (Q.S.Surat Al-Jatsiyah, 45:17).

2.2 Teori Yang Berhubungan dengan Penelitian

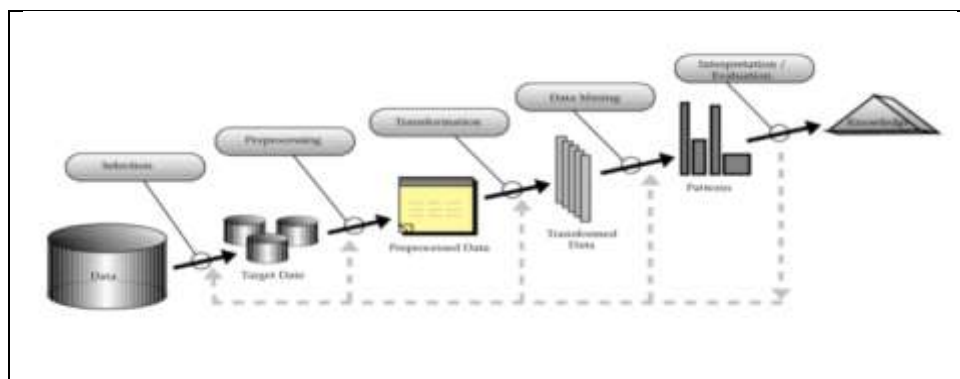
2.2.1 Pengertian *Data Mining*

Data mining adalah serangkaian proses yang menggali nilai tambah berupa informasi yang selama ini tidak diketahui secara manual dari suatu basis data. Informasi yang dihasilkan diperoleh dengan cara mengekstraksi dan mengenali pola yang penting atau menarik dari data yang terdapat pada basis data. Data

mining terutama digunakan untuk mencari pengetahuan yang terdapat dalam basis data yang besar sehingga sering disebut *Knowledge Discovery Database (KDD)* (Vulandari, 2017:1)

Data mining adalah proses interatif dan interaktif untuk menemukan pola atau model baru yang sah (sempurna), bermanfaat dan dapat dimengerti dalam suatu database yang besar (*massive database*). Data mining berisi pencarian trend atau pola yang diinginkan dalam database besar untuk membantu pengambilan keputusan di waktu yang akan datang (Hermawati, 2013 : 3).

Istilah data mining dan *Knowledge Discovery in Databases (KDD)* menurut Vulandari (2013:7), sering kali digunakan secara bergantian untuk menjelaskan proses penggalian informasi tersembunyi dalam suatu basis data yang besar. Sebenarnya kedua istilah tersebut memiliki konsep yang berbeda, tetapi berkaitan satu sama lain. Dan salah satu tahapan dalam keseluruhan proses KDD adalah *data mining* . Proses KDD (*Knowlegge discovery in databases*) menurut Vulandari, (2014: 3) secara garis besar dapat dijelaskan sebagai berikut:



Sumber : Vulandari, 2014: 3)

Gambar 2.1. Tahapan *Knowledge Discovery in Databases*

1. *Data Selection*

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam *knowledge data discovery* (KDD) dimulai. Data hasil seleksi yang akan digunakan untuk proses data mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. *Preprocessing* atau *Cleaning*

Sebelum proses *data mining* dapat dilaksanakan, perlu dilakukan proses *cleaning* pada data yang menjadi fokus *knowledge data discovery*. Proses *cleaning* mencakup antara lain membuang duplikasi data, memeriksa data yang *inkonsisten*, dan memperbaiki kesalahan pada data, seperti kesalahan cetak juga dilakukan proses *enrichment*, yaitu proses memperkaya data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan untuk KDD, seperti data atau informasi.

3. *Transformation*

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses *data mining*. Proses *coding* dalam *knowledge data discovery* merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. *Data mining*

Data mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam *data mining* sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat tergantung pada tujuan dan proses KDD secara keseluruhan.

5. *Interpretation* atau *evaluation*

Pola informasi yang dihasilkan dari proses *data mining* perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut *interpretation*. Tahap ini mencakup pemeriksaan apakah pola informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada pada sebelumnya.

2.2.2 **Konsep Naive Bayes Classifier**

Metode Naive Bayes Naive Bayes merupakan sebuah pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan menjumlahkan frekuensi dan kombinasi nilai dari dataset yang diberikan. Algoritma menggunakan teorema Bayes dan mengasumsikan semua atribut independen atau tidak saling ketergantungan yang diberikan oleh nilai pada variabel kelas. Definisi lain mengatakan Naive Bayes merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu memprediksi peluang di masa depan berdasarkan pengalaman di masa sebelumnya. (Saleh, 2015: 209).

Naive bayes adalah metode yang membagi permasalahan ke dalam sebuah kelas-kelas berdasarkan ciri-ciri persamaan dan perbedaan dengan statistik yang bisa memprediksi probabilitas sebuah kelas (Pratiwi, 2013:100).

Naive bayes classifier juga merupakan suatu klasifikasi berpeluang sederhana berdasarkan aplikasi teorema Bayes dengan asumsi antar variabel penjelas saling bebas (independen). Dalam hal ini diasumsikan bahwa kehadiran

atau ketiadaan dari suatu kejadian tertentu dari suatu kelompok tidak berhubungan dengan kehadiran atau ketiadaan dari kejadian lainnya (Pratiwi, 2013:100).

Teorema Bayes seperti berikut ini :

$$P(H|X) = \frac{P(X|H) \times P(H)}{P(X)}$$

Dimana:

X : Data dengan kelas yang belum diketahui

H : Hipotesis data X merupakan suatu kelas spesifik

P(H|X): Probabilitas hipotesis H berdasar kondisi X (*posterior probability*)

P(H) : Probabilitas hipotesis H (*prior probability*)

P(X|H): Probabilitas X berdasarkan kondisi pada hipotesis H

P(X) : Probabilitas X

2.3.3 Klasifikasi

Menurut Hermawati (2013:55), Klasifikasi merupakan proses pemberlakuan suatu fungsi tujuan (*target*) f yang memetakan tiap himpunan atribut x ke satu dari label kelas yang didefinisikan sebelumnya. Fungsi target disebut juga model klasifikasi.

Menurut Tan *et all* (2004), Klasifikasi adalah sebuah proses untuk menemukan model yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang kelasnya tidak diketahui. Di dalam klasifikasi diberikan sejumlah *record* yang dinamakan *training set*, yang terdiri dari beberapa atribut, atribut dapat berupa kontinyu ataupun kategoris, salah satu atribut menunjukkan kelas untuk *record*.

Klasifikasi adalah metode data mining yang dapat digunakan untuk proses pencarian sekumpulan model (fungsi) yang dapat menjelaskan dan membedakan kelas-kelas data atau konsep, yang tujuannya supaya model tersebut dapat digunakan memprediksi objek kelas yang labelnya tidak diketahui atau dapat memprediksi kecenderungan data-data yang muncul di masa depan. Metode klasifikasi juga bertujuan untuk melakukan pemetaan data ke dalam kelas yang sudah didefinisikan sebelumnya berdasarkan pada nilai atribut data (Han dan Kamber, 2006).

2.2.4 *Microsoft office excel*

Microsoft excel merupakan program *spreadsheet* yang digunakan untuk mengolah data. Antara lain pengolahan data yang besar, pembuatan dan pengaturan laporan, membuat perhitungan dengan rumus, pembuatan diagram atau *chart*, dan tersedianya berbagai template lembar kerja (Andi, 2013: 1)

Microsoft excel adalah sebuah program aplikasi lembar kerja spreadsheet yang dibuat dan didistribusikan oleh Microsoft Cooperation yang dapat dijalankan pada Microsoft Windows dan Mac OS (Sarwandi, 2017:1)

2.2.5 Promosi

Promosi adalah fungsi memberikan informasi, melakukan persuasi, dan mempengaruhi keputusan untuk membeli (Kurtz, 2010:150).

