

**KLASIFIKASI PENERIMA BEASISWA
MENGUNAKAN ALGORITMA NAIVE BAYES
CLASSIFIER**

(Studi Kasus : Universitas Ichsan Gorontalo)

Oleh

IMAM NORALIM ANSAR

T3115162

SKRIPSI



**PROGRAM SARJANA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS ICHSAN GORONTALO
GORONTALO
2020**

PENGESAHAN SKRIPSI
KLASIFIKASI PENERIMA BEASISWA MENGGUNAKAN
ALGORITMA NAIVE BAYES CLASSIFIER

(Studi Kasus : Universitas Ichsan Gorontalo)

Oleh

IMAM NORALIM ANSAR

T3115162

SKRIPSI

Untuk memenuhi salah satu syarat ujian

Guna memperoleh gelar Sarjana

Program Studi Teknik Informatika, dan telah disetujui oleh Tim Pembimbing

Gorontalo, 03 Desember 2020

Pembimbing I


Zohrahavaty, M.Kom
NIDN. 0912117702

Pembimbing II


Yulianty Lasena, M.Kom
NIDN. 0907078603

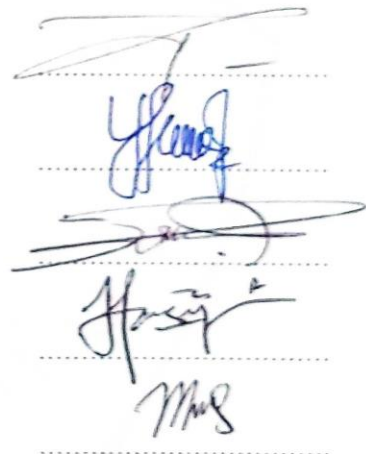
PERSETUJUAN SKRIPSI
KLASIFIKASI PENERIMA BEASISWA MENGGUNAKAN
ALGORITMA NAIVE BAYES CLASSIFIER

(Studi Kasus : Universitas Ichsan Gorontalo)

Oleh
IMAM NORALIM ANSAR
T3115162

Diperiksa Oleh Panitia Ujian Sastra Satu (S1)
Universitas Ichsan Gorontalo
Gorontalo, 07 Desember 2020

1. Pembimbing I
Zohrahayaty, M.Kom
2. Pembimbing II
Yulianty Lasena, M.Kom
3. Penguji I
Irvan Abraham Salihi, M.Kom
4. Penguji II
Haditsah Annur, M.Kom
5. Penguji III
Muis Nanja, M.Kom



Mengetahui :


Dekan Fakultas Ilmu Komputer

Zohrahayaty, M.Kom
NIDN. 0912117702

Ketua Program Studi

Irvan Abraham Salihi, M.Kom
NIDN. 0928028101

PERNYATAAN SKRIPSI

Dengan ini saya menyatakan bahwa:

1. Karya tulis (Skripsi) saya ini adalah asli dan belum pernah diajukan untuk mendapatkan gelar akademik (Sarjana) baik di Universitas Ichsan Gorontalo maupun diperguruan tinggi lainnya.
2. Karya Tulis (Skripsi) saya ini adalah murni gagasan, rumusan, dan penelitian saya sendiri tanpa bantuan pihak lain, kecuali arahan dari Tim Pembimbing.
3. Dalam Karya tulis (Skripsi) saya ini tidak terdapat karya atau pendapat yang telah di publikasikan orang lain, kecuali secara tertulis dicantumkan sebagai acuan/sitasi dalam naskah dan dicantumkan pula dalam daftar pustaka.
4. Pernyataan ini saya buat dengan sesungguhnya dan apabila dikemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya tulis ini, serta sanksi lainnya sesuai dengan norma - norma yang berlaku di Universitas Ichsan Gorontalo.

Gorontalo, 03 Desember 2020

Yang membuat pernyataan,


Imam Noraim Ansar

ABSTRACT

Data mining is a technique that utilizes large amounts of data to obtain valuable information that was previously unknown and can be used for important decision making. In this study, the authors attempted to mine data (data mining) of Ichsan University students in Gorontalo to determine the eligibility status of students in obtaining BBM scholarships (Student Learning Aid). Existing data were analyzed using the Naive Bayes algorithm. Naive Bayes is a method of probabilistic reasoning. The Naive Bayes algorithm aims to classify data in a certain class, then this pattern can be used to predict which customers will join, so that the company can make a decision to accept or reject the prospective customer.

Keywords: data mining, scholarship, classification, Naive Bayes algorithm

ABSTRAK

Data mining adalah teknik yang memanfaatkan data dalam jumlah yang besar untuk memperoleh informasi berharga yang sebelumnya tidak diketahui dan dapat dimanfaatkan untuk pengambilan keputusan penting. Pada penelitian ini, penulis berusaha menambang data (data mining) mahasiswa Universitas Ichsan Gorontalo untuk mengetahui status kelayakan mahasiswa dalam memperoleh beasiswa BBM (Bantuan Belajar Mahasiswa). Data yang ada dianalisis menggunakan algoritma Naive Bayes. Naive Bayes merupakan salah satu metode pada probabilistic reasoning. Algoritma Naive Bayes bertujuan untuk melakukan klasifikasi data pada kelas tertentu, kemudian pola tersebut dapat digunakan untuk memperkirakan nasabah yang bergabung, sehingga perusahaan bisa mengambil keputusan menerima atau menolak calon nasabah tersebut.

Kata Kunci : data mining, beasiswa, klasifikasi, algoritma Naive Bayes

KATA PENGANTAR

Puji syukur, penulis panjatkan kehadiran Allah SWT atas berkat rahmat dan hidayah-Nya, sehingga penulis dapat menyelesaikan skripsi ini dengan judul: **“Klasifikasi Penerima Beasiswa Menggunakan Algoritma Naive Bayes Classifier”**

untuk memenuhi salah satu syarat untuk Ujian Akhir guna memperoleh gelar Sarjana Komputer pada Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo. Salam dan taslim kepada junjungan kita, Nabi Muhammad SAW atas perjuangan beliau yang telah mengantarkan kita dari alam kebodohan ke alam yang penuh ilmu pengetahuan. Penelitian ini penulis beri judul: **“Klasifikasi Penerima Beasiswa Menggunakan Algoritma Naive Bayes Classifier.”**

Penulis menyadari sepenuhnya bahwa usulan penelitian ini tidak mungkin terwujud tanpa bantuan dan dorongan dari berbagai pihak, baik bantuan moril maupun materil. Untuk itu, dengan segala keikhlasan dan kerendahan hati, penulis mengucapkan terima kasih dan penghargaan yang setinggi-tingginya kepada:

1. Bapak Mohamad Ichsan Gaffar, SE, M.AK, selaku Ketua Yayasan Pengembangan Ilmu Pengetahuan dan Teknologi (YPIPT) Ichsan Gorontalo;
2. Bapak Dr. Abdul Gaffar La Tjokke, M.Si., selaku Rektor Universitas Ichsan Gorontalo;
3. Ibu Zohrahayaty, M.Kom, selaku Dekan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo, sekaligus sebagai pembimbing 1 dalam penelitian ini.
4. Sudirman S. Panna M.Kom, selaku Wakil Dekan I Bidang Akademik Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.
5. Ibu Irma Surya Kumala Idris, M.Kom., selaku Wakil Dekan II Bidang Kepegawaian, Administrasi Umum, dan Keuangan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.

6. Bapak Sudirman Melangi, M.Kom, selaku Wakil Dekan III Bidang Kemahasiswaan Fakultas Ilmu Komputer Universitas Ichsan Gorontalo.
7. Bapak Irvan Abraham Salihi, M.Kom., selaku Ketua Program Studi Teknik Informatika Fakultas Ilmu Komputer Universitas Ichsan Gorontalo;
8. Ibu Yulianty Lasena M.Kom., selaku dosen Pembimbing Pendamping yang telah membimbing penulis dalam penyusunan usulan penelitian ini;
9. Bapak dan Ibu Dosen Universitas Ichsan Gorontalo yang telah mendidik dan mengajarkan berbagai disiplin ilmu kepada penulis;
10. Kedua Orang Tua dan keluarga atas segala kasih sayang, jerih payah dan doa restunya dalam membesarkan, mendidik penulis, serta memberikan bantuan dan dukungan moril yang sangat besar kepada penulis;
11. Rekan-rekan seperjuangan yang telah banyak memberikan bantuan dan dukungan moril yang sangat besar kepada penulis;
12. Semua pihak yang ikut membantu dalam penyelesaian usulan penelitian ini yang tak sempat penulis sebutkan satu-persatu. Semoga Allah, SWT melimpahkan balasan atas jasa-jasa mereka dan kepada penulis. Selanjutnya, penulis menyadari sepenuhnya bahwa apa yang telah dicapai ini masih jauh dari kesempurnaan dan masih banyak terdapat kekurangan. Oleh karena itu, penulis sangat mengharapkan adanya kritik dan saran yang konstruktif.

Gorontalo, 03 Desember 2020

Imam Noralim Ansar

DAFTAR ISI

PENGESAHAN SKRIPSI	i
PERSETUJUAN SKRIPSI.....	ii
PERNYATAAN SKIRPSI.....	iii
ABSTRACT	iv
ABSTRAK	v
KATA PENGANTAR.....	vi
DAFTAR ISI.....	viii
DAFTAR GAMBAR.....	x
DAFTAR TABEL	xiii
DAFTAR LAMPIRAN	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Identifikasi Masalah	4
1.3 Rumusan Masalah	4
1.4 Tujuan Penelitian	4
1.5 Manfaat Penelitian	5
BAB II LANDASAN TEORI	6
2.1 Tinjauan Studi	6
2.2 Tinjauan Pustaka	6
2.2.1 Beasiswa.....	6
2.2.2 Data Mining	7
2.2.3 Naive Bayes Classifier	8
2.2.4 Confusion Matrix	12
2.2.5 Rekayasa Perangkat Lunak	12
2.2.6 Konstruksi Sistem	20
2.2.7 Pengujian Sistem dengan White Box Testing.....	20

2.2.8	Pengujian Sistem dengan Black Box Testing	21
2.2.9	Perangkat Lunak Pendukung.....	22
2.2.10	Kerangka Pikir	23
BAB III METODE PENELITIAN		24
3.1	Objek Penelitian	24
3.2	Model yang diusulkan	24
3.3	Pengumpulan Data	24
3.4	Klasifikasi	25
3.5	Pengujian Model	25
3.6	Pengembangan Sistem	25
3.6.1	Analisis Sistem.....	25
3.6.2	Desain Sistem.....	26
3.6.3	Konstruksi Sistem	26
3.6.4	Pengujian Sistem.....	27
3.6.5	White Box Testing	27
3.6.6	Black Box Testing.....	27
BAB IV HASIL PENELITIAN.....		28
4.1	Hasil Pengumpulan Data.....	28
4.2	Use Case.....	29
4.3	Activity Diagram.....	30
4.3.1	Digram Login admin	30
4.3.2	Diagram Data Training	30
4.3.3	Diagram Activity Data Testing	31
4.3.4	Diagram Activity Hasil Klasifikasi.....	31
4.3.5	Diagram Activity Probabilitas.....	32
4.4	Class Diagram	33
4.5	UML Sequence.....	34
4.5.1	Login	34
4.5.2	Training Atribut	35

4.5.4	Klasifikasi Data.....	37
4.5.5	Hasil Klasifikasi.....	38
4.6	Arsitektur Sistem.....	38
4.7	Mekanisme User.....	39
4.8	Format File.....	39
4.9	Program Desain.....	39
4.10	Flowchart Pengujian White Box.....	41
4.11	Flowgraph Pengujian White Box.....	42
4.12	Pengujian Black Box.....	44
BAB V	PEMBAHASAN PENELITIAN	45
5.1	Prosedur Pengoperasian Sistem	45
5.2	Penerapan Algoritma Naive Bayes	50
BAB VI	KESIMPULAN DAN SARAN.....	52
6.1	Kesimpulan	52
6.2	Saran	52
DAFTAR PUSTAKA.....		53

DAFTAR GAMBAR

Gambar 2.1 Siklus hidup pengembangan sistem.....	13
Gambar 2.2 Contoh Bagan Alir	20
Gambar 2.3 Kerangka Pikir	23
Gambar 3.1 Model klasifikasi yang diusulkan	24
Gambar 4.1 Use Case Diagram	29
Gambar 4.2 Diagram Login Admin.....	30
Gambar 4.3 Data Training	30
Gambar 4.4 Diagram Testing	31
Gambar 4.5 Diagram Hasil Klasifikasi.....	31
Gambar 4.6 Diagram Hasil Probabilitas.....	32
Gambar 4.7 Class Diagram.....	33
Gambar 4.8 UML sequence Login	34
Gambar 4.9 UML Sequence Training Atribut.....	35
Gambar 4.10 UML sequence Data Training	36
Gambar 4.11 UML Sequence Klasifikasi Data	37
Gambar 4.12 UML Sequence Hasil Klasifikasi	38
Gambar 4.13 Flowchart Pengujian White Box.....	41
Gambar 4.14 Flowchart Pengujian White Box.....	42
Gambar 5.1 XAMPP Control Panel	45
Gambar 5.2 Menu Login	46
Gambar 5.3 Halaman Utaman	46
Gambar 5.4 Menu Data Training.....	47
Gambar 5.5 Form Input Data Training.....	47
Gambar 5.6 Menu Probabilitas.....	48
Gambar 5.7 Halaman Data Testing	48
Gambar 5.8 Form Input Data Testing.....	49

Gambar 5.9 Halaman Proses Rekomendasi (klasifikasi).....	49
Gambar 5.10 Halaman Laporan.....	50

DAFTAR TABEL

Tabel 1.1 Dataset Penerima Beasiswa.....	3
Tabel 2.1 Data Training Mahasiswa.....	10
Tabel 2.2 Confusion Matrix	12
Tabel 2.3 Simbol-Simbol Activity Diagram	17
Tabel 2.4 Simbol-Simbol Class Diagram.....	18
Tabel 2.5 Simbol-Simbol Sequence Diagram	19
Tabel 2.6 Perangkat Lunak Pendukung.....	22
Tabel 3.1 Variabel Data Mahasiswa yang digunakan	24
Tabel 4.1 Data Permohonan Beasiswa BBM	28
Tabel 4.2 Interface design : Mekanisme User (Hasil Desain Sistem).....	39
Tabel 4.3 Program Desain	39
Tabel 4.4 Pengujian Black Box	44

DAFTAR LAMPIRAN

LAMPIRAN 1 : KODE PROGRAM NAÏVE BAYES	56
LAMPIRAN 2 : SURAT KETERANGAN PENELITIAN.....	58
LAMPIRAN 3 : SURAT REKOMENDASI BEBAS PLAGIASI	59
LAMPIRAN 4 : HASIL TURNITIN.....	60
LAMPIRAN 5 : SURAT KETERANGAN BEBAS PUSTAKA.....	62
LAMPIRAN 6 : RIWAYAT HIDUP PENELITI.....	63

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pendidikan merupakan salah satu faktor kemandirian bangsa. Untuk menciptakan generasi penerus yang berkualitas dari suatu bangsa, dapat dilihat kemajuan pendidikannya. Pendidikan bertujuan untuk mengembangkan potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, kreatif, mandiri dan menjadi warga negara yang demokratis serta bertanggung jawab. Untuk mewujudkan pendidikan yang berkualitas tentunya diperlukan biaya yang tidak sedikit, setiap peserta didik pada satuan pendidikan tertentu berhak mendapatkan biaya pendidikan bagi mereka yang orang tuanya tidak mampu membiayai pendidikannya dan berhak mendapatkan beasiswa bagi mereka yang berprestasi sesuai amanat Undang-undang pendidikan. Seiring banyaknya kesadaran masyarakat terhadap pentingnya pendidikan tinggi tentu berbanding lurus dengan tingginya jumlah mahasiswa pada suatu perguruan tinggi, hal ini dapat dimaksimalkan dengan kebijakan pemberian beasiswa.[1]

Beasiswa adalah pemberian berupa bantuan keuangan yang diberikan kepada perorangan yang memiliki tujuan untuk digunakan dalam keberlangsungan Pendidikan yang sedang ditempuh. Pemberian beasiswa merupakan program kerja yang ada di Kementerian Pendidikan dan Kebudayaan Republik Indonesia (Kemendikbud RI) maupun yang ada di institusi penyelenggara pendidikan tinggi [2]. Dengan adanya pemberian beasiswa ini diharapkan dapat meringankan beban mahasiswa dalam menempuh masa kuliah khususnya pada masalah beasiswa. Proses pendistribusian beasiswa kepada mahasiswa biasa dilakukan secara selektif sesuai dengan jenis beasiswa yang terdapat pada perguruan tinggi.[3]

Universitas Ichsan Gorontalo setiap tahunnya mendapatkan kuota program beasiswa dari Lembaga Layanan Pendidikan Tinggi Wilayah IX (LLDIKTI IX), yaitu Beasiswa Peningkatan Prestasi Akademik (PPA) dan Beasiswa Bantuan Belajar Mahasiswa (BBM) [4] . Data beasiswa yang ada di Universitas Ichsan

Gorontalo tidak banyak memiliki kegunaan, belum menjadi sebuah informasi dan pengetahuan yang dapat dijadikan untuk melihat indikator proses penyeleksian beasiswa. Data tentang beasiswa dapat memberikan informasi berguna bagi perguruan tinggi jika dimanfaatkan dengan maksimal. Setiap tahunnya terdapat peningkatan data penerima beasiswa sedangkan kuota penerima beasiswa yang diberikan dari pihak kementerian sangat terbatas. Melihat indikator dalam penyeleksian berkas pengajuan beasiswa dengan beberapa macam kriteria, namun masih menggunakan cara manual dalam proses penyeleksiannya, maka pemilihan mahasiswa yang berhak menerima beasiswa berdasarkan kelengkapan data yang dikumpulkan terkesan subjektif dari sisi pengambil keputusan, kurang efektif dan membutuhkan waktu yang relatif lebih lama.