Promosi merupakan teknik mengomunikasikan informasi mengenai produk dan menjadi bagian dari *bauran komunikasi*, yaitu pesan keseluruhan yang dikirimkan perusahaan kepada pelanggan mengenai produknya (Griffin, 2014:436)

Menurut Laksana (2008:135) pengertian promosi adalah suatu komunikasi dari penjual dan pembeli yang berasal dari informasi yang tepat dan bertujuan untuk merubah sikap dan tingkah laku pembeli, Promosi adalah usaha yang dilakukan pemasar untuk mempengaruhi pihak lain agar berpartisipasi dalam kegiatan pertukaran. Promosi juga merupakan usaha mengkomunikasikan informasi yang bermanfaat tentang suatu perusahaan atau produk dan jasa untuk mempengaruhi pembeli potensial (Dewanti 2008:144)

Promosi juga merupakan kegiatan untuk menyebarluaskan informasi tentang barang dan jasa yang dijual dengan maksud untuk merubah pola perilaku konsumen. Berbagai informasi yang di berikan kepada calon pembeli sangat mempengaruhi keputusan mereka dengan mempertimbangkan uang yang mereka miliki. Mereka akan memasukan barang atau jasa yang mereka ketahui ke dalam daftar pertimbangan, lalu memilih kombinasi yang paling optimal.

2.2.6 Rapidminer

Menurut Aprilla, dkk (2013:8) RapidMiner merupakan perangkat lunak yang bersifat terbuka (*open source*). *RapidMiner* adalah sebuah solusi untuk melakukan analisis terhadap *data mining*, *text mining* dan analisis prediksi. *RapidMiner* menggunakan berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan yang paling baik. *RapidMiner* memiliki kurang lebih 500 operator data mining, termasuk operator untuk input, output, data preprocessing dan visualisasi. *RapidMiner* merupakan software yang berdiri sendiri untuk analisis data dan sebagai mesin data mining yang dapat diintegrasikan pada produknya sendiri. *RapidMiner* ditulis

dengan menggunakan bahasa java sehingga dapat bekerja di semua sistem operasi.

RapidMiner sebelumnya bernama YALE (*Yet Another Learning Environment*), dimana versi awalnya mulai dikembangkan pada tahun 2001 oleh RalfKlinkenberg, Ingo Mierswa, dan Simon Fischer di Artificial Intelligence Unit dari *University of Dortmund*. *RapidMiner* didistribusikan di bawah lisensi AGPL (*GNU Affero General Public License*) versi 3. Hingga saat ini telah ribuan aplikasi yang dikembangkan menggunakan *RapidMiner* di lebih dari 40 negara. *RapidMiner* sebagai *software open source* untuk data mining tidak perlu diragukan lagi karena *software* ini sudah terkemuka di dunia. *RapidMiner* menempati peringkat pertama sebagai *Software* data mining pada polling oleh KDnuggets, sebuah portal *data-mining* pada 2010-2011 (Aprilla, dkk 2013:8)

RapidMiner menyediakan GUI (*Graphic User Interface*) untuk merancang sebuah pipeline analitis. GUI ini akan menghasilkan file XML (*Extensible Markup Language*) yang mendefinisikan proses analitis keinginan pengguna untuk diterapkan ke data. File ini kemudian dibaca oleh *RapidMiner* untuk menjalankan analisis secara otomatis (Aprilla, dkk 2013:9)

RapidMiner memiliki beberapa sifat menurut Aprilla, dkk (2013:10) sebagai berikut:

1. Ditulis dengan bahasa pemrograman Java sehingga dapat dijalankan di berbagai sistem operasi.
2. Proses penemuan pengetahuan dimodelkan sebagai operator trees
3. Representasi XML internal untuk memastikan format standar pertukaran data.

4. Bahasa scripting memungkinkan untuk eksperimen skala besar dan otomatisasi eksperimen.
5. Konsep multi-layer untuk menjamin tampilan data yang efisien dan menjamin penanganan data.

Beberapa Fitur dari *RapidMiner* menurut Aprilla, dkk (2013:10) antara lain:

1. Banyaknya algoritma *data mining*, seperti *decision tree* dan *self-organization map*.
2. Bentuk grafis yang canggih, seperti tumpang tindih diagram histogram, tree chart dan 3D Scatter plots.
3. Banyaknya variasi plugin, seperti text plugin untuk melakukan analisis teks.
4. Menyediakan prosedur data mining dan machine learning termasuk: ETL (*extraction, transformation, loading*), data *preprocessing, visualisasi, modelling* dan *evaluasi*
5. Proses data mining tersusun atas operator-operator yang nestable, dideskripsikan dengan XML, dan dibuat dengan GUI

2.3 Tinjauan Pustaka

Berbagai tinjauan pustaka yang berkaitan dengan penerapan data mining berisi beberapa dari jurnal skripsi dapat dilihat pada tabel.

No	Nama	Judul	Tahun	Isi
1	Dicky Nofriansyah, Kamil Erwansyah, Mukhlis Ramadhan	Penerapan Data Mining dengan Algoritma <i>Naive Bayes Clasifier</i> untuk Mengetahui Minat Beli Pelanggan terhadap Kartu <i>Internet XL</i> (Studi Kasus di CV. Sumber Utama Telekomunikasi)	2016	tujuan penelitian ini adalah untuk memberikan gambaran mengenai minat beli pelanggan terhadap kartu <i>internet XL</i> pada CV. Sumber Utama Telekomunikasi.
2	Saleh, Alfa	Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga	2015	Penelitian ini bertujuan untuk memprediksi penggunaan listrik rumah tangga dengan metode naive bayes dengan menggunakan tools Weka

3	Asrul	Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi)	2016	Penelitian ini akan menggunakan metode algoritma naïve bayes untuk melakukan prediksi peluang kelulusan mahasiswa baru. Kelulusan yang dimaksud adalah diterimanya seorang mahasiswa pada salah satu program studi di Perguruan Tinggi.
4	Elvira Asril dkk	Analisis Data Lulusan dengan Data Mining untuk Mendukung Strategi Promosi Universitas Lancang Kuning	2015	Penelitian ini bertujuan untuk menganalisa dan mengolah data mahasiswa yang telah lulus dari Universitas Lancang Kuning untuk mendapatkan informasi yang penting dan bermanfaat. Dengan pengolahan menggunakan teknik data mining dengan algoritma <i>K-Means clustering</i>

5	Siska haryati, dkk	Implementasi data mining untuk memprediksi masa studi mahasiswa menggunakan algoritma c4.5 (studi kasus: universitas dehasen bengkulu)	2015	Tujuan dari penelitian ini adalah dengan menggunakan pohon keputusan berbasis algoritma c4.5 dan dimplementasikan ke suatu aplikasi yaitu rapidminer diharapkan dapat meningkatkan keakuratan analisa masa studi mahasiswa
6	Nuraeni	Penentuan Kelayakan Kredit Dengan Algoritma Naïve Bayes Classifier: Studi Kasus Bank Mayapada Mitra Usaha Cabang PGC	2017	Penelitian ini akan menggunakan algoritma data mining Naive Bayes Classifier sehingga diharapkan akan menghasilkan akurasi yang lebih tinggi
7	Wahyudi, Deni	Implementasi <i>data mining</i> dengan <i>naive bayes classifier</i> untuk mendukung strategi promosi (studi kasus universitas bina darma Palembang)	2015	Penelitian mengenai Implementasi <i>data mining</i> dengan <i>Naïve Bayes Classifier</i> dalam mendukung strategi promosi dengan menggunakan Tools Rapidminer

Tabel 2.1 Menjelaskan tentang perbedaan dari beberapa penelitian yang telah dilaksanakan sebelumnya, maka perbedaan yang dimiliki dari pengajuan penelitian ini adalah Penerapan Data Mining Dengan Metode *Naive Bayes Classifier* Untuk Mendukung Strategi Promosi Fakultas Sains dan Teknologi UIN Raden Fatah Palembang

BAB III

METODOLOGI PENELITIAN

3.1 Gambaran Umum Lokasi Penelitian

3.1.1 Sejarah Fakultas Saintek dan Teknologi

Perkembangan pendidikan keagamaan di Sumatera Selatan, khususnya perkembangan Perguruan Tinggi Islam mengalami perubahan transformasi. Perguruan Tinggi keagamaan yang ada di Palembang dulu berstatus Institut Agama Islam Negeri (IAIN) Raden Fatah Palembang yang berdiri sejak 13 Nopember 1964 meliputi program studi keagamaan. Dari tahun 1964 sampai dengan 2014, Institut Agama Islam Negeri masih memiliki 6 Fakultas yaitu Fakultas Syariah dan Hukum, Ilmu Tarbiyah dan Keguruan, Ushuludin dan Pemikiran Islam, Adab dan Humaniora, Dakwah dan Komunikasi Islam serta Ekonomi dan Bisnis Islam. Dalam rangka memenuhi tuntutan perkembangan ilmu pengetahuan dan teknologi, dan proses integrasi ilmu agama Islam dengan ilmu lain serta mewujudkan sumber daya manusia yang berkualitas, perlu dibentuk suatu perguruan tinggi universal. Oleh karena itu Institut Agama Islam Negeri Raden Fatah perlu mencari solusi untuk menghadapi perubahan tersebut.

Perkembangan selanjutnya adalah transformasi keilmuan, yaitu menyatukan keilmuan umum dan agama dalam suatu universitas. Tantangan ini terjawab dengan dikeluarkannya peraturan presiden Nomor 29 tahun 2014 tertanggal 17 Oktober 2014, tentang perubahan status dari Institut Agama Islam Negeri menjadi Universitas Islam Negeri Raden Fatah.

Berdasarkan peraturan tersebut, maka disusunlah organisasi tata kerja Universitas Islam Negeri Raden Fatah yang termaktub dalam statuta Universitas Islam Negeri Raden Fatah Palembang untuk menambah Fakultas baru, salah satunya Fakultas Sains dan Teknologi.

Fakultas Sains dan teknologi merupakan Fakultas baru yang cikal bakalnya berada di Fakultas Ilmu tarbiyah dan Keguruan. Berdasarkan surat Peraturan Menteri Agama RI nomor 53 tahun 2015, pasal 10 tertanggal 2 September 2015 yang isinya salah satu Fakultas yang ada di Universitas Islam Negeri Raden Fatah adalah Fakultas Sains dan Teknologi. Awal berdirinya Fakultas Sains dan Teknologi memiliki 2 Program Studi yaitu Program studi Biologi dan Kimia, ini berdasarkan surat dari Keputusan Menteri Pendidikan dan Kebudayaan Nomor 273A/P/2014 tertanggal 9 Oktober 2014 tentang izin prodi-prodi di Universitas Islam Negeri Raden Fatah Palembang.

Untuk mengelola Fakultas baru, maka dipilih dan di lantiklah kepengurusan Dekan dan Wakil Dekan berdasarkan keputusan Rektor untuk Fakultas Sains dan Teknologi unsur pimpinan Fakultas pada awal terpilih adalah Dr. Dian Erlina, M.Hum, (Dekan), Dr. Irham Falahudin, M.Si. (Wakil dekan Bidang Akademik dan Kelembagaan), Gusmelia Testiana, M.Kom (Wakil Dekan II Bidang Administrasi. Keuangan dan Kepegawaian) dan Wakil Dekan III Syarifah S.Si, M.Kes (Bidang Kemahasiswaan dan Kerjasama). Sejak tahun 2015 telah melakukan operasional yaitu menerima mahasiswa baru untuk Progam Studi Biologi dan Kimia.

3.1.2 Visi dan Misi Fakultas Sains Teknologi

3.1.2.1 Visi Fakultas Sains Teknologi

“Unggul dalam bidang Sains dan Teknologi, berkarakter islami dan berstandar internasional pada tahun 2025”.

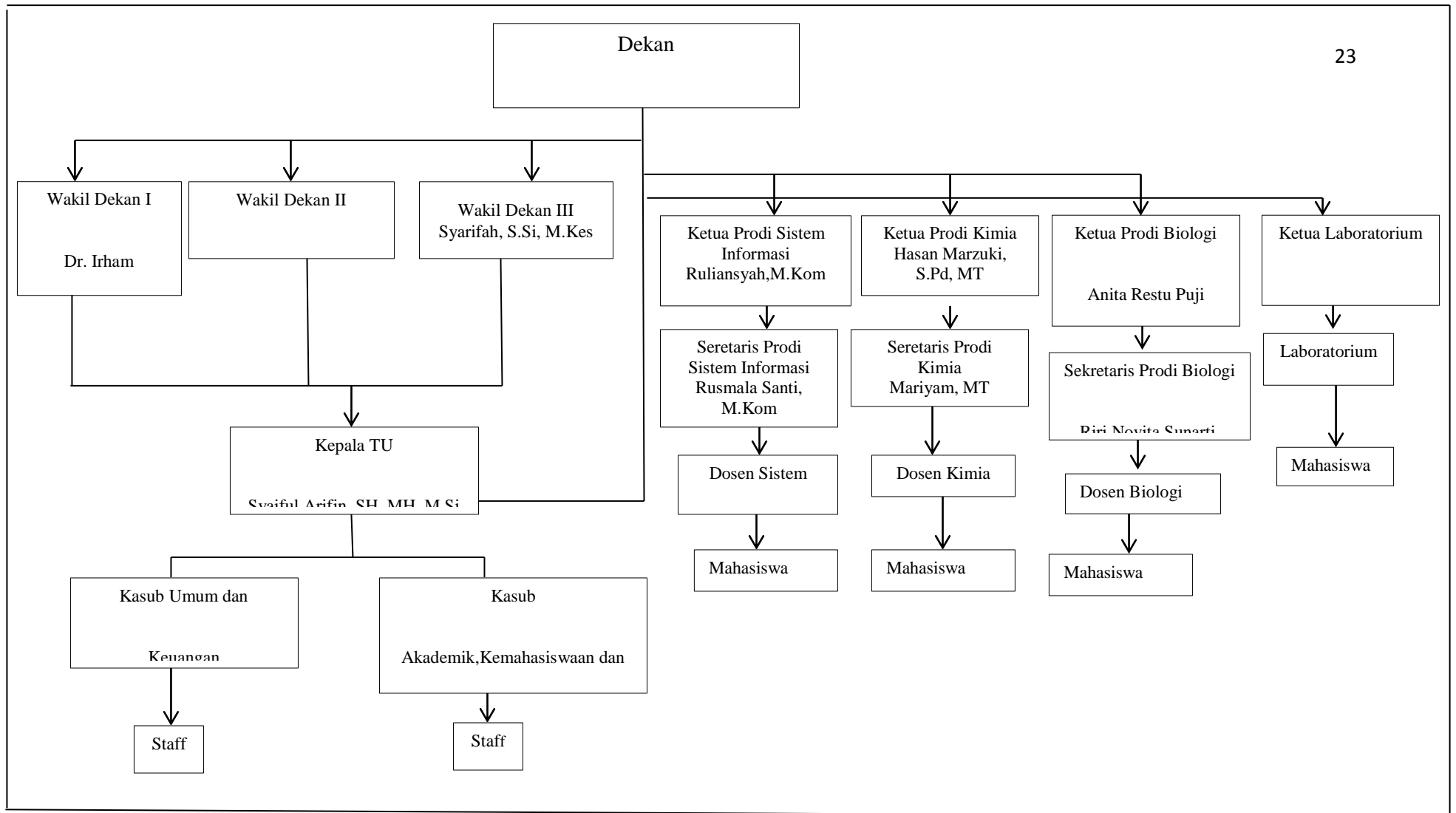
3.1.2.2 Misi Fakultas Sains Teknologi

1. Pendidikan dan Pengajaran :

- a. Menyelenggarakan pendidikan dan pengajaran yang bermutu tinggi untuk menghasilkan lulusan yang unggul dalam bidang sains dan teknologi;
- b. Menyelenggarakan tata kelola fakultas yang bermutu tinggi berdasarkan prinsip prinsip good governance

2. Penelitian : Mengembangkan penelitian dalam bidang sains dan teknologi yang bermutu tinggi.

3. Pengabdian : Melakukan pengabdian masyarakat berbasis riset sains dan teknologi dengan mempertimbangkan local wisdom



Gambar 3.1 Struktur Organisasi Fakultas Sains dan Teknologi

Berdasarkan PMA No. 53 tahun 2015, tentang organisasi tata kerja UIN Raden Fatah Palembang maka disusunlah tata pamong organisasi Fakultas Sains dan teknologi. Adapun unsur pimpinan pada Fakultas Sains dan teknologi adalah sebagai berikut:

Pimpinan

Dekan	: Dr. Dian Erlina, M.Hum
Wakil Dekan Bid. Akademik dan Kelembagaan	: Dr. Irham Falahudin, M.Si
Wakil Dekan Bid. Adm. Umum, Perencanaan dan keuangan	: Gusmelia Testiana, M.Kom
Wakil Dekan Bid. Kemahasiswaan dan Kerjasama	: Syarifah, S.Si, M.Kes

Unsur Tenaga Kependidikan Bagian Tata Usaha

Kepala Bagian Tata Usaha	: Syaiful Arifin, SH., MH., MSi
Kasubbag Adm. Umum dan Keuangan	: Dra. Sri Mulyati
Kasubbag Akademik, Kemahasiswaan dan Alumni	: Marsi Sulistiawati, M.Kom
Ketua Prodi Biologi	: Anita Restu Puji Restu, M.Si, BioMed,Sc
Ketua Prodi Kimia	: Hasan Marzuki, S.Pd, M.T.
Ketua Prodi Sistem Informasi	: Ruliansyah, M.Kom
Ketua Laboratorium MIPA	: Ahmad Zaky, S.Si.
Ketua Laboratorium Biota	: Mariyamah, M.T.
Subbag. Akademik, Kemahasiswaan dan Alumni	: Ahmad Zaky, S.Si.