Untuk mempermudah proses penyeleksian agar lebih objektif, maka diperlukan adanya klasifikasi untuk penyeleksian penerimaan beasiswa apakah pemohon diterima atau ditolak untuk program beasiswa. Klasifikasi merupakan salah satu dari teknik data mining yang merupakan proses penempatan objek atau konsep tertentu ke dalam satu set kategori berdasarkan objek yang digunakan. Salah satu metode yang dapat digunakan untuk proses klasifikasi data yaitu *Naive Bayes Classifier*. Metode Naïve Bayes Classifier merupakan salah satu metode yang dapat digunakan dalam hal pengambilan keputusan untuk mendapatkan hasil yang lebih baik pada suatu permasalahan klasifikasi [5], [6].

Pada penelitian [3] kelayakan calon penerima beasiswa bidikmisi ditentukan dengan menerapkan metode Naïve Bayes. 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%. Dengan menerapkan metode naive bayes classifier diharapkan mampu memberikan sebuah hasil rekomendasi kepada perguruan tinggi terhadap suatu proses penyeleksian program beasiswa dengan pemohon yang sangat banyak. Berikut adalah dataset data penerima beasiswa :

Tabel 1.1 Dataset Penerima Beasiswa

No	NIM	NAMA MAHASISWA	ANGKATAN	SEMESTER	PRODI	IPK
1	S2118128	Avany Dhea Arum Sastrawaty	2018	2	S1 - ilmu Pemerintah	3.54
2	S2216042	Dhea Rizky Syukur	2018	6	S1 – Ilmu komunikasi	3.15
3	S2218016	Silvinka Nur Kosasih	2018	2	S1 – Ilmu komunikasi	3.92
4	S2218020	Cindy Claudia Mohamad	2018	2	S1 – Ilmu komunikasi	3.83
5	S2218037	Aldiyansah Kuice	2018	2	S1 – Ilmu komunikasi	3.92
6	T1116020	M. Yusuf	2016	6	S1 – Teknik Arsitektur	3.77
7	T1118007	Jasir Labinta	2018	2	S1 – Teknik Aristektur	3.52
8	T1118014	Mawarsari Chisty Theodora Tiwow	2018	2	S1 – Aristektur	3.64
9	T1118063	Shafira Nur Kmarulah	2018	2	S1 – Teknik Aristektur	3.23
10	T1118064	Nurmala Masry	2018	2	S1 – Teknik Aristektur	3.42
11	T1118070	Pangeran Muh. Sultan Ilyas	2018	2	S1 – Teknik Aristektur	3.45
12	T2117014	Israwani Ahmadi	2017	4	S1 – Teknik Elektro	3.42
...	...	...	...	...	...	...

Metode klasifikasi lain yang sering digunakan untuk penentuan penerima mahasiswa adalah dengan algoritma C4.5 [1], dimana diperoleh akurasi yang cukup baik namun masih dibutuhkan atribut yang lebih banyak agar rules lebih akurat.

Pada penelitian ini, peneliti membatasi masalah pada klasifikasi pemberian beasiswa BBM, menyesuaikan data primer yang diperoleh, dengan menggunakan algoritma *Naive Bayes Classifier* untuk proses klasifikasinya. Peneliti mengusulkan judul penelitian ***Klasifikasi Penerima Beasiswa menggunakan Naive Bayes Classifier.***

1.2 Identifikasi Masalah

Berdasarkan latar belakang yang telah diuraikan, maka identifikasi masalah dalam penelitian ini adalah Algoritma Naïve Bayes sangat sering digunakan untuk menyelesaikan masalah klasifikasi data namun belum pernah diuji coba pada data penerima beasiswa di Universitas Ichsan Gorontalo.

1.3 Rumusan Masalah

Berdasarkan identifikasi masalah, maka dapat dirumuskan masalah penelitian yaitu :

1. Bagaimana hasil uji coba (pemodelan) *Naive Bayes Classifier* untuk klasifikasi penerima beasiswa di Universitas Ichsan Gorontalo?
2. Bagaimana mereayasa aplikasi untuk klasifikasi penerima beasiswa di Universitas Ichsan Gorontalo?

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, maka tujuan penelitian ini adalah: Memperoleh Model *Algoritma Naïve Bayes Classifier* yang paling baik kinerjanya untuk klasifikasi pemberian beasiswa di Universitas Ichsan Gorontalo.

1.5 Manfaat Penelitian

Manfaat penelitian ini adalah:

1. Manfaat Teoritis: Memberikan masukan bagi perkembangan ilmu pengetahuan dan teknologi, khususnya pada bidang ilmu komputer, yaitu berupa model algoritma *Naive Bayes Classifier* untuk data penerima beasiswa.
2. Manfaat Praktis: sebagai sumbangan pemikiran, karya, bahan pertimbangan, atau solusi bagi institusi perguruan tinggi guna mendukung pengambilan keputusan dalam rangka proses penyeleksian penerima beasiswa.

BAB II

LANDASAN TEORI

2.1 Tinjauan Studi

Beberapa penelitian yang terkait dengan penyeleksian beasiswa dengan beberapa metode di data mining, antara lain:

1. Pada penelitian [3], Penelitian ini berhasil menerapkan metode Naïve Bayes untuk membantu menyeleksi penerima beasiswa bidikmisi, dengan tingkat akurasi sebesar 85.56%.
2. Pada penelitian [8], pada kasus penentuan kelulusan mahasiswa juga menggunakan algoritma Naive Bayes Classifier untuk melakukan prediksi dengan perolehan Akurasi pengujian data yang diperoleh dalam penelitian ini sebesar 80,72% dari 1162 data yang digunakan untuk data training dan 587 data untuk data testing.
3. Sedangkan pada penelitian [1], penerapan algoritma c.45 untuk rekomendasi penerima beasiswa sudah memiliki akurasi yang cukup baik dimana penyeleksian nya telah memiliki aturan yang telah dilakukan proses klasifikasi dengan data mining. Namun pada penelitian, lebih ke development system tanpa mencantumkan pengukuran akurasinya.

2.2 Tinjauan Pustaka

2.2.1 Beasiswa

Beasiswa merupakan suatu bantuan untuk membantu pelajar atau mahasiswa yang masih sekolah atau kuliah supaya mereka bisa menyelesaikan tugasnya dalam mencari ilmu pengetahuan sampai selesai. Beasiswa dalam bentuk bantuan dapat berupa dana sebagai penunjang biaya yang harus dikeluarkan oleh pelajar atau mahasiswa selama menempuh masa pendidikan di tempat belajar. [2]

Undang-undang Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional, Bab V pasal 12 (1.c), menyebutkan bahwa setiap peserta didik pada setiap satuan pendidikan berhak mendapatkan beasiswa bagi yang berprestasi yang orang tuanya tidak mampu membiayai pendidikannya.

2.2.2 Data Mining

Klasifikasi adalah suatu proses untuk menemukan model atau fungsi yang menguraikan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui. Untuk mencapai tujuan tersebut, proses klasifikasi membentuk suatu model yang mampu membedakan data ke dalam kelas-kelas yang berbeda berdasarkan aturan atau fungsi tertentu. Model itu sendiri bisa berupa aturan “jika-maka”, pohon keputusan, atau formula matematis [9].

Tugas dalam data mining terbagi atas [10]:

1. *Clustering*. Membagi rangkaian data menjadi beberapa grup berdasarkan kesamaan-kesamaan yang telah ditentukan sebelumnya. *Clustering* ini termasuk dalam *unsupervised learning*.
2. *Classification*. Mirip dengan *clustering*, yaitu bertujuan untuk pengelompokkan, perbedaannya adalah *prediction* memiliki *target/class/label* atau dengan kata lainnya output pengelompokkannya sudah ditentukan. Komponen-komponen dalam *prediction* terdiri dari atribut, class, dataset latih, dan dataset uji. *Classification* termasuk dalam tugas *prediction*.
3. *Estimation*. Jika pada *classification*, class-nya bertipe nominal, maka dalam *estimation*, class-nya bertipe numerik. *Estimation* termasuk dalam tugas *prediction*.
4. *Regression*. Mempelajari hubungan antara suatu variabel dengan variabel lainnya. Regresi dimanfaatkan untuk menentukan hubungan kuantitatif antar variabel dan untuk meramalkan nilai dari variabel di masa yang akan datang. *Regression* termasuk dalam tugas *prediction*.
5. *Assosiation*. Dalam satu grup tertentu, tiap anggota memiliki satu hubungan dengan yang lain, juga terhadap grup-grup yang lain. Keterhubungan itu selanjutnya dibuat hokum/aturan yang dapat dijadikan patokan kebijakan tertentu.
6. *Anomaly Detection*. Terkadang kita ingin mengetahui adanya suatu penyimpangan pada data sebelum data tersebut diolah dengan teknik-teknik tertentu.

2.2.3 Naïve Bayes Classifier

Naïve Bayes adalah suatu pengklasifikasian probabilistik sederhana yang menghitung sekumpulan probabilitas dengan cara 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 dari *Naïve Bayes* adalah pengklasifikasian dengan menggunakan metode probabilitas dan statistik yang dikemukakan oleh ilmuwan Inggris Thomas Bayes, yaitu cara memprediksi peluang di masa depan berdasarkan hasil pengalaman di masa sebelumnya [11].

Naïve Bayes didasarkan pada asumsi penyederhanaan bahwa nilai atribut secara kondisional saling bebas jika diberikan nilai output. maka dari itu, diberikan nilai output, probabilitas mengamati secara bersama merupakan produk dari probabilitas individu. Keuntungan menggunakan metode *Naïve Bayes* adalah hanya membutuhkan jumlah data pelatihan (*Training Data*) yang kecil untuk menentukan estimasi parameter yang diperlukan dalam proses pengklasifikasian. *Naïve Bayes* dapat bekerja jauh lebih baik dalam suatu situasi dunia nyata yang kompleks daripada yang diharapkan, dengan kata lain model *Naïve Bayes* yang digunakan adalah model fitur independen.

Persamaan dari teorema Bayes adalah sebagai berikut :

$$1. \quad P(H|X) = \frac{P(X|H).P(H)}{P(X)}$$

Keterangan :

X : Data dengan kelas yang belum diketahui

H : Hipotesis data X adalah suatu kelas spesifik

$P(H|X)$: Probabilitas hipotesis data H berdasarkan kondisi X (posteriori probability)

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

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

$P(X)$: Probabilitas X

Untuk menjelaskan teorema Naïve Bayes, perlu diketahui bahwa proses klasifikasi perlu sejumlah petunjuk untuk menentukan kelas yang cocok bagi sampel yang dianalisis. Maka dari itu, teorema bayes di atas disesuaikan seperti berikut:

$$2. \quad P(C|F_1 \dots F_n) = \frac{P(C)P(F_1 \dots F_n|C)}{P(F_1 \dots F_n)}$$

Diketahui variabel C merepresentasikan kelas, sementara variabel $F_1 \dots F_n$ merepresentasikan karakteristik petunjuk yang dibutuhkan untuk melakukan klasifikasi. Maka dari itu, rumus tersebut menjelaskan bahwa peluang masuknya sampel karakteristik tertentu dalam kelas C (Posterior) adalah peluang munculnya kelas C (sebelum masuknya sampel, biasanya disebut prior), dikali dengan peluang munculnya karakteristik sampel pada kelas C (biasanya disebut likelihood), dibagi dengan peluang munculnya karakteristik sampel secara global (biasanya disebut evidence). Maka rumus diatas bisa ditulis dengan sederhana yaitu sebagai berikut:

$$3. \quad \text{Posterior} = \frac{\text{Prior} \times \text{Likelihood}}{\text{Evidence}}$$

Nilai Evidence akan tetap untuk setiap kelas pada satu sampel. Nilai dari posterior tersebut akan dibandingkan dengan nilai-nilai posterior kelas lainnya untuk dapat menentukan ke kelas suatu sampel yang akan diklasifikasikan.

Setelah itu ada alur dari metode Naïve Bayes yaitu sebagai berikut:

1. Baca data training
2. Hitung jumlah dan probabilitas, jika data numerik maka harus dilakukan:
 - a. Cari nilai mean dan standar deviasi dari masing-masing parameter yang data numerik
 - b. Cari nilai probabilitas dengan cara menghitung jumlah data yang sesuai dari kategori yang sama dibagi dengan jumlah data pada kategori tersebut
 - c. Menghasilkan nilai dalam tabel mean, standar deviasi, dan probabilitas
 - d. Solusi dapat dihasilkan.

Contoh penggunaan algoritma Naive Bayes Classifier :

Tabel 2.1 Data Training Mahasiswa

No	Jenis Kelamin	Status Pekerjaan	Status Nikah	IPK	Status Lulus
1.	Laki-laki	Mahasiswa	Belum	3.17	Tepat
2.	Laki-laki	Bekerja	Belum	3.30	Tepat
3.	Perempuan	Mahasiswa	Belum	3.01	Tepat
4.	Perempuan	Mahasiswa	Menikah	3.25	Tepat
5.	Laki-laki	Bekerja	Menikah	3.20	Tepat
6.	Laki-laki	Bekerja	Menikah	2.50	Terlambat
7.	Perempuan	Bekerja	Menikah	3.00	Terlambat
8.	Perempuan	Bekerja	Belum	2.70	Terlambat
9.	Laki-laki	Bekerja	Belum	2.40	Terlambat
10.	Perempuan	Mahasiswa	Menikah	2.50	Terlambat
11.	Perempuan	Mahasiswa	Belum	2.50	Terlambat
12.	Perempuan	Mahasiswa	Belum	3.50	Tepat
13.	Laki-laki	Bekerja	Menikah	3.30	Tepat
14.	Laki-laki	Mahasiswa	Menikah	3.25	Tepat
15.	Laki-laki	Mahasiswa	Belum	2.20	Terlambat

Data Testing :

JKEL	STATUS	PERNIKAHAN	IPK	KETERANGAN
Laki-laki	Mahasiswa	Belum	2.70	????