Subbag. Akademik, Kemahasiswaan dan Alumni : Erpani

Subbag. Akademik, Kemahasiswaan dan Alumni : Ir. Ledis Heru Saryono Putro, M.Si

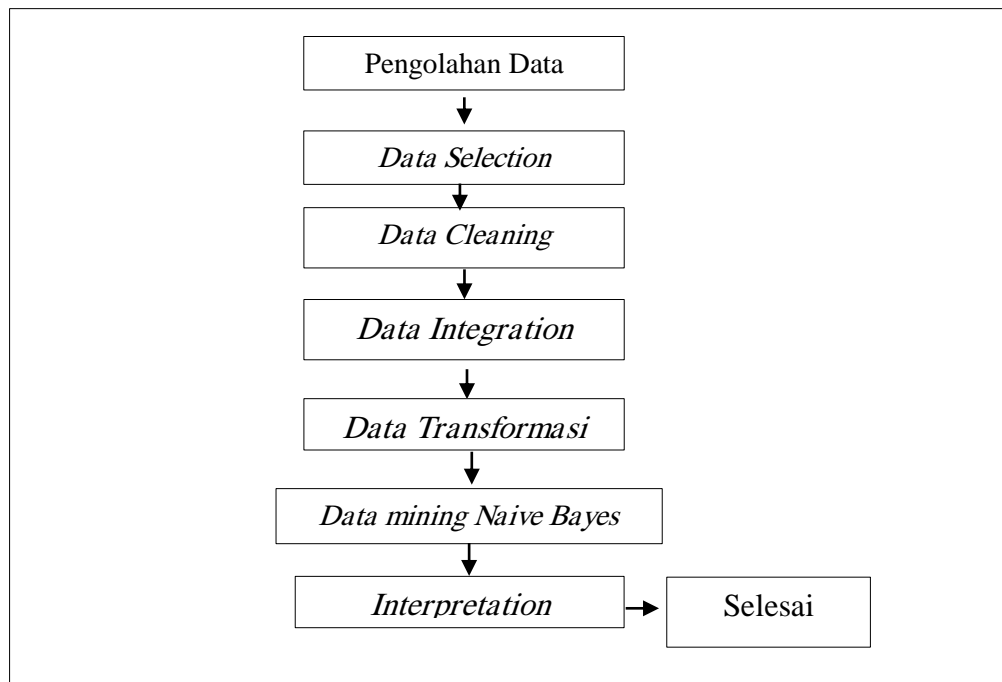
Subbag. Teknologi Informasi : M. Leandry Dalafranka, S.SI

Subbag. Umum dan Keuangan : Ruslan

Subbag. Umum dan Keuangan : dr. Hj. Delia Yusfarani, M.Kes.

Bendahara : Saniah, SE

3.2 Kerangka Berfikir



Gambar 3.2 Kerangka Berfikir

3.3 Data Selection

Pada tahap ini data yang digunakan akan diseleksi dengan cara melihat kecenderungan data, kesesuaian data judul penelitian yang akan diteliti oleh penulis, semua atribut yang ada pada data terdiri dari atribut Tahun, Nama, Alamat, Asal_Sekolah, Program Studi dan Status

Tabel 3.1 Data Penerimaan Mahasiswa Baru dari 3229 data pendaftar

TAHUN	NAMA	ALAMAT	ASAL SEKOLAH	PROGRAM STUDI	STATUS
2013	BERTA ERAWANTI	KABUPATEN PALI	SMA	BIOLOGI	TIDAK LULUS
2014	NINI SUMARNI	KOTA PAGAR ALAM	SMA	SISTEM INFORMASI	LULUS
2015	RETNO DWI LESTARI	KABUPATEN BANYUASIN	SMA	KIMIA	DAFTAR

3.4 Data Pre Processing atau Data Cleaning

Pada tahap ini akan dilakukan pembersihan data, yakni membuang data yang tidak konsisten dan *noise* atau *redundancy data*. Pada data yang didapat, terdapat beberapa data yang tidak konsisten dan *noise*, berikut kesalahan pada yang dimaksud, sample data yang tidak konsisten tersebut dapat di lihat pada gambar dibawah :

Tabel 3.2 Proses Data Cleaning Pada Atribut Alamat dan Asal Sekolah dari 5026

TAHUN	NAMA	ALAMAT	ASAL SEKOLAH	PROGRAM STUDI	STATUS
2013	BERTA ERAWANTI	NULL	NULL	BIOLOGI	LULUS
2014	NINI SUMARNI	NULL	NULL	SISTEM INFORMASI	TIDAK LULUS
2015	BAGUS DERMAWAN	NULL	NULL	KIMI	LULUS
2016	AYU LESTARI	NULL	NULL	BIOLOGI	TIDAK LULUS

Pada data alamat dan asal sekolah dapat dilihat adanya data yang kosong sehingga data tersebut harus di buang agar pada saat proses klasifikasi dapat dilakukan perhitungan untuk proses *data mining*.

3.5 Data Integration

Tahap *integrasi* data adalah tahap penggabungan data. Data pendaftar Universitas Islam Negeri Raden Fatah Palembang khususnya Fakultas Sains dan Teknologi berasal dari satu sumber yaitu PUSTIPD UIN Raden Fatah Palembang dan dari data yang ada tersebut dapat digunakan untuk proses pengolahan *data mining* dengan menggunakan metode *Naïve Bayes*. Data ini terdiri dari Calon pendaftar Program Studi Sistem informasi dari tahun 2013 sampai dengan tahun 2017, Calon Mahasiswa pendaftar Program Studi Biologi pada tahun 2015 sampai

dengan 2017, Calon Mahasiswa pendaftar Program Studi Kimia pada tahun 2015 sampai dengan 2017 yang kemudian digabungkan kedalam satu *datase*.

Tabel 3.3 Gabungan data dari 5026

TAHUN	NAMA	ALAMAT	ASAL SEKOLAH	PROGRAM STUDI	STATUS
2013	BERTA ERAWANTI	KOTA PRABUMULIH	MAN	BIOLOGI	LULUS
2014	NINI SUMARNI	KOTA PAGAR ALAM	SMA	SISTEM INFORMASI	LULUS
2015	SYARIF	KOTA PAGAR ALAM	SMK	SISTEM INFORMASI	TIDAK LULUS
2016	DWI LESTARI	KABUPATEN OKUS	SMA	SISTEM INFORMASI	LULUS
2017	RETNO	KABUPATEN BANYUASIN	MAN	SISTEM INFORMASI	LULUS

3.6 Data Transformasi.

Ada pun pada tahap ini data akan diubah menjadi bentuk yang sesuai untuk proses *data mining*. Karena dalam penelitian ini akan dilakukan uji coba secara teoritis dan menggunakan *software data mining* yaitu *RapidMiner*, maka data yang telah melalui proses sebelumnya akan di *transformasi* agar dapat sesuai dengan algoritma yang dipakai yaitu algoritma *Naïve Bayes*.

Pada tahapan ini atribut yang di pakai akan diberi label mengikuti kondisi data - data pada atribut tersebut :

1. Klasifikasi Atribut Alamat

Berdasarkan hasil data set yang siap untuk di olah terdapat 56 daerah asal dari calon mahasiswa yang mendaftar di Fakultas Sains dan Teknologi UIN Raden Fatah pada tahun 2013 – 2017.

Tabel 3.4 Klasifikasi Alamat

NO	ALAMAT
1	KABUPATEN EMPAT LAWANG
2	KABUPATEN BANYUASIN
3	KABUPATEN OGAN KOMERING ULU SELATAN
4	KABUPATEN ACEH BARAT DAYA
5	KABUPATEN ACEH TENGGARA
6	KABUPATEN AGAM
7	KABUPATEN BANDUNG
8	KABUPATEN BANGKA
9	KABUPATEN BANGKA BARAT
10	KABUPATEN BANGKA SELATAN
11	KABUPATEN CIAMIS
12	KABUPATEN JOMBANG
13	KABUPATEN KUDUS
14	KABUPATEN LAMPUNG BARAT
15	KABUPATEN LAMPUNG SELATAN
16	KABUPATEN LAMPUNG TENGAH
17	KABUPATEN LAMPUNG UTARA
18	KABUPATEN LEBAK
19	KABUPATEN LAHAT
20	KABUPATEN MUARA ENIM
21	KABUPATEN MUSI BANYUASIN
22	KABUPATEN MUSI RAWAS
23	KABUPATEN MUSI RAWAS UTARA
24	KABUPATEN OGAN ILIR
25	KABUPATEN OGAN KOMERING ILIR
26	KABUPATEN OGAN KOMERING ULU
27	KABUPATEN OGAN KOMERING ULU TIMUR
28	KABUPATEN PANDEGLANG
29	KABUPATEN PANGKAL PINANG
30	KABUPATEN PATI
31	KABUPATEN PASAMAN
32	KABUPATEN PENUNGGAL ABAB LEMATANG ILIR
33	KABUPATEN PESAWARAN
34	KABUPATEN PONOROGO
35	KABUPATEN PRINGSEWU

36	KABUPATEN SOLOK SELATAN
37	KABUPATEN SUKABUMI
38	KABUPATEN SUMENEP
39	KABUPATEN WAY KANAN
40	KOTA BOGOR
41	KOTA BUKIT TINGGI
42	KOTA GUNUNGSITOLI
43	KOTA JAKARTA BARAT
44	KOTA JAKARTA TIMUR
45	KOTA JAMBI
46	KOTA LANGSA
47	KOTA LUBUK LINGGAU
48	KOTA MAKASAR
49	KOTA PADANG
50	KOTA PARIAMAN
51	KOTA PAYAKUMBUH
52	KOTA PEMATANG SIANTAR
53	KOTA SABANG
54	KOTA PAGAR ALAM
55	KOTA PALEMBANG
56	KOTA PRABUMULIH

2. Klasifikasi Atribut Asal Sekolah

Dikelompokan berdasarkan SMA, SMK, MADRASAH

Tabel 3.5 Tabel Klasifikasi Asal Sekolah

No	Klasifikasi Asal Sekolah
1	SMA
2	SMK
3	MA

3. Klasifikasi Atribut Progd

Pada proses klasifikasi program studi akan dilakukan berdasarkan program studi yang ada pada masing – masing fakultas yang ada di lingkungan UIN Rafah Palembang

Tabel 3.6 Klasifikasi Atribut Program Studi

No	Klasifikasi Atribut Program Studi
1	Sistem Informasi
2	Biologi
3	Kimia

4. Klasifikasi Status

Pada tahapan ini terdapat 2 status yaitu Lulus dan Tidak Lulus

Tabel 3.7 Klasifikasi Status

No	Klasifikasi Status
1	Lulus
2	Tidak Lulus

3.7 Data Mining

Tahap ini merupakan proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu berdasarkan proses KDD secara keseluruhan.

Data yang akan digunakan pada proses *data mining* menggunakan Algoritma *Naïve Bayes* merupakan data pendaftar yang merupakan gambaran secara nyata mengenai penerimaan mahasiswa baru pada Fakultas Sains dan teknologi UIN Raden Fatah Palembang.

Adapun tahapan pengerjaan metode *Naïve Bayes* menurut Pratiwi (2015:101) dalam penelitian ini adalah sebagai berikut :

1. Menentukan kriteria apa saja yang digunakan, kriteria sangat penting dalam melakukan perhitungan *Naïve Bayes*, karena digunakan untuk mengumpulkan data
2. Menyiapkan data yang digunakan untuk melakukan perhitungan *Naïve Bayes*
3. Membuat tabel aturan, hanya dengan batasan aturan dan data pada tabel aturan diambil dari data yang ada
4. Membuat tabel probabilitas kemunculan setiap atribut dari semua kriteria yang ada
5. Menghitung nilai likelihood ya dan likelihood tidak diambil dari tabel probabilitas kemunculan setiap nilai atribut
6. Menghitung nilai probabilitas dimana dapat dihitung dengan melakukan normalisasi ini bias mengetahui hasil akhir dari perhitungan dengan menggunakan metode *Naïve Bayes* layak atau tidak layak berdasarkan dari probabilitas ya atau tidak

3.8 Interpretation atau Evaluasi

Pola informasi yang dihasilkan dari proses data mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

BAB IV HASIL DAN PEMBAHASAN

4.1 HASIL

Setelah melakukan analisa terhadap perancangan dengan tahapan data mining untuk menentukan strategi promosi pada Fakultas Sains dan Teknologi Universitas Islam Negeri Raden Fatah Palembang dengan menggunakan *Algoritma Naive Bayes*, analisa ini berakhir dengan melakukan proses *data mining* yang sesungguhnya, maka hasil yang dicapai oleh penulis adalah untuk mengetahui data calon pendaftar berdasarkan data atribut dari Alamat, Asal Sekolah dan Program Studi. Peneliti menggunakan aplikasi *Rapidminer* untuk mempermudah proses data mining yang kemudian menghasilkan informasi data pendaftar.

4.2 PEMBAHASAN

4.2.1 *Naive Bayes*

Berdasarkan dari tahapan data mining untuk algoritma *Naive Bayes*, adapun hasil dari langkah-langkah *Naive Bayes Classifier* :

1. Menentukan kriteria apa saja yang digunakan, kriteria sangat penting dalam melakukan perhitungan *Naive Bayes*, dapat dilihat pada tabel 4.1 dibawah ini :

Tabel 4.1 Atribut Pendaftar

No	Atribut	Keterangan
1	Tahun Penerimaan	Tahun Penerimaan Mahasiswa
2	Nama	Nama Mahasiswa
3	Alamat	Alamat Mahasiswa

4	Asal Sekolah	Asal Sekolah Mahasiswa
5	Program Studi	Pilihan Program Studi
6	Status	Status Penerimaan Calon Mahasiswa

2. Menyiapkan data yang digunakan untuk melakukan perhitungan *Naïve Bayes* dapat dilihat pada tabel 4.2 dibawah ini :

Tabel 4.2 Data Pendaftar Calon Mahasiswa setelah melalui tahapan data mining (7 dari 3296)

TAHUN	NAMA	ALAMAT	ASAL SEKOLAH	PROGRAM STUDI	STATUS
2013	DWI SEPTI ANGGRAINI	KOTA PALEMBANG	SMA	SISTEM INFORMASI	LULUS
2013	ELIZA	KOTA PALEMBANG	SMK	SISTEM INFORMASI	TIDAK LULUS
2013	EMILIA	KOTA PALEMBANG	SMK	SISTEM INFORMASI	TIDAK LULUS
2013	WINDA DWI LESTARI	KOTA PALEMBANG	SMK	SISTEM INFORMASI	LULUS
2013	WIWIN SULASTRI	KOTA PALEMBANG	SMK	SISTEM INFORMASI	LULUS
2013	SERI UNTARI	KOTA PALEMBANG	SMK	SISTEM INFORMASI	LULUS
2013	AGUNG PRANDIKO	KOTA PALEMBANG	SMK	SISTEM INFORMASI	LULUS

3. Membuat Atribut status atribut status terdapat 2 kriteria status yaitu Lulus dan Tidak Lulus. Pada label status akan di jadikan sebagai proses untuk melakukan proses eksekusi pada program *Rapidminer* dapat dilihat pada tabel 4.3 dibawah ini:

Tabel 4.3 Status

STATUS
LULUS
TIDAK LULUS

4. Membuat tabel probabilitas kemunculan setiap atribut dari semua kriteria yang ada pada tabel 4.4 dibawah ini :

Tabel 4.4 Data Klasifikasi Penerimaan Mahasiswa Baru

NO	ALAMAT	LULUS	TIDAK LULUS
1	KABUPATEN EMPAT LAWANG	22	63
2	KABUPATEN BANYUASIN	161	142
3	KABUPATEN OGAN KOMERING ULU SELATAN	24	100
4	KABUPATEN ACEH BARAT DAYA	1	1
5	KABUPATEN ACEH TENGGARA	1	3
6	KABUPATEN AGAM	5	9
7	KABUPATEN BANDUNG	1	2
8	KABUPATEN BANGKA	2	15
9	KABUPATEN BANGKA BARAT	2	7
10	KABUPATEN BANGKA SELATAN	1	20
11	KABUPATEN CIAMIS	1	7
12	KABUPATEN JOMBANG	2	6
13	KABUPATEN KUDUS	1	1
14	KABUPATEN LAMPUNG BARAT	1	5
15	KABUPATEN LAMPUNG SELATAN	2	2
16	KABUPATEN LAMPUNG TENGAH	2	5
17	KABUPATEN LAMPUNG UTARA	3	8
18	KABUPATEN LEBAK	1	3
19	KABUPATEN LAHAT	30	115
20	KABUPATEN MUARA ENIM	62	84
21	KABUPATEN MUSI BANYUASIN	84	97
22	KABUPATEN MUSI RAWAS	13	60
23	KABUPATEN MUSI RAWAS UTARA	9	21
24	KABUPATEN OGAN ILIR	68	61
25	KABUPATEN OGAN KOMERING ILIR	97	86
26	KABUPATEN OGAN KOMERING ULU	37	96
27	KABUPATEN OGAN KOMERING ULU TIMUR	52	111
28	KABUPATEN PANDEGLANG	1	1
29	KABUPATEN PANGKAL PINANG	1	30
30	KABUPATEN PATI	1	1
31	KABUPATEN PASAMAN	1	1

32	KABUPATEN PENUNGGAL ABAB LEMATANG ILIR	28	51
33	KABUPATEN PESAWARAN	3	17
34	KABUPATEN PONOROGO	5	23
35	KABUPATEN PRINGSEWU	2	5
36	KABUPATEN SOLOK SELATAN	2	1
37	KABUPATEN SUKABUMI	2	1
38	KABUPATEN SUMENEP	1	3
39	KABUPATEN WAY KANAN	1	2
40	KOTA BOGOR	1	3
41	KOTA BUKIT TINGGI	2	4
42	KOTA GUNUNGSITOLI	1	3
43	KOTA JAKARTA BARAT	1	1
44	KOTA JAKARTA TIMUR	1	1
45	KOTA JAMBI	2	3
46	KOTA LANGSA	1	1
47	KOTA LUBUK LINGGAU	14	61
48	KOTA MAKASAR	1	23
49	KOTA PADANG	1	1
50	KOTA PARIAMAN	1	1
51	KOTA PAYAKUMBUH	1	3
52	KOTA PEMATANG SIANTAR	2	7
53	KOTA SABANG	1	8
54	KOTA PAGAR ALAM	8	54
55	KOTA PALEMBANG	535	396
56	KOTA PRABUMULIH	33	121
ASAL SEKOLAH			
SMA		871	1114
SMK		183	346
MAN		285	497
PROGRAM STUDI			
BIOLOGI		184	371
KIMIA		166	163
SISTEM INFORMASI		989	1423
STATUS		1339	1957

5. Menghitung nilai likelihood ya dan likelihood tidak diambil dari tabel probabilitas kemunculan setiap nilai atribut, dengan *Naïve Bayes* dengan data sejumlah 3296 record. Perhitungan pemilihan Lulus dan Tidak Lulus. Dalam data terdapat 1339 Lulus dan 1957 Tidak Lulus dengan menggunakan rumus dibawah ini:

$$P(H|X) = \frac{P(X|H) \times P(H)}{P(H)}$$

$$P(H|X) = (P(X|H) \times P(H))$$

$$P(\text{LULUS}) = 1339/3296 = 0.406$$

$$P(\text{TIDAK LULUS}) = 1957/3296 = 0.594$$

6. Menghitung nilai probabilitas dimana dapat dihitung dengan melakukan normalisasi terhadap *Likelihood* dan dengan menghitung nilai probabilitas ini bisa mengetahui hasil akhir dari perhitungan dengan menggunakan metode *Naïve Bayes* Lulus atau Tidak Lulus :

$$P(H|X) = \frac{P(X|H) \times P(H)}{P(H)}$$

Dimana:

X : Data dengan kelas yang belum diketahui

H : Hipotesis data X merupakan suatu kelas spesifik

P(H|X): Probabilitas hipotesis H berdasar kondisi X (*posterior probability*)

P(H) : Probabilitas hipotesis H (*prior probability*)

$P(X|H)$: Probabilitas X berdasarkan kondisi pada hipotesis H

$P(X)$: Probabilitas X

Proses pengujian 1

Diketahui :

X_1 = Alamat

X_2 = Asal Sekolah

X_3 = Program Studi

Y = Status

Jika Diketahui Alamat = Empat Lawang , Asal Sekolah = MAN, Program Studi = Kimia.

$$P(\text{LULUS}) = 1339/3296 = 0.406$$

$$P(\text{TIDAK LULUS}) = 1957/3296 = 0.593$$

$$P(\text{Empat Lawang} | \text{LULUS}) = 22 / 1339 = 0.016$$

$$P(\text{Empat Lawang} | \text{TIDAK LULUS}) = 63 / 1957 = 0.032$$

$$P(\text{MAN} | \text{LULUS}) = 285 / 1339 = 0.212$$

$$P(\text{MAN} | \text{TIDAK LULUS}) = 497 / 1957 = 0.253$$

$$P(\text{Kimia} | \text{LULUS}) = 166 / 1339 = 0.123$$

$$P(\text{Kimia} | \text{TIDAK LULUS}) = 163 / 1957 = 0.083$$

Kesimpulan :

$$P(\text{LULUS}) = (22/1339) * (285/1339) * (166/1339) * (1339/3296) = 0.00016$$

$$P(\text{TIDAK LULUS}) = (63/1957) * (497/1957) * (163/1957) * (1957/3296) = 0.00039$$

Kesimpulan $P(\text{LULUS}) < P(\text{TIDAK LULUS})$ sehingga Calon

Mahasiswa dengan kriteria tersebut dapat di prediksi **Tidak Lulus**

Proses Pengujian 2

Jika Diketahui Alamat = Palembang, Asal Sekolah = SMA, Progdi = Sistem Informasi.

$$P(\text{LULUS}) = 1339/3296 = 0.406$$

$$P(\text{TIDAK LULUS}) = 1957/3296 = 0.593$$

Pengujian

$$P(\text{Palembang} | \text{LULUS}) = 535 / 1339 = 0,399$$

$$P(\text{Palembang} | \text{TIDAK LULUS}) = 396 / 1957 = 0.202$$

$$P(\text{SMA} | \text{LULUS}) = 871 / 1339 = 0.650$$

$$P(\text{SMA} | \text{TIDAK LULUS}) = 1114 / 1957 = 0.569$$

$$P(\text{Sistem Informasi} | \text{LULUS}) = 989 / 1339 = 0.738$$

$$P(\text{Sistem Informasi} | \text{TIDAK LULUS}) = 1423/ 1957 = 0.727$$

Kesimpulan :

$$P(\text{LULUS}) = (535/1339)*(871 / 1339)* (989/1339)*(1339/3296) = 0.077$$

$$P(\text{TIDAK LULUS}) = (396/ 1957)*(1114/ 1957)* (1423/1957)*$$

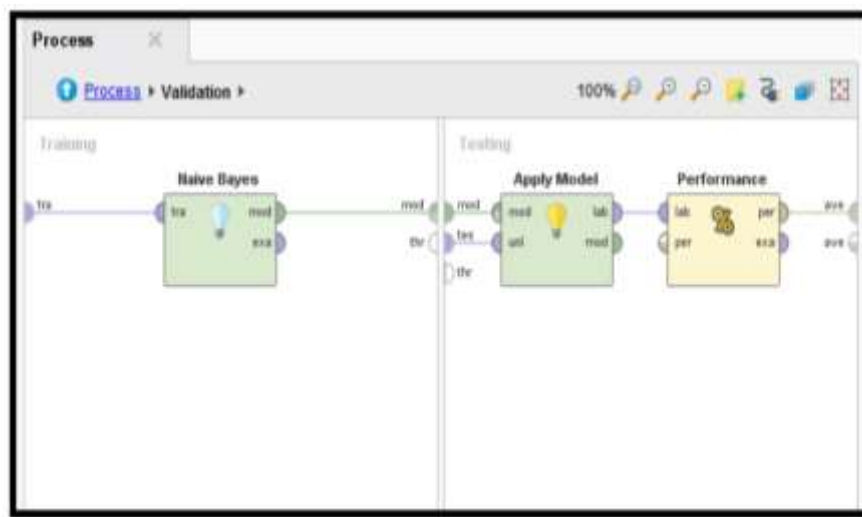
$$(1339/3296) = 0.049$$

Kesimpulan $P(\text{LULUS}) > P(\text{TIDAK LULUS})$ sehingga Calon Mahasiswa dengan

kriteria tersebut dapat di prediksi **LULUS**.