$P(C_i)$

$P(\text{STATUS KELULUSAN} = \text{"TEPAT"}) = 8/15$

$P(\text{STATUS KELULUSAN} = \text{"TERLAMBAT"}) = 7/15$

$P(X|C_i)$

$P(\text{JENIS KELAMIN} = \text{"LAKI - LAKI"} \mid \text{STATUS KELULUSAN} = \text{"TEPAT"}) = 5/8$

$$P(\text{JENIS KELAMIN} = \text{"LAKI - LAKI"} \mid \text{STATUS KELULUSAN} = \text{"TERLAMBAT"}) = 3/7$$

$$P(\text{STATUS MAHASISWA} = \text{"MAHASISWA"} \mid \text{STATUS KELULUSAN} = \text{"TEPAT"}) = 5/8$$

$$P(\text{STATUS MAHASISWA} = \text{"MAHASISWA"} \mid \text{STATUS KELULUSAN} = \text{"TERLAMBAT"}) = 3/7$$

$$P(\text{STATUS PRENIKAHAN} = \text{"BELUM"} \mid \text{STATUS KELULUSAN} = \text{"TEPAT"}) = 4/8$$

$$P(\text{STATUS PRENIKAHAN} = \text{"BELUM"} \mid \text{STATUS KELULUSAN} = \text{"TERLAMBAT"}) = 4/7$$

$$P(\text{IPK} = \text{"2.70"} \mid \text{STATUS KELULUSAN} = \text{"TEPAT"}) = 0/8$$

$$P(\text{IPK} = \text{"2.70"} \mid \text{STATUS KELULUSAN} = \text{"TERLAMBAT"}) = 1/7$$

$$\begin{aligned} P(X \mid \text{STATUS KELULUSAN} = \text{"TEPAT"}) &= P(\text{KELAMIN} = \text{"LAKI - LAKI"}, \\ &\text{STATUS MAHASISWA} = \text{"MAHASISWA"}, \text{STATUS PERNIKAHAN} = \\ &\text{"BELUM"}, \text{IPK} = 2.70 \mid \text{STATUS KELULUSAN} = \text{"TEPAT"}) \\ &= 5/8 * 5/8 * 4/8 \& 0/8 \\ &= 0 \end{aligned}$$

$$\begin{aligned} P(X \mid \text{STATUS KELULUSAN} = \text{"TERLAMBAT"}) &= P(\text{KELAMIN} = \text{"LAKI - LAKI"}, \\ &\text{STATUS MAHASISWA} = \text{"MAHASISWA"}, \text{STATUS PERNIKAHAN} = \\ &\text{"BELUM"}, \text{IPK} = 2.70 \mid \text{STATUS KELULUSAN} = \text{"TERLAMBAT"}) \\ &= 3/7 * 3/7 * 4/7 * 1/7 \\ &= 0.43 * 0.43 * 0.57 * 0.14 \\ &= 0,014 \end{aligned}$$

$$P(X|Ci)*P(Ci)$$

$$P(X \mid \text{STATUS KELULUSAN} = \text{"TEPAT"}) * P(\text{STATUS KELULUSAN} = \text{"TEPAT"}) = 0 * 8/15 = 0$$

$$P(X \mid \text{STATUS KELULUSAN} = \text{"TERLAMBAT"}) * P(\text{STATUS KELULUSAN} = \text{"TERLAMBAT"}) = 0.014 * 7/15 = 0.0069$$

Untuk KELAMIN = “LAKI – LAKI”, STATUS MAHASISWA = “MAHASISWA”, STATUS PERNIKAHAN = “BELUM”, IPK = 2.70, masuk ke kelas STATUS KELULUSAN = “TERLAMBAT”

2.2.4 Confusion Matrix

Pada permasalahan klasifikasi statistik, confusion matrix, yang juga dikenal sebagai matriks kesalahan, merupakan tata letak tabel yang spesifik dan memungkinkan visualisasi dari kinerja algoritma, yang biasanya merupakan tipikal dari supervised learning (dalam unsupervised learning biasanya disebut matriks pencocokan). Setiap kolom dari matriks mewakili contoh dalam kelas yang diprediksi, sementara setiap baris mewakili contoh di kelas sebenarnya (atau sebaliknya). Nama tersebut berasal dari fakta bahwa mudah untuk melihat apakah sistem memiliki dua kelas yang membingungkan (biasanya salah mengartikannya sebagai yang lain). Confusion matrix adalah sebuah metode untuk evaluasi yang menggunakan tabel matrix yang terdapat pada Tabel 2.1.

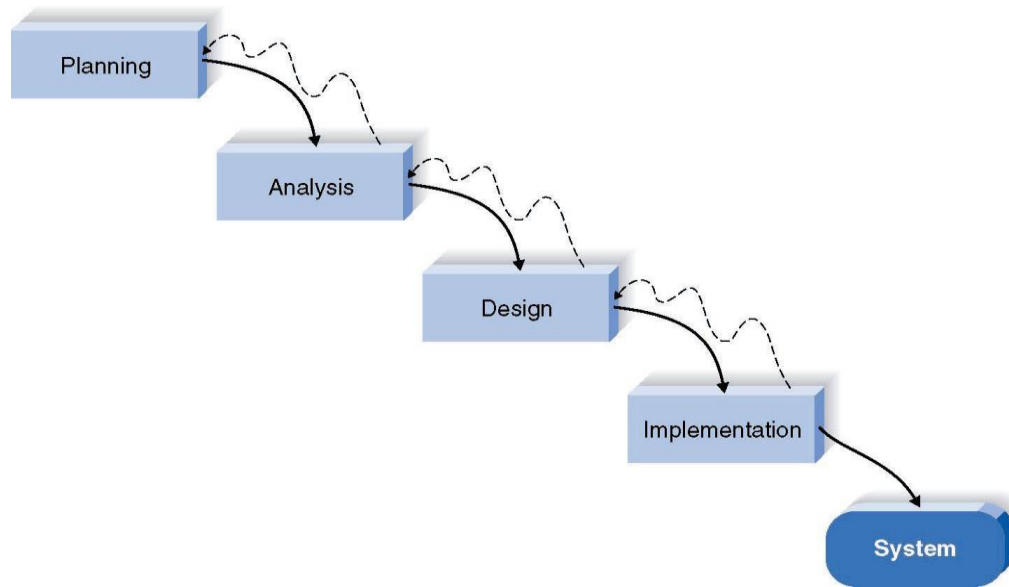
Tabel 2.2 Confusion Matrix

	True Condition	
Predicted Condition	Positive	Negative
Positive	True Positive	False Positive
Negative	False Negative	True Negative

2.2.5 Rekayasa Perangkat Lunak

Proses pengembangan sistem melewati beberapa tahapan dari mulai sistem itu direncanakan sampai dengan sistem tersebut diterapkan, dioperasikan dan dipelihara. Bila operasi sistem yang sudah dikembangkan masih timbul kembali permasalahan-permasalahan yang kritis serta tidak dapat diatasi dalam tahap pemeliharaan sistem, maka perlu dikembangkan kembali suatu sistem untuk mengatasinya dan proses ini kembali ke tahap yang pertama, yaitu tahap perencanaan sistem. Siklus ini disebut dengan siklus hidup suatu sistem (*systems*

life cycle). Daur atau siklus hidup dari pengembangan sistem merupakan suatu bentuk yang digunakan untuk menggambarkan tahapan utama dan langkah-langkah didalam tahapan tersebut dalam proses pengembangannya [12].



Gambar 2.1 Siklus hidup pengembangan sistem

Software (perangkat lunak) berhubungan dengan: (1) Perintah (program komputer) yang bila dieksekusi memberikan fungsi dan unjuk kerja seperti yang diinginkan; (2) Struktur data yang memungkinkan program memanipulasi informasi secara proporsional; dan (3) dokumen yang menggambarkan operasi dan kegunaan program. “*Software engineering* adalah suatu disiplin rekayasa (rancang-bangun) yang terkait dengan semua aspek produksi perangkat lunak. Sedangkan menurut IEEE (Institute of Electrical and Electronic Engineers) “*Software engineering* adalah: (1) Aplikasi dari sebuah pendekatan yang bersifat kuantifiabel, disiplin, dan sistematis bagi pengembangan, operasi, dan pemeliharaan perangkat lunak; (2) Studi tentang pendekatan-pendekatan seperti pada [12]”. Dengan demikian bidang penelitian software engineering mengacu pada kedua hal tersebut.

Perbedaan utama paradigma berorientasi objek dibandingkan konvensional adalah pada penyatuan proses/fungsi dan data ke dalam bentuk yang terenkapsulasi, sedangkan paradigma konvensional memisahkan data dengan

proses. Data merupakan sekumpulan atribut yang di-enkapsulasi (dikemas/dibungkus) bersama algoritma (operasi/metode/servis) untuk melakukan suatu proses berdasarkan pesan (stimulan/message/event) yang masuk ke objek/class. Pesan ini merupakan sarana interface antar-objek [13].

Objek-objek merupakan realitas yang unik sehingga tidak efisien bila dideskripsikan satu persatu sementara mereka memiliki beberapa kesamaan. Misal kerbau dengan sapi adalah dua objek yang berbeda, tetapi keduanya adalah pemakan rumput. Oleh sebab itu diciptakan kelas yang memuat karakteristik yang sama dari kerbau dan sapi. Dengan demikian kelas ini dapat menginformasikan berbagai binatang pemakan rumput (tetapi tidak bisa menyebutkan binatang-binatang di kebun binatang). Kelas ini dengan mudah dikembangkan untuk merekrut objek/binatang-binatang lain pemakan rumput secara efisien tanpa mengulang-ulang deskripsi yang sama. Bahkan dapat dibuat hirarkhinya sebagai superkelas atau subkelas yang masing-masing memiliki warisan (inheritance) dari kelas di atasnya. Sedangkan *Polymorphism* diperlukan guna memperluas sistem (software engineering) berbasis objek yang ada secara efisien, yakni dengan nama yang sama tetapi operasinya berbeda [13].

Pertanyaan pada object oriented analysis adalah: *who, what, when, where will the system be* [14]. Tahap ini mengidentifikasi dan mengumpulkan kebutuhan sistem dengan cara menjawab: (1) *Who will use the system?* (2) *What will the system do?* dan (3) *When will it be used?* Menginvestigasi sistem yang sedang berjalan, kemudian mengidentifikasi/mengusulkan kemungkinan-kemungkinan perbaikannya, dan akhirnya membangun konsep untuk sistem yang baru. Beberapa hal yang dilakukan pada tahap ini adalah sebagai berikut [14] :

1. *Requirements determination*

- Laporan berupa daftar persyaratan fungsional dan non-fungsional.
- Semua persyaratan tersebut harus dilacak kembali.

2. *Functional modeling*

- *Business process modeling with Activity Diagram.*
- Dengan *Use Case Diagram*:
 - Bagaimana bisnis berinteraksi dengan lingkungannya;

- Kegiatan yang dilakukan oleh pengguna;
- Aktivitas sistem atau apa yang akan dilakukan oleh sistem;
- Digunakan untuk dokumentasi dan pembangunan sistem;
- Dapat digunakan dengan baik sebagai komunikasi dengan pengguna.

3. *Structural modeling*

- Menggambarkan struktur data;
- Struktur data yang direpresentasikan melalui *Class Diagram*.
- Mengurangi "*semantic gap*" antara dunia nyata dan dunia *software*.
- Digunakan sebagai kosakata untuk analisis dan pengguna.

4. *Behavioral modeling*

- Untuk proses bisnis digambarkan dengan *Sequence Diagram* dan *Collaboration Diagram*.
- Untuk perubahan data digambarkan dengan *Statechart Diagram*.
- Tujuan *behavioral modeling* adalah:
 - Bagaimana objek berkolaborasi
 - Menggambarkan pandangan internal proses bisnis
 - Menunjukkan efek dari variasi-variasi proses pada sistem.

Sedangkan pertanyaan pada *object oriented design* adalah: *How will the system work* [14]. Dibandingkan dengan pendekatan konvensional, arsitektur *object oriented design* tidak memperlihatkan hirarkhi kontrol tetapi kolaborasi antar objek dengan aliran kontrol.

Beberapa hal yang dilakukan pada tahap ini adalah sebagai berikut [13]:

1. *Moving on to Design*

Tujuan dari tahap *object oriented analysis* adalah untuk mencari tahu apa kebutuhan bisnis. Sedangkan tujuan dari tahap *object oriented design* adalah cara membuat cetak biru sistem.

2. *Physical Architecture Layer Design (Architectur Design)*

- *Server based, client based, atau client server based*
- *Network model*
- *Hardware and software specification*
- *Global support and security plan*

3. *Human Computer Interaction Layer Design (Interface Design)*

- *User interface*
- *System interface*
- *Navigation mechanism*
- *Input mechanism (Form)*
- *Output mechanism (Report)*
- *Graphical User Interface (GUI)*

4. *Data Management Layer Design (Data Design)*

Pada dasarnya format yang digunakan untuk data, yaitu:

- *File*
- *Relational database (SQL)*
- *Object relational database (Extended SQL: ad hoc atau SQL3)*
- *Object oriented database (Sistem manajemen database yang terpisah)*

5. *Class and Method Design (Program Design)*

Kelas yang sudah ada perlu dipahami dan terorganisir dengan baik, pada umumnya beberapa kode, masih memerlukan *instantiate* kelas baru. Komponen-komponen pada paradigma berbasis objek antara lain: (1) *Classes*; (2) *Objects*; (3) *Attributes*; (4) *States*; (5) *Methods*; (6) *Messages/Event*.

Seperti yang telah dipaparkan sebelumnya, alat bantu yang biasanya digunakan pada tahap *object oriented analysis and design* adalah *The Unified Modelling Language (UML)*.

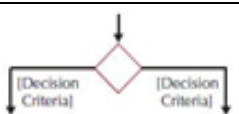
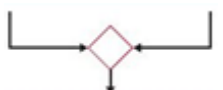
“UML can be used for modeling all processes in the development life cycle and across different implementation technologies. UML is the standard language for visualizing, specifying, constructing, and documenting the artifacts of a software-intensive system [14]

UML merupakan *tools* pemodelan yang bermanfaat dalam rekayasa perangkat lunak. UML menawarkan sebuah standar untuk merancang model sebuah sistem. Dengan menggunakan UML dapat dibuat model untuk semua jenis aplikasi perangkat lunak, yang dapat beroperasi pada perangkat keras, sistem operasi, dan jaringan apapun, serta ditulis dalam bahasa pemrograman apapun. Tetapi karena UML juga menggunakan *class* dan *operation* dalam konsep

dasarnya, maka ia lebih cocok untuk pemodelan perangkat lunak dalam bahasa berorientasi objek seperti C++, Java, C#, atau VB.NET, dsb., meskipun tidak menutup kemungkinan untuk pemodelan aplikasi prosedural dalam VB, Pascal, atau C.