4.2.2. Rapidminer

Rapidminer merupakan salah satu *software data mining* pengolahan data set untuk mencari pola data sesuai dengan tujuan dari pengolahan data tersebut, tidak semua algoritma yang ada dapat sesuai atau dapat mengolah data set yang ada, harus dilakukan penyesuaian pola data dan sesuai dengan tujuan dari pengolahan data tersebut.



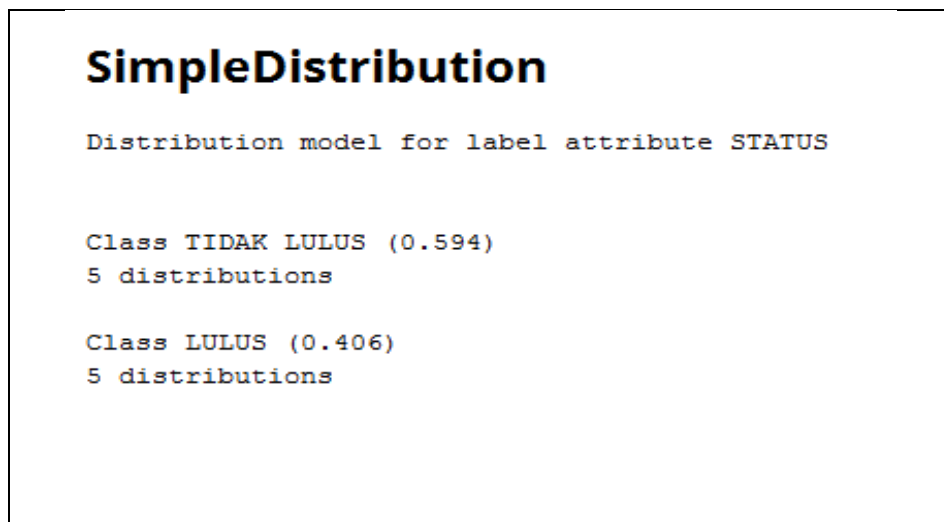
Gambar 4.1. Model Sub Proses Pada *Cross Validation Rapid Miner*

Adapun hasil *Accuracy* dari *performance vector* sebesar 62.69 % untuk proses yang di laksanakan pada model diatas dapat dilihat pada Gambar 4.2 dibawah ini:

accuracy: 62.69%			
	true TIDAK LULUS	true LULUS	class precision
pred. TIDAK LULUS	337	119	73.90%
pred. LULUS	250	283	53.10%
class recall	57.41%	70.40%	

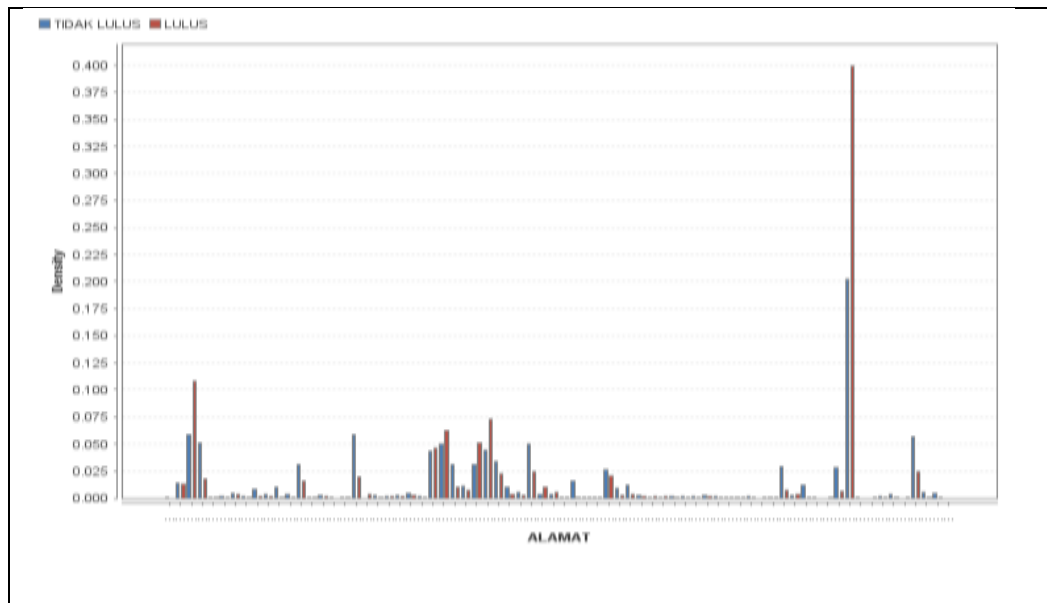
Gambar 4.2 Hasil *Performance Vector*

Berdasarkan hasil dari perhitungan dengan *Rapidminer* dengan model *Naive Bayes* maka nilai dari *simple Distribution* yang didapat adalah seperti pada gambar di bawah, yang hasilnya nilai class Tidak Lulus 0,594 dan nilai Class Lulus 0,406. Dimana perhitungan manual dengan perhitungan pada *Rapidminer* hasilnya sama.