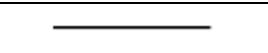
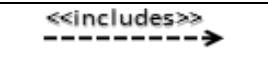
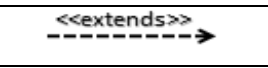
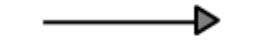
Seperti bahasa-bahasa lainnya, UML mendefinisikan notasi dan *syntax*/semantik. Notasi UML merupakan sekumpulan bentuk khusus untuk menggambarkan berbagai diagram perangkat lunak. Notasi UML terutama diturunkan dari 3 notasi yang telah ada sebelumnya: Grady Booch OOD (Object-Oriented Design), Jim Rumbaugh OMT (Object Modeling Technique), dan Ivar Jacobson OOSE (Object-Oriented Software Engineering). Simbol-simbol yang digunakan pada UML dapat ditunjukkan pada berbagai tabel berikut [14]:

Tabel 2.3 Simbol-Simbol Activity Diagram

Simbol	Keterangan
	<i>Action/activity</i>
	<i>Object/class</i>
	<i>Control/object flow</i>
	<i>Initial node</i>
	<i>Activity final node</i>
	<i>Decision node</i>
	<i>Merge node</i>
	<i>Fork node</i>
	<i>Join node</i>

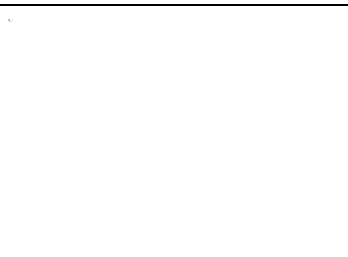
Sumber: [14]

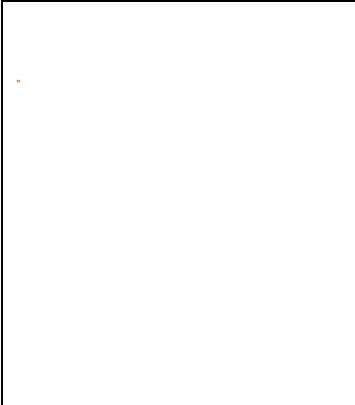
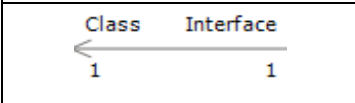
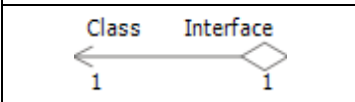
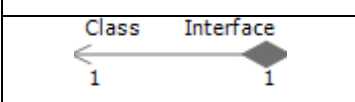
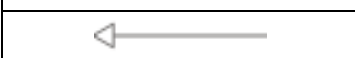
Tabel 2.4 Simbol-Simbol Use Case Diagram

Simbol	Keterangan
	<i>Actor</i>
	<i>UseCase</i>
	<i>Association relationship</i>
	<i>Include relationship</i>
	<i>Extend relationship</i>
	<i>Generalization relationship</i>

Sumber: [14]

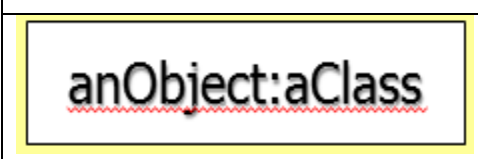
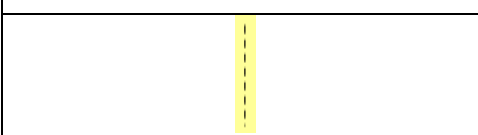
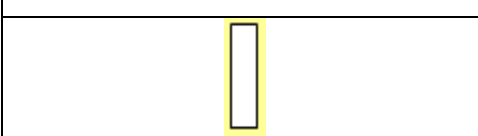
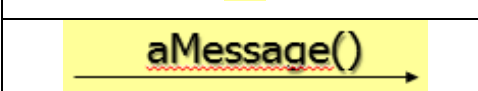
Tabel 2.5 Simbol-Simbol Class Diagram

Simbol	Keterangan
	<i>Class/Object</i>
	<i>Interface</i>
	<i>Enumeration</i>

	<i>Package</i>
	<i>Association</i>
	<i>Aggregation</i>
	<i>Composition</i>
	<i>Dependency</i>
	<i>Inheritance</i>

Sumber: [14]

Tabel 2.6 Simbol-Simbol Sequence Diagram

Simbol	Keterangan
	<i>Actor</i>
	<i>Object</i>
	<i>Lifeline</i>
	<i>Focus of control</i>
	<i>Message</i>

X	<i>Object destruction</i>
----------	---------------------------

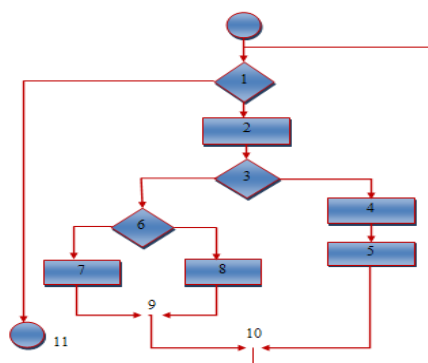
Sumber: [14]

2.2.6 Konstruksi Sistem

Tahapan implementasi merupakan tahap dimana dilakukan transformasi/penerjemahan dari bahasa *modeling* ke suatu bahasa pemrograman. hal ini merupakan tugas dari pemrogram, pada pengembangan sistem/perangkat lunak berorientasi objek penerjemahan dari setiap diagram-diagram UML yang telah dirancang pada tahap analisis dan desain harus diterjemahkan ke dalam bahasa pemrograman sama persis dengan diagram-diagram yang ada guna menghindari terjadinya perubahan fungsi/tujuan dari pengembangan sistem/perangkat lunak.

2.2.7 Pengujian Sistem dengan White Box Testing

Pengujian *White Box*, adalah metode pengujian yang menggunakan struktur kontrol desain prosedur untuk memperoleh *test case*. Dengan menggunakan metode *white box*, perekraya sistem dapat melakukan *test case* yang memberikan jaminan bahwa semua jalur independen pada suatu modul telah digunakan paling tidak satu kali, menggunakan semua keputusan logis pada sisi *true* dan *false*, mengeksekusi semua *loop* pada batasan mereka dan pada batas operasional mereka, dan menggunakan struktur data internal untuk menjamin validitasnya [13].



Gambar 2.2 Contoh Bagan Alir

Bagan alir digunakan untuk menggambarkan struktur kontrol program dan untuk menggambarkan grafik alir, harus memperhatikan representasi desain prosedural pada bagan alir. Pada gambar dibawah ini, grafik alir memetakan

bagan alir tersebut ke dalam grafik alir yang sesuai (dengan mengasumsikan bahwa tidak ada kondisi senyawa yang diisikan di dalam diamond keputusan dari bagan alir tersebut). Masing-masing lingkaran, yang disebut *simpul* grafik alir, merepresentasikan satu atau lebih statemen prosedural. Urutan kotak proses dan permata keputusan dapat memetakan simpul tunggal. Anak panah tersebut yang disebut *edges* atau *links*, merepresentasikan aliran kontrol dan analog dengan anak panah bagan alir. Edge harus berhenti pada suatu simpul, meskipun bila simpul tersebut tidak merepresentasikan statemen prosedural [13].

2.2.8 Pengujian Sistem dengan Black Box Testing

Black box approach adalah suatu sistem dimana *input dan output*-nya Dapat didefinisikan tetapi prosesnya tidak diketahui atau tidak terdefinisi. Metode ini hanya dapat dimengerti oleh pihak dalam (yang menangani sedangkan pihak luar hanya mengetahui masukan dan hasilnya). Sistem ini terdapat pada subsistem tingkat terendah. Metode ujicoba *black box* memfokuskan pada keperluan fungsional dari *software*. Karena itu ujicoba *black box* memungkinkan pengembang *software* untuk membuat himpunan kondisi *input* yang akan melatih seluruh syarat-syarat fungsional suatu program. Ujicoba *black box* bukan merupakan alternatif dari ujicoba *white box*, tetapi merupakan pendekatan yang melengkapi untuk menemukan kesalahan lainnya, selain menggunakan metode *white box*.

Uji coba *black box* berusaha untuk menemukan kesalahan dalam beberapa kategori, diantaranya:

1. Fungsi-fungsi yang salah atau hilang
2. Kesalahan *interface*
3. Kesalahan dalam struktur data atau akses *database eksternal*
4. Kesalahan performa
5. Kesalahan inisialisasi dan terminasi

Tidak seperti metode *white box* yang dilaksanakan diawal proses, ujicoba *black box* diaplikasikan dibeberapa tahapan berikutnya karena ujicoba *black box*

dengan sengaja mengabaikan struktur kontrol, sehingga perhatiannya difokuskan pada informasi *domain*. Ujicoba didesain untuk dapat menjawab pertanyaan-pertanyaan berikut:

1. Bagaimana validitas fungsionalnya diuji?
2. Jenis *input* seperti apa yang akan menghasilkan kasus uji yang baik?
3. Apakah sistem secara khusus sensitif terhadap nilai input tertentu?
4. Bagaimana batasan-batasan kelas data diisolasi?
5. Berapa rasio data dan jumlah data yang dapat ditoleransi oleh sistem?
6. Apa akibat yang akan timbul dari kombinasi spesifik data pada operasi sistem?

Dengan mengaplikasikan uji coba *black box*, diharapkan dapat menghasilkan sekumpulan kasus uji yang memenuhi kriteria berikut:

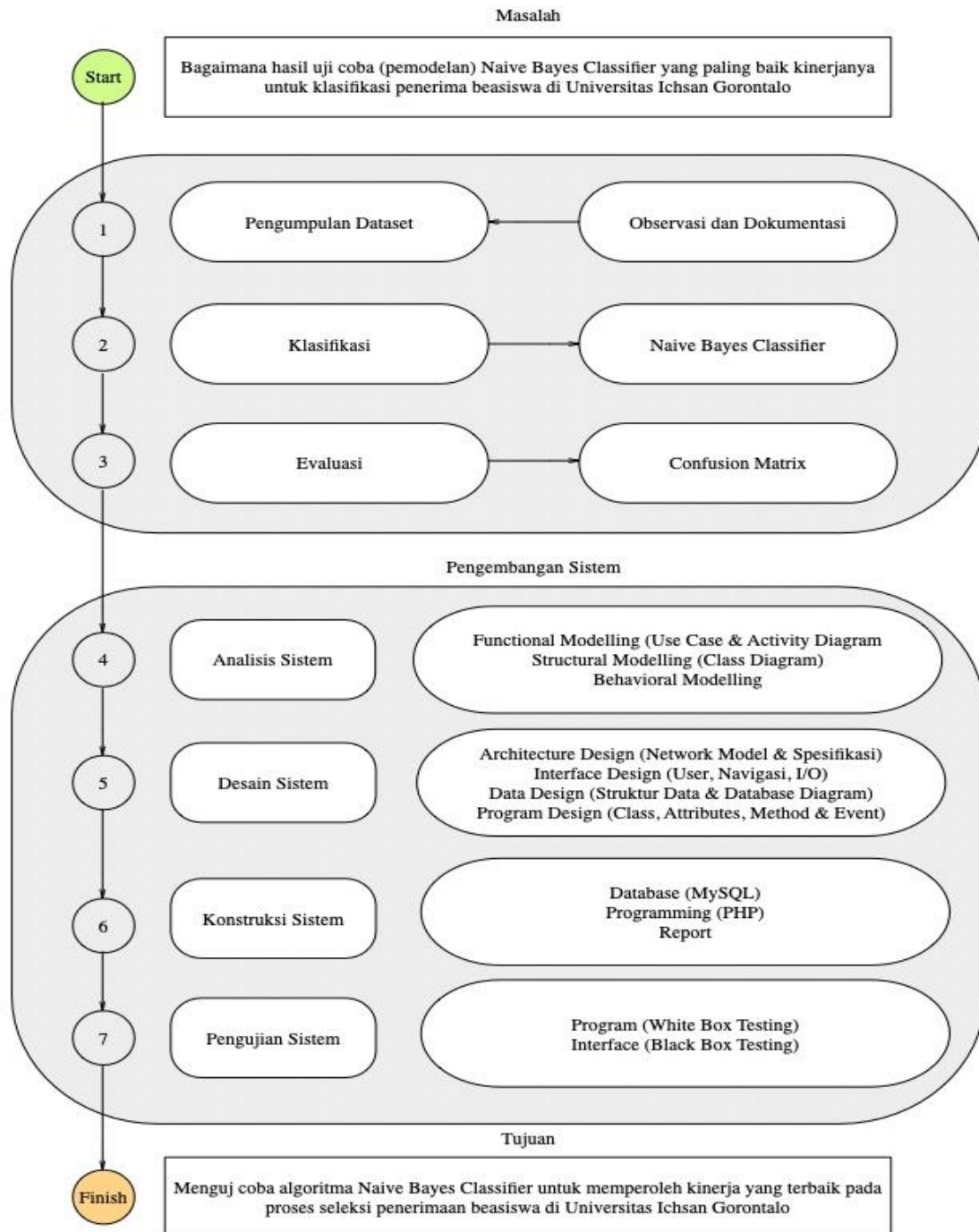
1. Kasus uji yang berkurang, jika jumlahnya lebih dari 1, maka jumlah dari uji kasus tambahan harus didesain untuk mencapai uji coba yang cukup beralasan.
2. Kasus uji yang memberitahukan sesuatu tentang keberadaan atau tidaknya suatu jenis kesalahan, dari pada kesalahan yang terhubung hanya dengan suatu uji coba yang spesifik.

2.2.9 Perangkat Lunak Pendukung

Tabel 2.7 Perangkat Lunak Pendukung

No	TOOLS	KEGUNAAN
1.	Rapid Miner	Digunakan untuk melakukan pemodelan Naïve Bayes untuk klasifikasi penerima beasiswa.
2.	Visual Studio Code (Editor)	Digunakan untuk mengembangkan sistem klasifikasi penerima beasiswa dengan Naive Bayes
3.	MySQL	Digunakan untuk mengolah <i>database</i>

2.2.10 Kerangka Pikir



Gambar 2.3 Kerangka Pikir

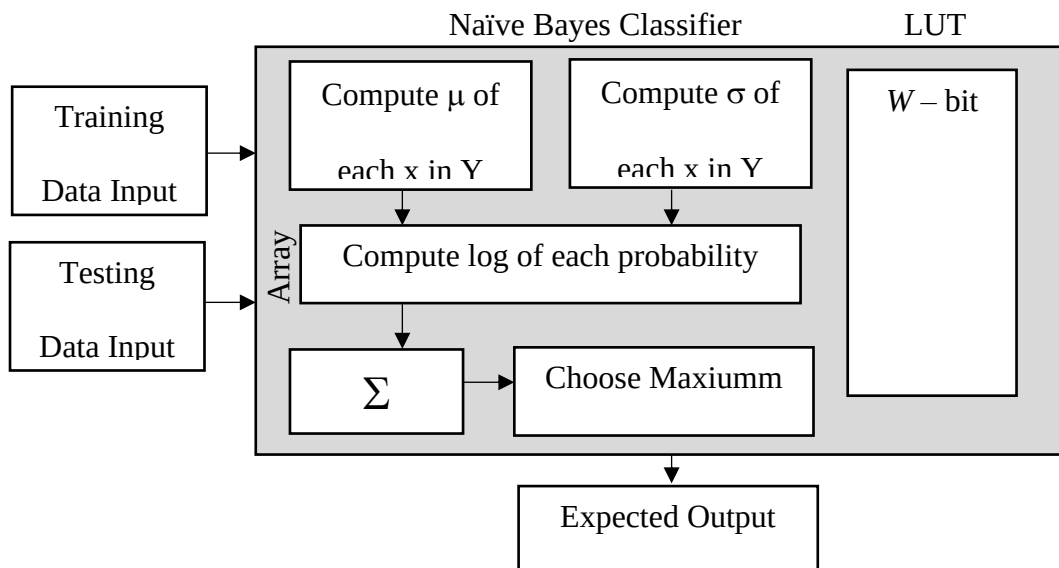
BAB III METODE PENELITIAN

3.1 Objek Penelitian

Penelitian ini menggunakan metode penelitian Deskriptif. Objek penelitian ini adalah data mahasiswa Universitas Ichsan Gorontalo. Penelitian ini dimulai dari Bulan Agustus 2020 s/d November 2020.

3.2 Model yang diusulkan

Gambar berikut menunjukkan model klasifikasi dengan menggunakan Naive Bayes Classifier.



Gambar 3.1 Model klasifikasi yang diusulkan

3.3 Pengumpulan Data

Data dikumpulkan dari BAAK Universitas Ichsan Gorontalo. Adapun variabel-variabel pada data mahasiswa ditunjukkan pada Tabel 3.1 berikut ini.

Tabel 3.1 Variabel Data Mahasiswa yang digunakan

No	ID	Nama Variabel	Tipe	Keterangan
1.	X1	NIM	Real	Variabel Input
2.	X2	Nama	Polynomial	Variabel Input

3.	X3	Tanggal Lahir	Real	Tidak Diolah
4.	X4	Angkatan	Real	Variabel Input
5.	X5	Program Studi	Real	Variabel Input
6.	X6	IPK	Real	Variabel Input
7.	X7	SKS	Real	Variabel Input
8.	X8	Penghasilan Orang Tua	Real	Variabel Input
9.	X9	Jumlah Piagam	Real	Variabel Input

Seperti yang telah dijelaskan sebelumnya bahwa kriteria penerima beasiswa berdasarkan buku pedoman penerimaan beasiswa [2]. Dimana untuk kriteria penerima beasiswa BBM adalah :

1. IPK minimal 2,75
2. penghasilan orang tua kategori rendah
3. Rekening Listrik

3.4 Klasifikasi

Setelah data ditransformasi, kemudian dianalisis polanya menggunakan algoritma Naïve Bayes Classifier untuk memperoleh pengelompokan yang terbaik. Alat bantu yang digunakan adalah Rapid Miner.

3.5 Pengujian Model

Setiap kali dilakukan percobaan menggunakan Naïve Bayes, dilanjutkan dengan pengujiannya menggunakan teori probabilitas. Alat bantu yang digunakan adalah Rapid Miner.

3.6 Pengembangan Sistem

3.6.1 Analisis Sistem

Analisis sistem menggunakan pendekatan berorientasi objek yang digambarkan dalam bentuk.

1. *Functional Modelling*, menggunakan alat bantu UML, dalam bentuk:
 - *Use Case Diagram*.

- *Activity Diagram*.
- 2. *Structural Modelling*, menggunakan alat bantu UML, dalam bentuk:
 - *Class Diagram*.
- 3. *Behavioral Modelling*, menggunakan alat bantu UML, dalam bentuk:
 - *Sequence Diagram*.

3.6.2 Desain Sistem

Desain sistem menggunakan pendekatan berorientasi objek yang digambarkan dalam bentuk.

4.1 *Architecture Design*, menggunakan alat bantu Visio dalam bentuk:

- a. Model jaringan dari sistem *client server*.
- b. Spesifikasi *hardware* dan *software* yang direkomendasikan.

4.2 *Interface Design*, menggunakan alat bantu Visio dalam bentuk:

- a. Mekanisme *User*.
- b. Mekanisme Navigasi.
- c. Mekanisme Input (Page jika web atau Form jika desktop).
- d. Mekanisme Output (Report).

4.3 *Data Design*, menggunakan alat bantu Visio dalam bentuk:

- a. Format data yang digunakan *Relational Database* (SQL)
- b. Struktur Data.
- c. *Database Diagram*.

4.4 *Program Design*, menggunakan alat bantu Visual Studio dalam bentuk:

- a. *Class*
- b. *Attributes*
- c. *Method*
- d. *Event*

3.6.3 Konstruksi Sistem

Pada tahap ini menerjemahkan hasil pada tahap analisis dan desain ke dalam kode-kode program komputer kemudian membangun aplikasinya. Model yang telah diperoleh kemudian dibuatkan aplikasinya. Alat bantu yang digunakan

pada tahap ini adalah Visual Studio Code dengan Bahasa pemrograman PHP, alat bantu data base yang digunakan MySQL.

3.6.4 Pengujian Sistem

Pengujian sistem menggunakan metode White Box dan Black Box Testing.

3.6.5 White Box Testing

Software yang telah direkayasa kemudian diuji dengan metode *White Box Testing* pada kode program proses penerapan metodenya/modelnya. Kode program tersebut dibuatkan *flowchart* programnya, kemudian dipetakan kedalam bentuk *flowgraph* (bagan alir kontrol) yang tersusun dari beberapa node dan edge. Berdasarkan *flowgraph*, ditentukan jumlah *Region* dan *Cyclomatic Complexity* (CC). Apabila *independent path* = $V(G) = (CC) = \text{Region}$, di mana setiap *path* hanya dieksekusi sekali dan sudah benar, maka sistem dinyatakan efisien dari segi kelayakan logika pemrograman.

3.6.6 Black Box Testing

Selanjutnya software diuji pula dengan metode *Black Box Testing* yang fokus pada keperluan fungsional dari *software* dan berusaha untuk menemukan kesalahan dalam beberapa kategori, diantaranya: (1) Fungsi-fungsi yang salah atau hilang; (2) Kesalahan *interface*; (3) Kesalahan dalam struktur data atau akses basis data eksternal; (4) Kesalahan performa; (5) Kesalahan inisialisasi dan terminasi. Jika sudah tidak ada kesalahan-kesalahan tersebut, maka sistem dinyatakan efisien dari segi kesalahan komponen-komponen sistem.

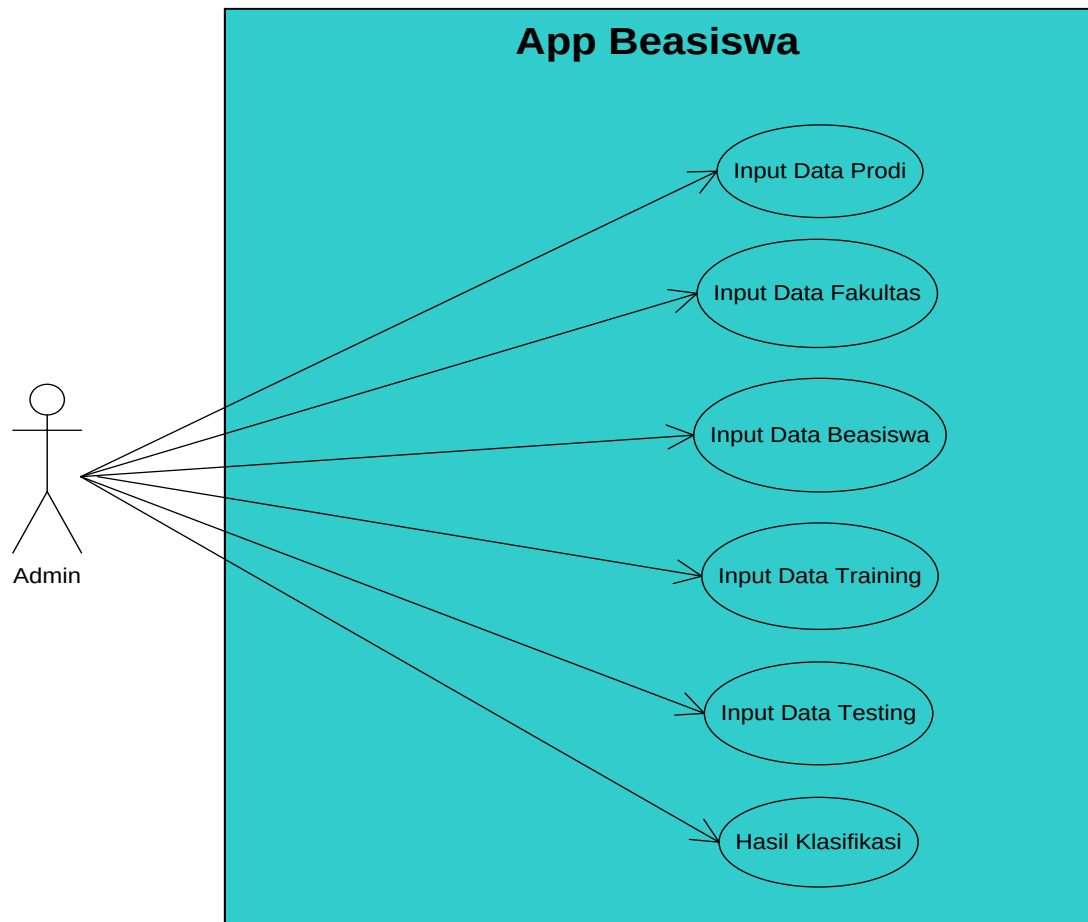
BAB IV HASIL PENELITIAN

4.1 Hasil Pengumpulan Data

Tabel 4.1 Data Permohonan Beasiswa BBM

No	NIM	Angkatan	Program Studi	IPK	SKS	Penghasilan Orang Tua	Jumlah Piagam
1.	T3118297	2018	Teknik Informatika	3.42	92	Rp.1.000.000	4
2.	E1118003	2018	Akuntansi	3,35	98	Rp.5.000.000	5
3.	E1118004	2018	Akuntansi	3.12	76	Rp.2.000.000	1
4.	H1118064	2018	Ilmu Hukum	3,46	95	Rp.4.500.000	4
5.	H1118014	2018	Ilmu Hukum	3.33	90	Rp.2.000.000	3
6.	P2218017	2018	Agribisnis	3.00	88	Rp.2.000.000	3
7.	P2218018	2018	Agribisnis	3.20	89	Rp.1.000.000	3
8.	T3118003	2018	Teknik Informatika	2.24	72	Rp.1.000.000	3
9.	T3118003	2018	Teknik Informatika	2.24	72	Rp.1.000.000	3
10.	P2218001	2018	Agribisnis	2.87	71	Rp.2.000.000	3
11	S2118021	2018	Ilmu Pemerintahan	3.00	80	Rp.3.000.000	2
12	S2118022	2018	Ilmu Pemerintahan	3.13	78	Rp.2.500.000	4
13	T2118001	2018	Teknik Elektro	3.23	92	Rp.1.500.000	3
14	T2118002	2018	Teknik Elektro	2.85	80	Rp.2.000.000	2
15	S2218003	2018	Ilmu Komunikasi	2.90	86	Rp.2.500.000	5
16	S2218004	2018	Ilmu Komunikasi	3.44	82	Rp.4.000.000	3
17	T1118016	2018	Teknik Arsitektur	3.26	81	Rp.1.000.000	2
18	T1118018	2018	Teknik Arsitektur	3.32	84	Rp.2.000.000	5
...	...	...	...	...	...	...	...