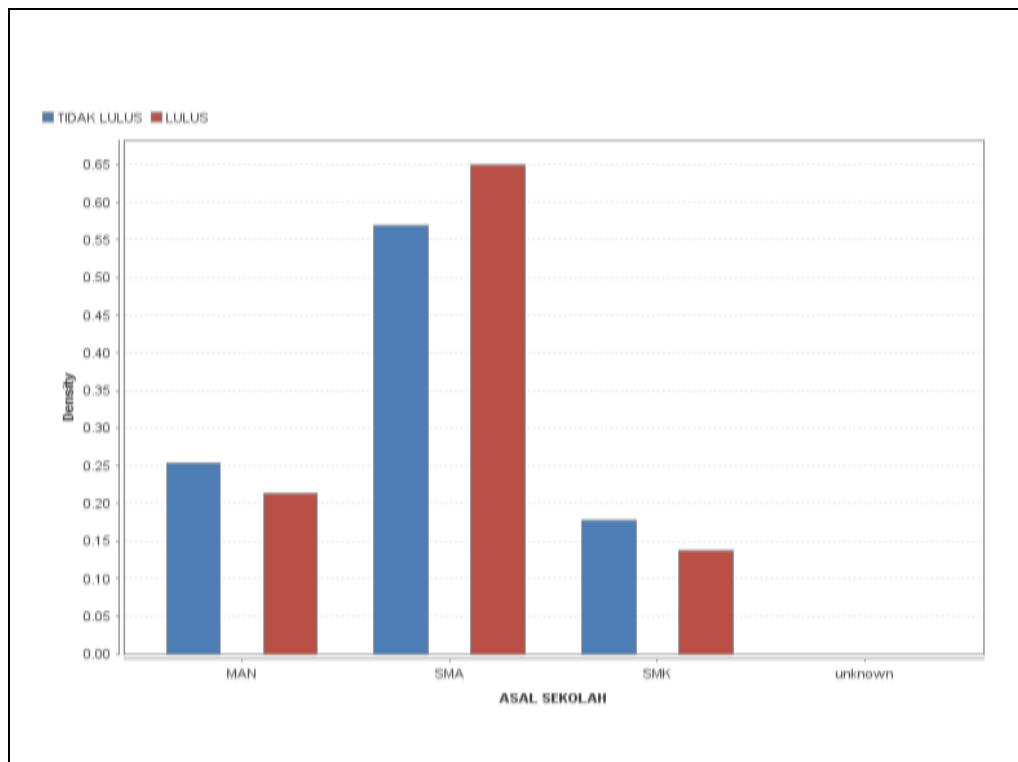


Gambar 4.3 Hasil *Simple Distribution*

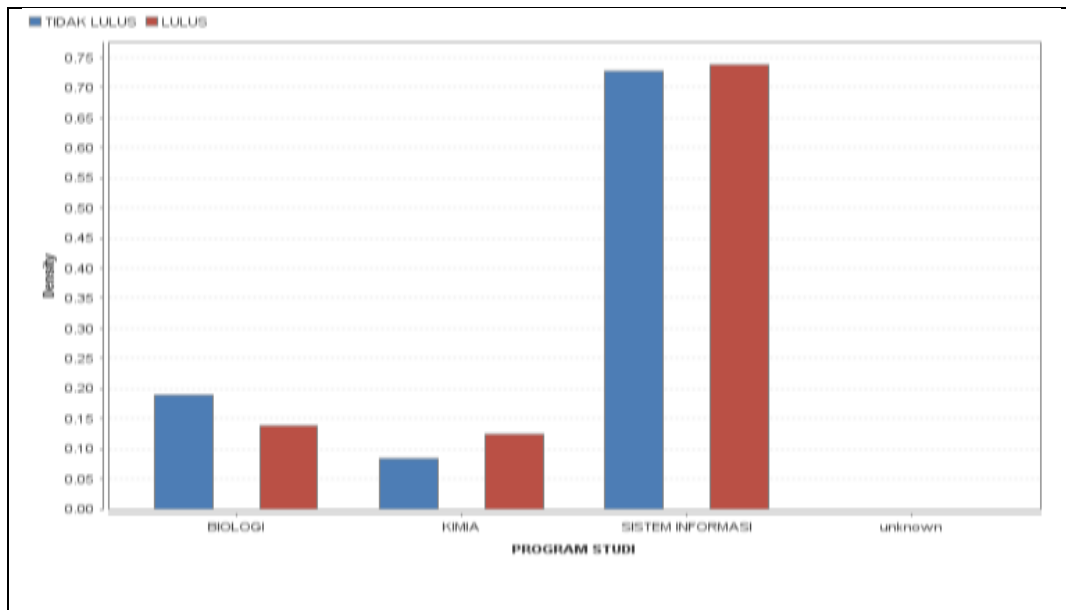
Pada gambar grafik dibawah dapat menunjukkan hasil dari tingkat rendah dan tingginya pendaftar di Fakultas Sains dan Teknologi UIN Raden Fatah Palembang berdasarkan Alamat, Asal Sekolah dan Program Studi.



Gambar 4.4 Grafik Berdasarkan Alamat



Gambar 4.5. Grafik Berdasarkan Asal Sekolah



Gambar 4.6. Grafik Berdasarkan Program Studi

Berdasarkan data hasil dan pembahasan dapat kita jadikan sebagai acuan untuk menentukan strategi promosi pada Fakultas Sains dan Teknologi Universitas Islam Negeri Raden Fatah Palembang dengan melihat data atribut dari Alamat, Asal Sekolah dan Program Studi. Dimana dari data Penerimaan Mahasiswa Baru pada Fakultas Sains dan Teknologi pada tahun 2013 sampai dengan 2017 dengan jumlah data 3296 yang Lulus 1339 dan Tidak Lulus 1957. Dari atribut alamat data mahasiswa yang mendaftar tertinggi adalah beralamat Kota Palembang dengan jumlah pendaftar 931, yang mendaftar berdasarkan tingkat sekolah asal adalah SMA dengan jumlah pendaftar 871 dan pilihan jurusan yang tertinggi adalah Program Studi Sistem Informasi dengan jumlah 989 dan peminat yang paling rendah khususnya provinsi sumatera selatan adalah Kabupaten Musi Rawas Utara dengan jumlah data 30, yang Lulus 9 dan Tidak Lulus 21.

Jadi dalam menentukan strategi promosi yang harus kita lihat adalah dari alamat dan asal sekolah yang tingkat pendaftar rendah, dan yang lebih harus ditonjolkan di Fakultas Sains dan Teknologi adalah Program Studi yang pendaftarnya rendah, Sehingga pendaftar akan tertarik dengan Fakultas Sains dan Teknologi Universitas Islam Negeri Raden Fatah Palembang.

BAB V

PENUTUP

5.1 Kesimpulan

Pada kesimpulan dari penelitian mengenai penerapan *data mining* dengan metode *Naïve Bayes Classifier* dalam mendukung strategi promosi Fakultas Sains dan Teknologi UIN Raden Fatah Palembang dalam memprediksi calon pendaftar pada tahun yang akan datang dengan memanfaatkan data pendaftar pada Fakultas Sains dan Teknologi UIN Raden Fatah Palembang menggunakan data set yang sudah melalui tahap awal *Knowledge Discovery in Database (KDD)* sebanyak 3296 dengan Lulus 1339 dan Tidak Lulus 1957.

Adapun yang paling banyak peminatnya berdasarkan atribut alamat adalah beralamat Kota Palembang dengan jumlah pendaftar 949, yang mendaftar berdasarkan tingkat sekolah asal adalah SMA dengan jumlah pendaftar 871 dan pilihan jurusan yang tertinggi adalah Program Studi Sistem Informasi dengan jumlah 989 dan peminat yang paling rendah khususnya provinsi Sumatera Selatan adalah Kabupaten Musi Rawas Utara dengan jumlah data 30 yang Lulus 9 dan Tidak Lulus 21. Jadi dalam menentukan strategi promosi pada Fakultas Sains dan Teknologi UIN Raden Fatah Palembang dapat melihat data pendaftar yang peminat yang paling rendah berdasarkan atribut alamat, asal sekolah, dan Program Studi agar lebih ditonjolkan lagi.

5.2. Saran.

Berdasarkan kesimpulan yang telah diuraikan tersebut, maka ada beberapa saran yang dapat di sampaikan yaitu :

1. Pada penelitian selanjutnya dapat di uji coba dengan dataset yang lebih besar sehingga nilai menghasilkan tingkat akurasi yang lebih tinggi
2. Pada penelitian yang akan datang, dapat juga ditambahkan penerapan data mining dengan menggunakan algoritma yang berbeda sehingga akan menghasilkan hasil yang lebih variatif dan tentunya bernilai informasi yang sangat tinggi.
3. Penelitian ini disarankan dapat menjadi bahan referensi yang dipergunakan dan dikembangkan untuk penelitian selanjutnya

DAFTAR PUSTAKA

- Andi. 2016. *Kerja Bareng dan Berbagi Data Pada Microsoft Office*. Yogyakarta: Penerbit Andi.
- Appriilia. 2013. *Belajar Data Mining Dengan Rapid Miner*. Jakarta: Penerbit Gava Media.
- Asrul. 2016. *Metode Naive Bayes Untuk Prediksi Kelulusan (Studi Kasus: Data Mahasiswa Baru Perguruan Tinggi)*. Bandung: PT Elex Media Komputindo.
- Dicky Nofriansyah, dkk. 2016. *Penerapan Data Mining dengan Algoritma Naive Bayes Clasifier Untuk Mengetahui Minat Beli Pelanggan terhadap Kartu Internet XL (Studi Kasus di CV. Sumber Utama Telekomunikasi)*. Bandung: CV. Sumber Utama Telekomunikasi.
- Elvira Asril dkk. 2015. *Analisis Data Lulusan dengan Data Mining untuk Mendukung Strategi Promosi Universitas Lancang Kuning*. Yogyakarta: Gava Media.
- Griffin. 2014. *Pengantar Bisnis*. Jakarta: Penerbit Salemba.
- Hermawati. 2013. *Data Mining*. Yogyakarta: Penerbit ANDI, Jl. Beo.
- Kurtz. 2010. *Pengantar Bisnis Kontemporer*. Jakarta Selatan: Penerbit Salemba Empat.
- Nuraeni. 2017. *Penentuan Kelayakan Kredit Dengan Algoritma Naïve Bayes Classifier*. Bandung: Mitra Buana.
- Pratiwi. 2013. *Buku Ajar Sistem Pendukung Keputusan*. Jakarta: Penerbit, Media.
- Saleh, Alfa. 2015. *Implementasi Metode Klasifikasi Naïve Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga*. Jakarta: Sumber Utama.
- Sarwandi.2017. *Jago Microsoft exel 2016*. Jakarta: PT Elex Media Komputindo.
- Siska haryati, dkk. 2015. *Implementasi Data Mining Untuk Memprediksi Masa Studi Mahasiswa Menggunakan Algoritma C4.5 (Studi Kasus: Universitas Dehasen Bengkulu)*. Bengkulu:Universitas Dehasen Bengkulu.

- Sugiyono.2017. *Metode Penelitian Kuantitatif, Kualitatif dan R&D*.Bandung: Penerbit Alfabeta.
- Vulandari. 2017.*Datan Mining*. Jakarta: Penerbit Gava Media.
- Vulandari.2017. *Data Mining Teori dan Aplikasi Rapidminer*. Yogyakarta: Penerbit Gava Media.
- Wahyudi, Deni. 2015. *Implementasi data mining dengan naive bayes classifier untuk mendukung strategi promosi (studi kasus universitas bina darma palembang)*. Palembang: Universitas bina darma.