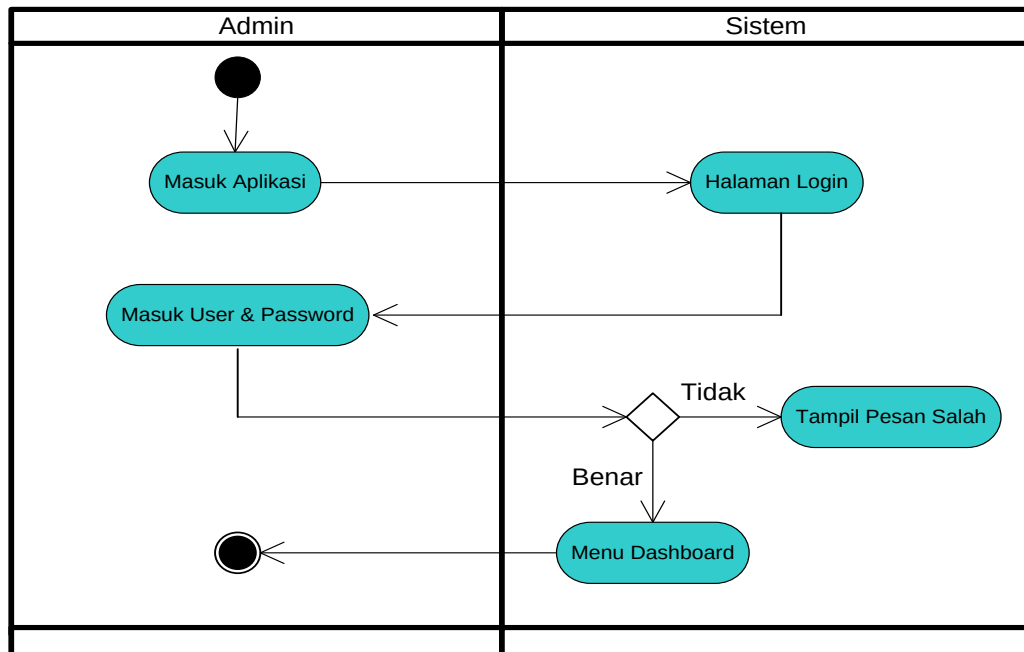
4.2 Use Case



Gambar 4.1 Use Case Diagram

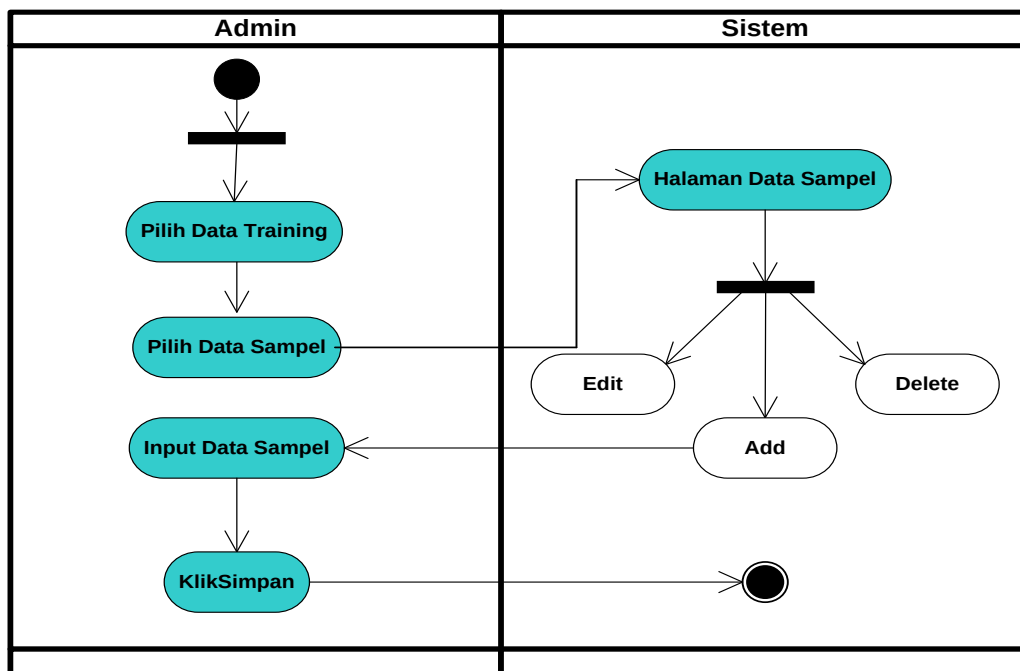
4.3 Activity Diagram

4.3.1 Diagram Login admin



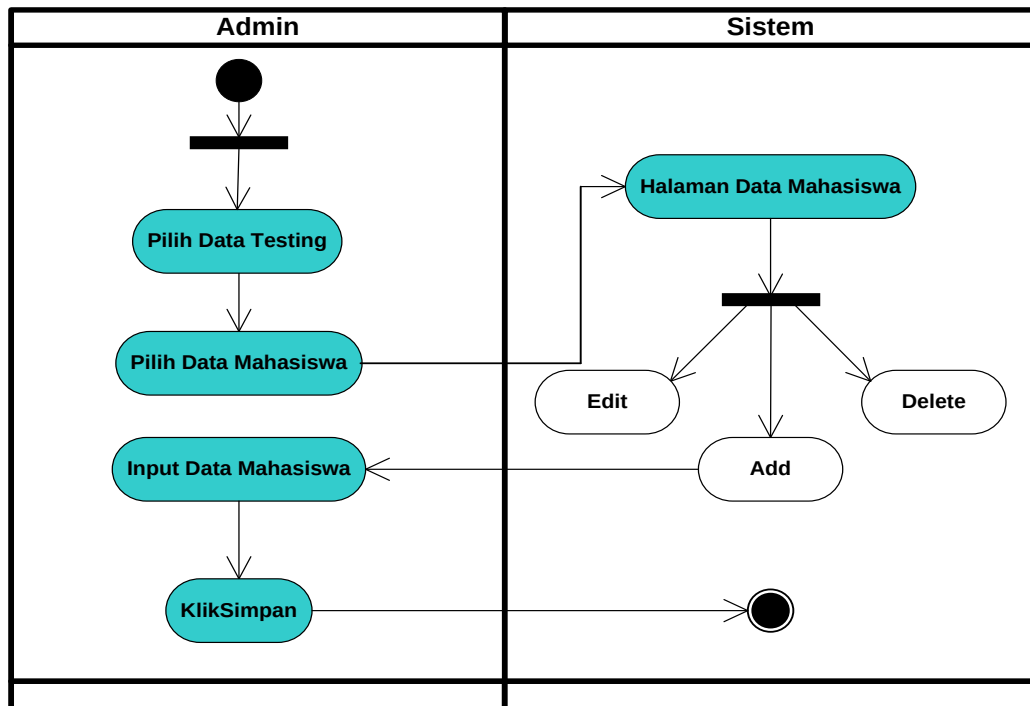
Gambar 4.2 Diagram Login Admin

4.3.2 Diagram Data Training



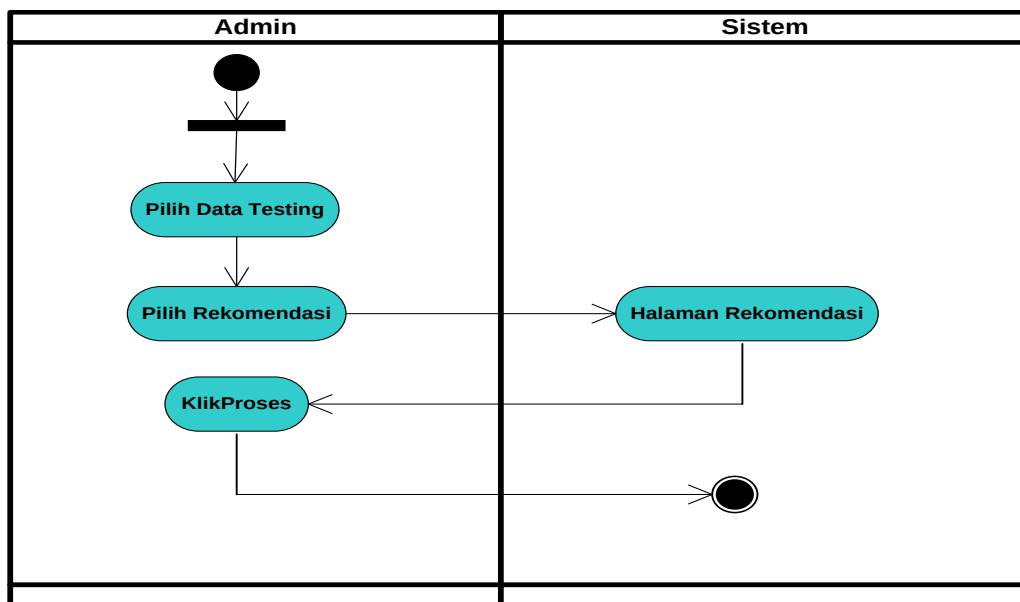
Gambar 4.3 Data Training

4.3.3 Diagram Activity Data Testing



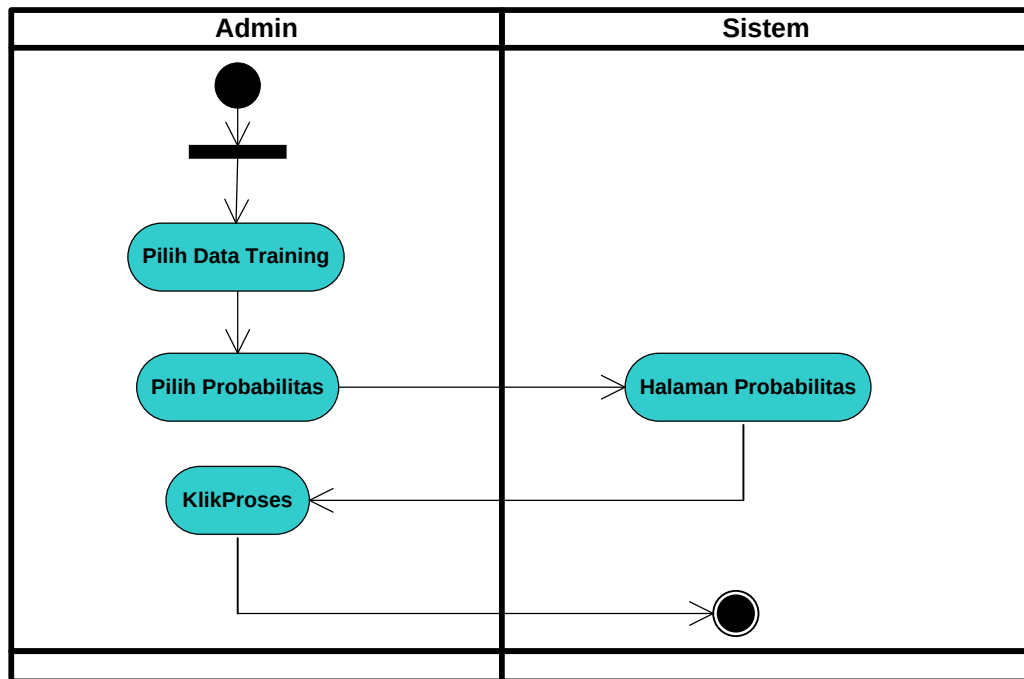
Gambar 4.4 Diagram Testing

4.3.4 Diagram Activity Hasil Klasifikasi



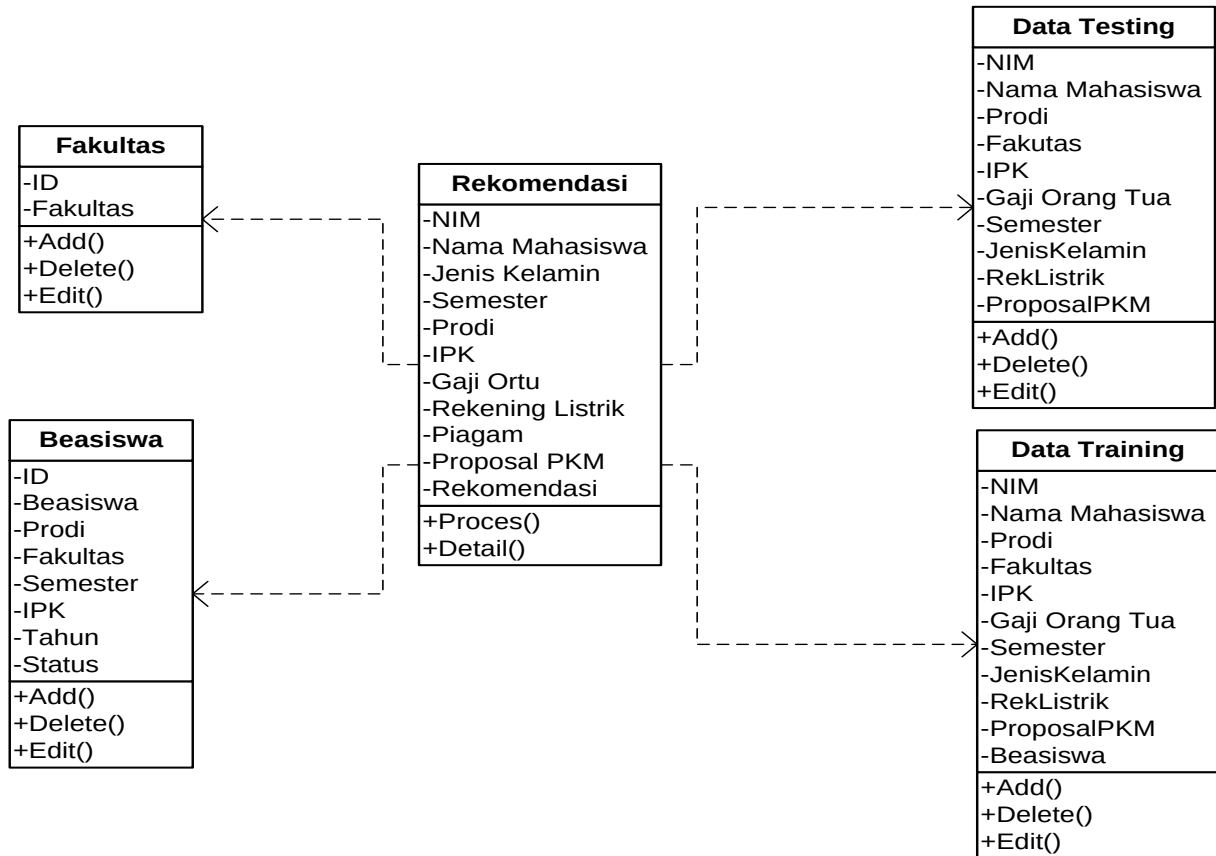
Gambar 4.5 Diagram Hasil Klasifikasi

4.3.5 Diagram Activity Probabilitas



Gambar 4.6 Diagram Hasil Probabilitas

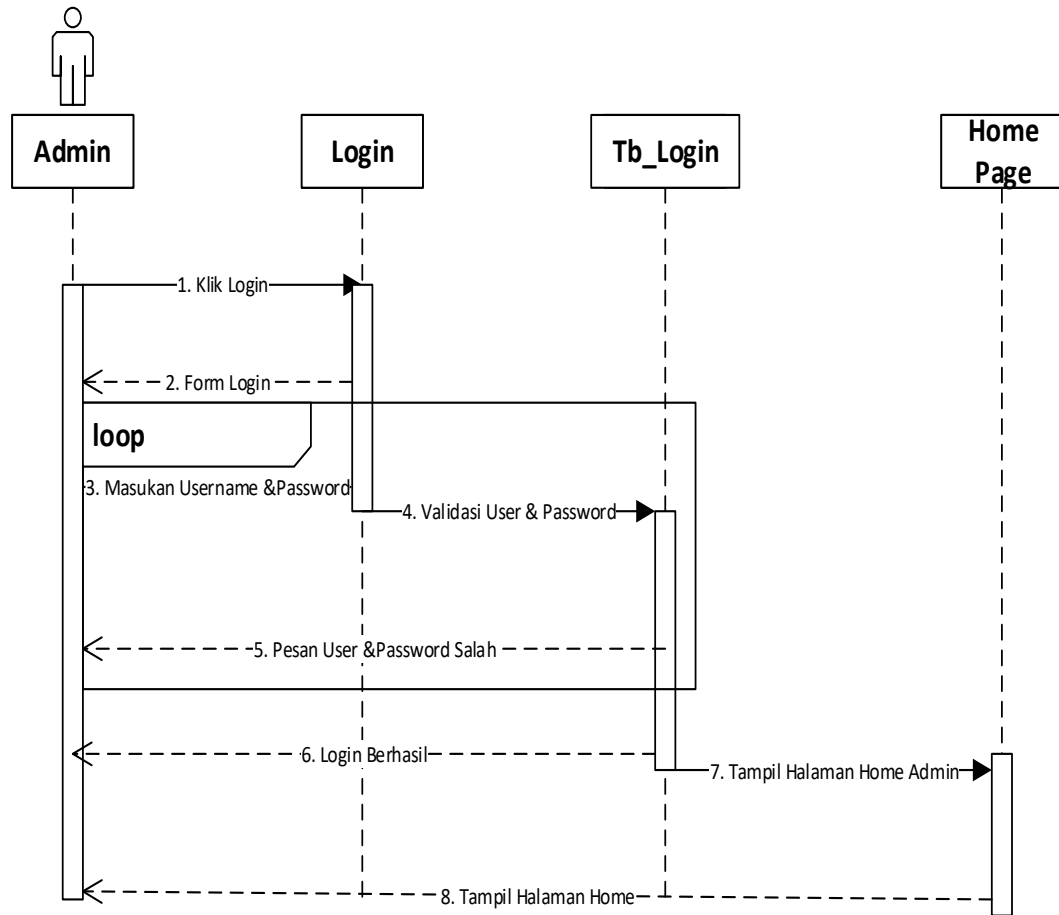
4.4 Class Diagram



Gambar 4.7 Class Diagram

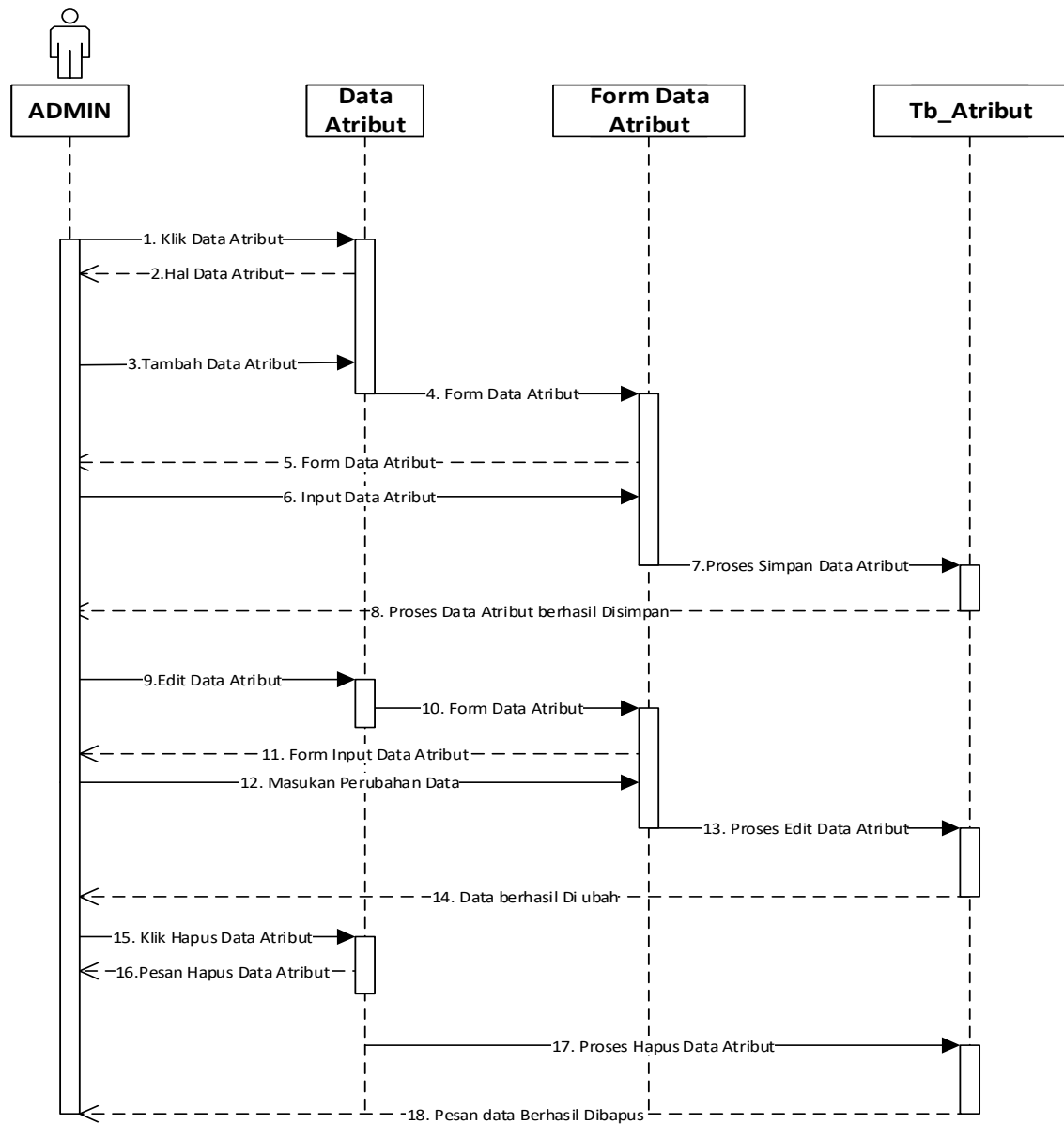
4.5 UML Sequence

4.5.1 Login



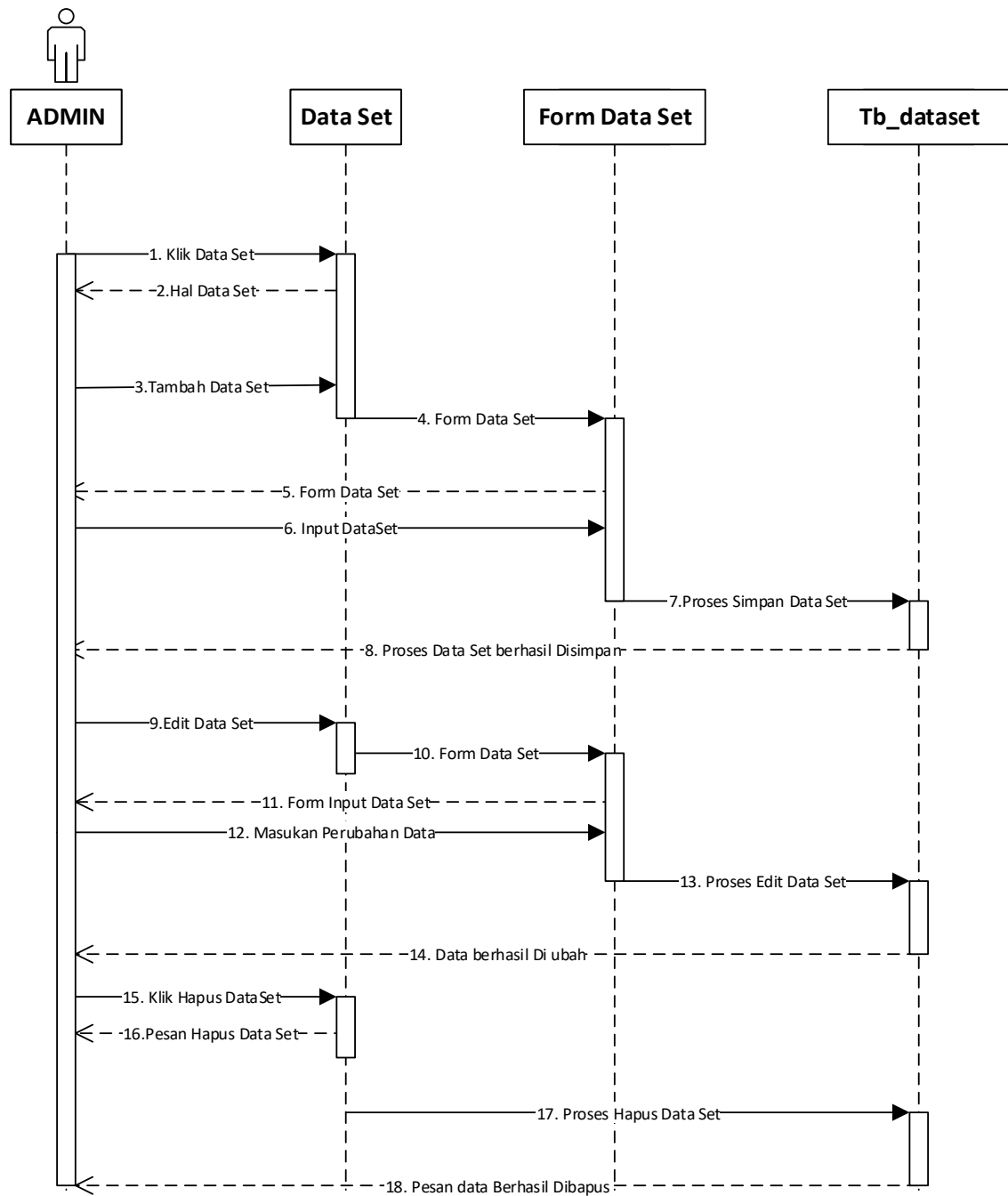
Gambar 4.8 UML sequence Login

4.5.2 Training Atribut



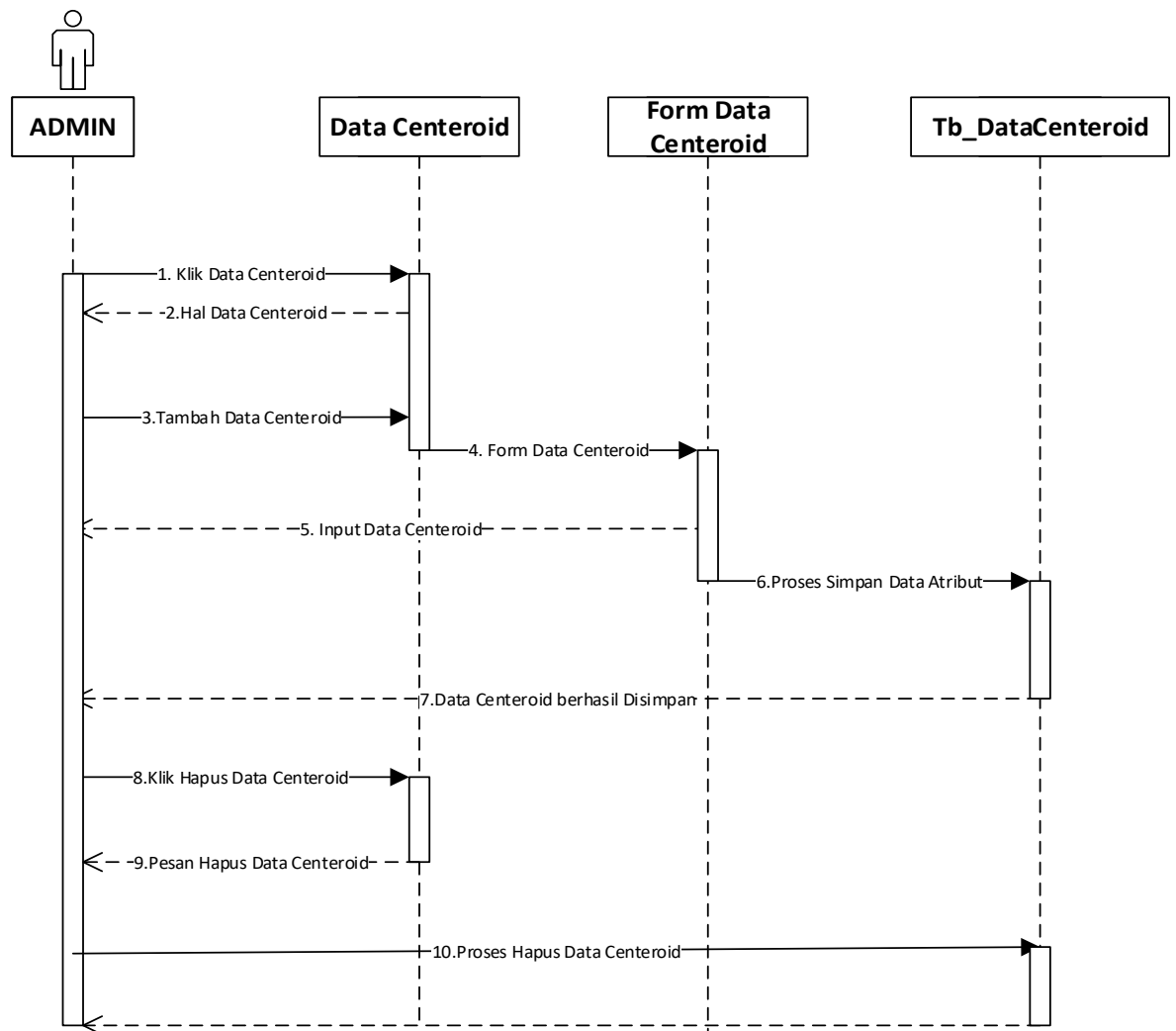
Gambar 4.9 UML Sequence Training Atribut

4.5.3 Data Training



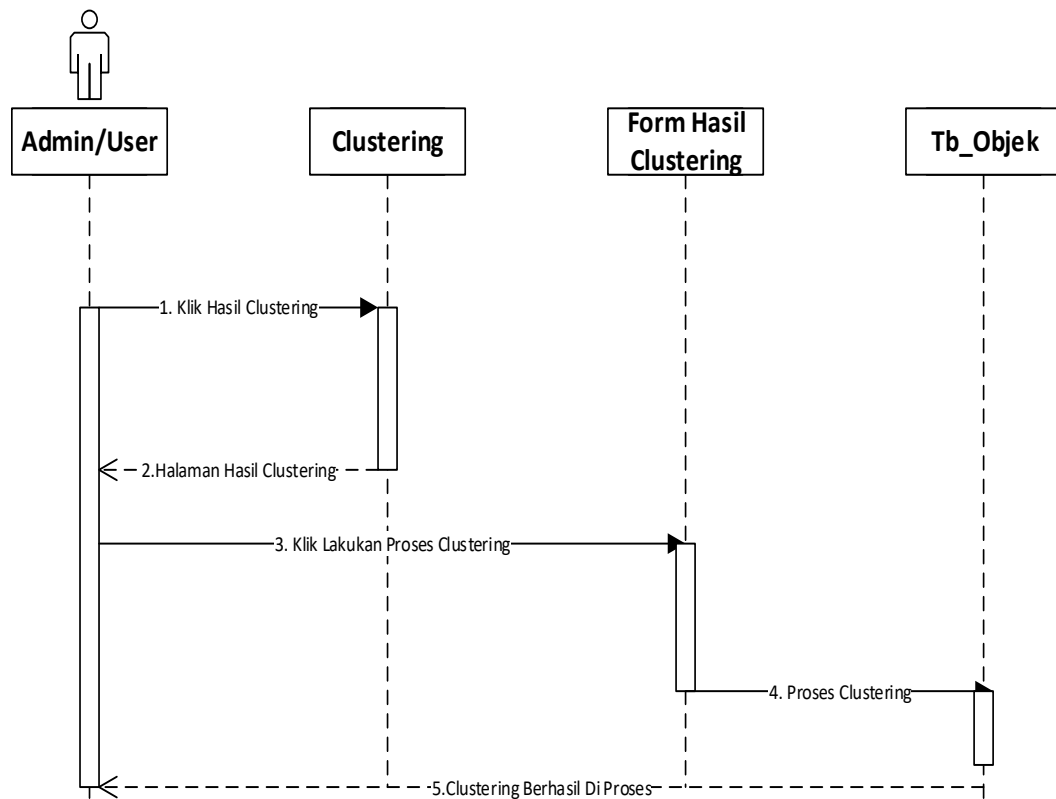
Gambar 4.10 UML sequence Data Training

4.5.4 Klasifikasi Data



Gambar 4.11 UML Sequence Klasifikasi Data

4.5.5 Hasil Klasifikasi



Gambar 4.12 UML Sequence Hasil Klasifikasi

4.6 Arsitektur Sistem

Sistem Klasifikasi penerima beasiswa menggunakan *client server*.

Sedangkan spesifikasi *hardware* dan *software* yang direkomendasikan, yaitu :

- | | |
|----------------------|--|
| 4.1 Processor | : Intel Core 2 duo |
| 4.2 RAM | : 2 GB |
| 4.3 VGA | : Nvidia GeForce GT 705 |
| 4.4 Haridisk | : 20GB |
| 4.5 Operating System | : Windows 7 |
| 4.6 Tools | : Rapid Miner, Visual Studio Code(Editor), dan MySQL |

4.7 Mekanisme User

Tabel 4.2 Interface design : Mekanisme User (Hasil Desain Sistem)

USER	KATEGORI	AKSES INPUT	AKSES OUTPUT
Admin	Administrator	All	All

4.8 Format File

Data yg diolah pada sistem klasifikasi penerima beasiswa ini menggunakan format :

1. Visual Studio Code Sebagai tempat penyimpanan eksternal
2. Xampp untuk mengolah dan menyimpan data secara temporary

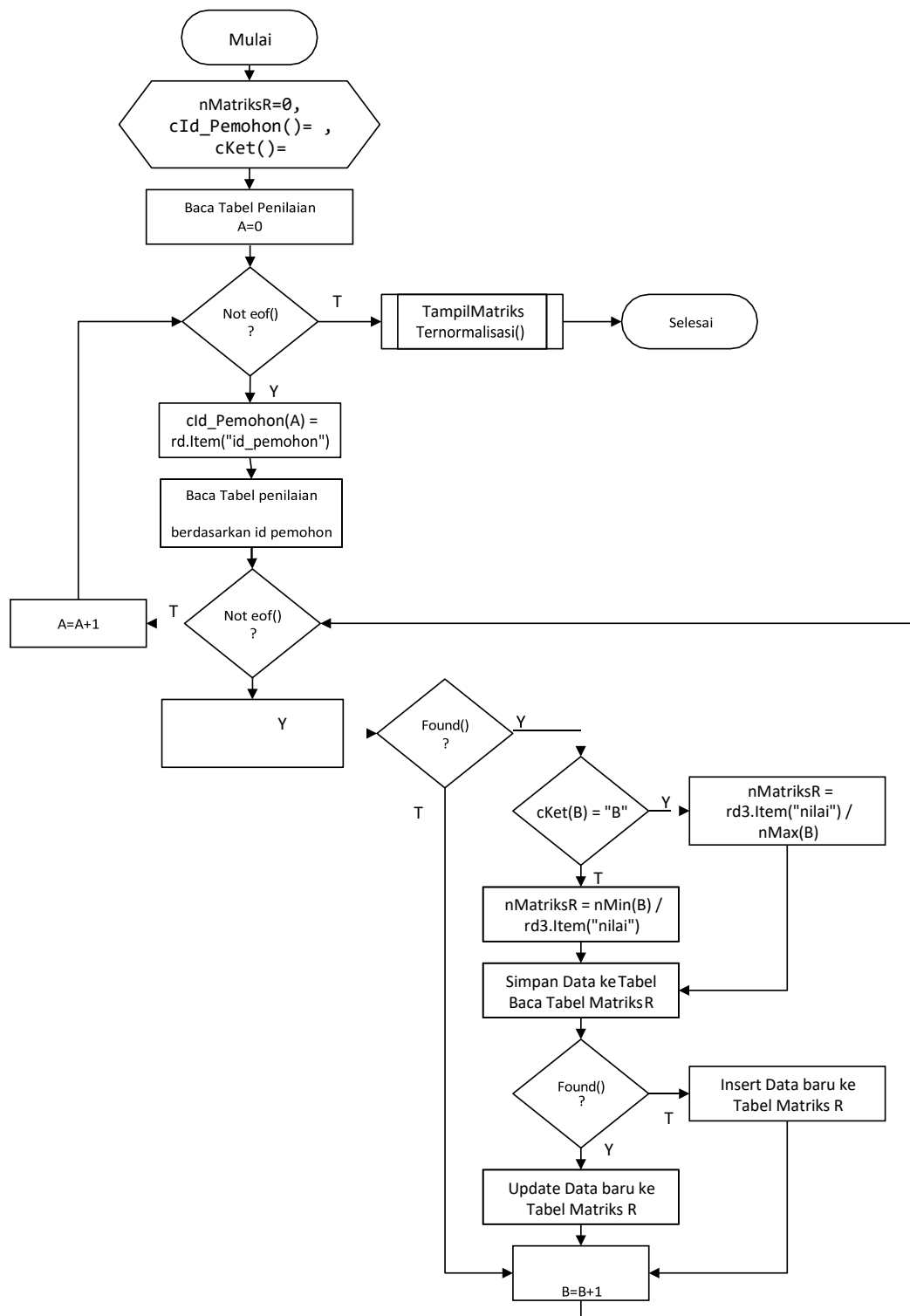
4.9 Program Desain

Tabel 4.3 Program Desain

CLASS/TYPE	ATTRIBUTES	METHODS
Main Menu	Admin [Menu]	Admin [Click]
	Laporan [Menu]	Laporan [Click]
	Data Training[Menu]	Data Training[Click]
	Data Testing[Menu]	Data Testing[Click]
	Data Master [Menu]	Data Master [Click]
	Home [Menu]	Home [Click]
Menu Login	ID[TextBox]	ID[Click]
	Username[Textbox]	Username[Click]
	Password[Textbox]	Password[Click]

	Login[BUTTON]	Login[Click]
Data Training	Data Sampel [Menu]	Data Sampel [Click]
	Probabilitas [Menu]	Probabilitas [Click]
Data Testing	Data Mahasiswa [Menu]	Data Mahasiswa [Click]
	Rekomendasi [Menu]	Rekomendasi[Click]

4.10 Flowchart Pengujian White Box



Gambar 4.13 Flowchart Pengujian White Box

4.11 Flowgraph Pengujian White Box

a. Menghitung Nilai *Cyclomatic Complexity* (CC)

Dimana :

$$\text{Region(R)} = 7$$

$$\text{Node(N)} = 16$$

$$\text{Edge(E)} = 21$$

$$\text{Predicate Node(P)} = 6$$

$$V(G) = E - N + 2$$

$$= 21 - 16 + 2$$

$$= 5 + 2$$

$$= 7$$

$$V(G) = P + 1$$

$$= 6 + 1$$

$$= 7$$

b. Menentukan Basis Path :

Basis set yang dihasilkan dari jalur independent path secara linier adalah jalur sebagai berikut :

Jalur 1 : 1-2-4-6-7-16

Jalur 2 : 1-2-3-2-4-6-7-16

Jalur 3 : 1-2-4-5-4-6-7-16

Jalur 4 : 1-2-4-6-7-8-9-11-5-7-16

Jalur 5 : 1-2-4-6-7-8-9-10-9-11-15-7-16

Jalur 6 : 1-2-4-6-7-8-9-11-12-14-11-15-7-16

Jalur 7 : 1-2-4-6-7-8-9-11-12-13-11-15-7-16

4.12 Pengujian Black Box

Tabel 4.4 Pengujian Black Box

Input/Event	Fungsi	Hasil yg Diharapkan	Hasil Uji
Input nama user dan password yg benar	Menampilkan halaman menu utama	Halaman menu utama tampil	Sesuai
Input nama user yg salah	Menampilkan pesan kesalahan “Maaf..., Username atau Password Salah ...”	Pesan kesalahan input nama user tampil	Sesuai
Input password yg salah	Menampilkan pesan kesalahan “Maaf..., Username atau Password Salah ...”	Pesan kesalahan input password tampil	Sesuai
Klik Data Training	Menampilkan Form Data Training	Halaman form Data Training	Sesuai
Klik Data Testing	Menampilkan form Data Testing	Halaman form data testing	Sesuai
Klik Klasifikasi	Menampilkan form Proses Klasifikasi	Halaman form klasifikasi	Sesuai
Klik Dashboard	Menampilkan form Menu Utama	Halaman form Menu Utama	Sesuai

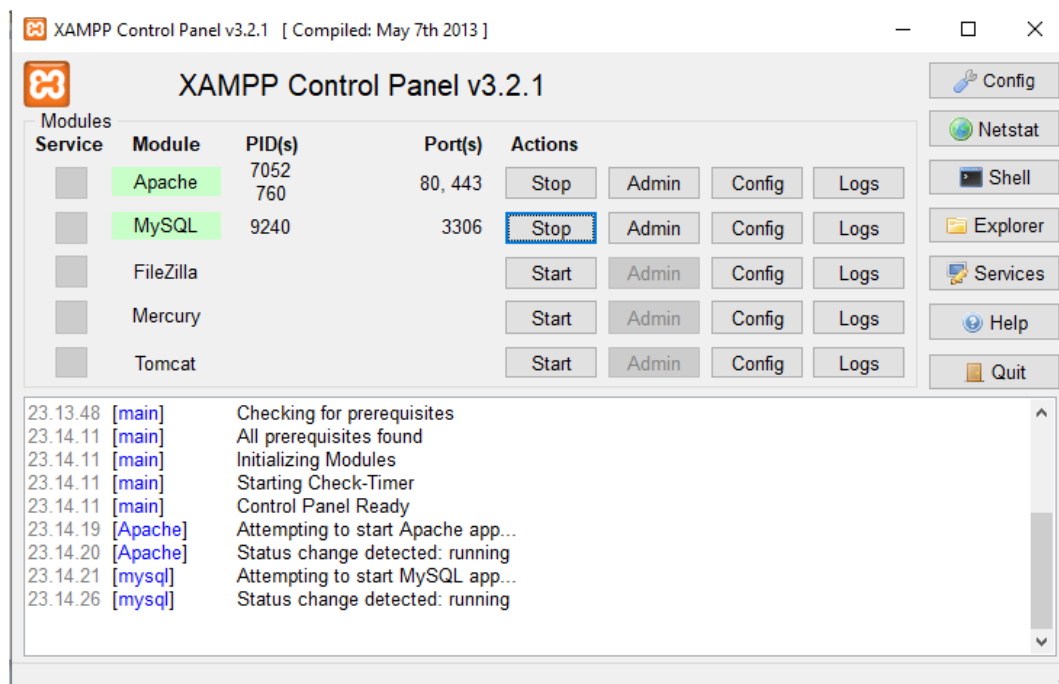


BAB V

PEMBAHASAN PENELITIAN

5.1 Prosedur Pengoperasian Sistem

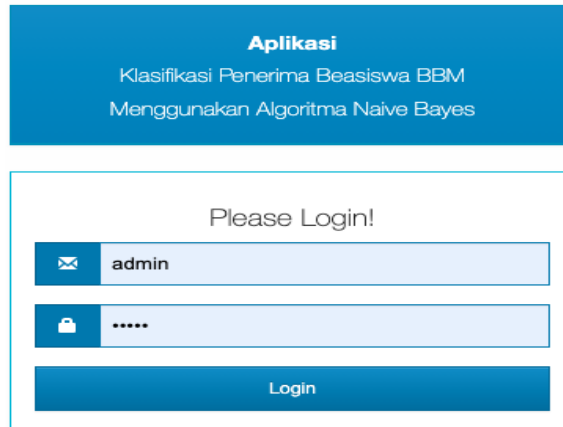
Untuk menjalankan program ini, buka program XAMPP terlebih dahulu. Klik start pada Apache dan MySQL



Gambar 5.1 XAMPP Control Panel

- **Halaman Login**

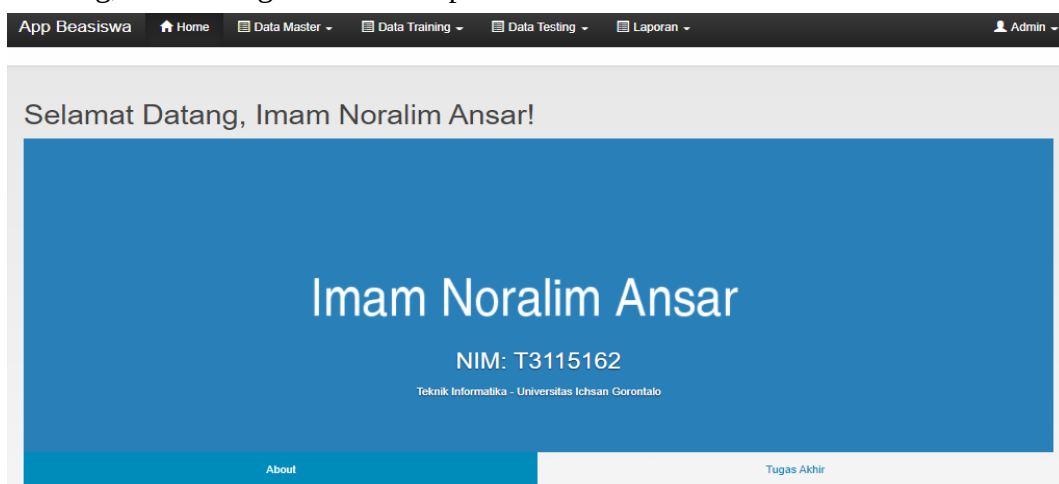
Pada tampilan halaman login ini, user menginput username dan password untuk masuk ke Menu Utama untuk memasukan. Apabila salah maka akan tampil pesan kesalahan input User ID dan password pada layar, kemudian ulangi lagi.



Gambar 5.2 Menu Login

- **Halaman Utama**

Halaman ini berfungsi untuk menampilkan seluruh menu utama yang terdapat pada Aplikasi Klasifikasi Beasiswa. Form ini terdiri atas menu-menu yang terdapat pada lajur atas, yang digunakan untuk menginput seluruh data-data peserta yang memasukkan data beasiswa. terdapat menu data master, menu data training, data testing dan menu Laporan



Gambar 5.3 Halaman Utaman

- **Menu Data Training**

Pada halaman data training ini digunakan untuk memasukkan data training berupa data mahasiswa yang memiliki atribut sesuai kriteria penerima beasiswa BBM.

Data Sample													
Data Sample													
Previous 1 2 3 Next													
#	NIM	Nama Lengkap	L/P	SMT	Fakultas	Prodi	IPK	Gaji Ortu	R. Listrik	Piagam	Prop. PKM	Beasiswa	Edit Hapus
1	2118017	FIRMAN ABDUL	L	5	TEKNIK	S1 ELEKTRO	2.42	8.071.801	143.722	0	TIDAK	TIDAK	 
2	2118020	PENEHAS EDOWAI	L	5	TEKNIK	S1 ELEKTRO	2.5	7.033.090	390.457	2	TIDAK	TIDAK	 
3	2118022	PRAMADYA HERRY MENIX SEMBIRING	L	5	TEKNIK	S1 ELEKTRO	2.45	7.177.981	464.352	2	TIDAK	TIDAK	 
4	2118023	RENDI ABD. LATIF PAKAYA	L	5	TEKNIK	S1 ELEKTRO	3.42	2.905.331	193.302	7	YA	BBM	 
5	2118024	ZIDANE FIERA GUSTIAN PAKAYA	L	5	TEKNIK	S1 ELEKTRO	3.61	3.243.983	281.931	7	YA	BBM	 
6	2118025	RIVAL S. ALAM	L	5	TEKNIK	S1 ELEKTRO	3.16	4.015.187	380.104	5	YA	BBM	 
7	2118026	MOH. ERDIYANSYAH	L	5	TEKNIK	S1 ELEKTRO	3.58	5.036.231	260.736	5	YA	BBM	 

Gambar 5.4 Menu Data Training

Untuk menambahkan data sample/training, pilih tombol tambah untuk melakukan penginputan.

+ Tambah Data Sample

NIM

:

Nama Lengkap

:

Fakultas

:

TEKNIK

▼

Prodi

:

S1 ELEKTRO

▼

IPK

:

Gaji Orang Tua

:

Semester

:

1

▼

Jenis Kelamin

:

Laki-laki

Perempuan

Rekening Listrik

:

Jumlah Piagam

:

Proposal PKM

:

Ya

Tidak

Beasiswa

:

TIDAK MENDAPAT

▼

Simpan

Batal

Gambar 5.5 Form Input Data Training

Setelah melakukan proses penginputan data training selanjutnya adalah menghitung probabilitas setiap kelas/atribut data, dengan memilih submenu probabilitas dan mengklik tombol proses untuk memperoleh hasil perhitungan probabilitasnya.

Probabilitas

Tabel Probabilitas			
Prodi	Semester	Jenis Kelamin	Indeks Prestasi
Gaji Orang Tua	Tagihan Listrik	Plagam	Proposal PKM
Warning: Division by zero in /Applications/XAMPP/xamppfiles/htdocs/beasiswanb/functions.php on line 51			
Warning: Division by zero in /Applications/XAMPP/xamppfiles/htdocs/beasiswanb/functions.php on line 51			
#	Indeks Prestasi	BBM	TIDAK
1	3.5	1/2	0/0
2	3.7	1/2	0/0
Tertinggi		1/2	0/0
Total		2	0
Mean (Rata-rata)		1,000	0,000
Standar Deviasi		0,000	0,000

Gambar 5.6 Menu Probabilitas

- **Menu Data Testing**

Pada halaman data testing ini digunakan untuk mengunggah data uji dari mahasiswa yang akan mengusulkan permohonan beasiswa BBM.

Data Mahasiswa

Data Mahasiswa													+ Tambah	▼ Filter
#	NIM	Nama Lengkap	L/P	SMT	Fakultas	Prodi	Alamat	IPK	Gaji Ortu	R. Listrik	Plagam	Prop. PKM	Edit	Hapus
1	3118165	MOH. ARIF N. TOLODO	L	3	ILMU KOMPUTER	S1 TEKNIK INFORMATIKA	JLN MANGGA	2.84	7.465.321	343.000	1	TIDAK		
2	1118034	ANAS KUDIAI	L	5	ILMU KOMPUTER	S1 TEKNIK INFORMATIKA	JLN DURIAN	2.22	7.500.000	230.000	2	TIDAK		
3	123827	MUIS NANJA	L	5	ILMU KOMPUTER	S1 TEKNIK INFORMATIKA	JLN MANGGA	3.55	3.000.000	250.000	5	YA		
4	3115783	IRVAN SALIHI	L	5	ILMU KOMPUTER	S1 TEKNIK INFORMATIKA	LIMBOTO	2.59	4.000.000	350.432	2	TIDAK		
5	321678	ZOHRAHAYATY	P	2	ILMU KOMPUTER	S1 TEKNIK INFORMATIKA	TELAGA	2.11	7.000.000	1.000.000	0	TIDAK		
6	1118025	RINDY ZURIATY PUTRI KUNU	P	5	HUKUM	S1 ILMU HUKUM	JLN MANGGA	2.79	6.203.747	444.996	7	YA		
7	3118003	REFENDI HUSIN	L	5	HUKUM	S1 ILMU HUKUM	JLN DURIAN	3.18	2.676.000	307.674	5	YA		

Gambar 5.7 Halaman Data Testing

Terdapat submenu data mahasiswa untuk menambahkan mengetahui apakah mahasiswa tersebut mendapatkan rekomendasi layak untuk memperoleh beasiswa dengan memasukkan atribut sesuai kriteria penerima beasiswa.

Tambah Data Mahasiswa

+ Tambah Data Mahasiswa

Warning: Creating default object from empty value in C:\xampp\htdocs\beasiswa\page-data-mahasiswa.php on line 14

NIM :

Nama Lengkap :

Fakultas :

Prodi :

Alamat :

IPK :

Gaji Orang Tua :

Semester :

Jenis Kelamin : ☒ Laki-laki ☐ Perempuan

Rekening Listrik :

Jumlah Piagam :

Proposal PKM : ☐ Ya ☒ Tidak

Gambar 5.8 Form Input Data Testing

- **Menu Rekomendasi**

Pada menu rekomendasi akan ditampilkan hasil klasifikasi penerima beasiswa apakah berhak untuk direkomendasikan untuk memperoleh beasiswa.

Rekomendasi

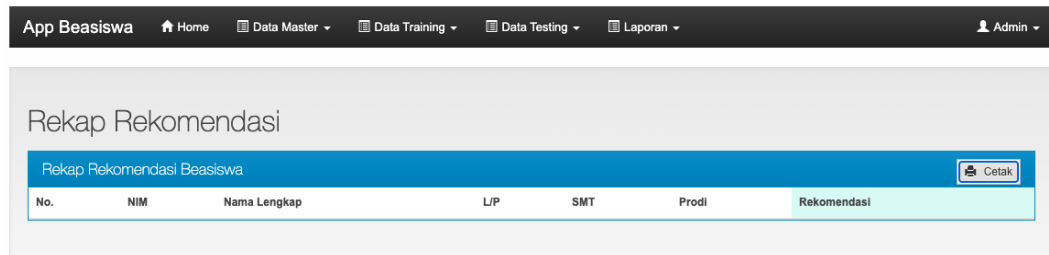
Rekomendasi Beasiswa Filter Process

#	NIM	Nama Lengkap	L/P	SMT	Prodi	IPK	Gaji Ortu	R. Listrik	Piagam	Prop. PKM	Rekomendasi	Detail
---	-----	--------------	-----	-----	-------	-----	-----------	------------	--------	-----------	-------------	--------

Gambar 5.9 Halaman Proses Rekomendasi (klasifikasi)

- **Halaman Laporan**

Halaman ini digunakan untuk merekap hasil proses perhitungan klasifikasi menggunakan algoritma Naive Bayes bagi penerima beasiswa BBM dengan rekomendasi layak atau tidak layak.



Gambar 5.10 Halaman Laporan

5.2 Penerapan Algoritma Naive Bayes

Tahap awal cara kerja dari proses perhitungan Naive Bayes adalah dengan melakukan pengambilan data training dari data mahasiswa. Klasifikasi mahasiswa apabila diberi input berupa jenis kelamin, IPK, SKS Lulus, Dan Gaji orang tua

- Menghitung jumlah class/label dari jenis kelamin
 $P(P=\text{Jenis Kelamin}=\text{Laki-laki}|C=\text{BBM})=15/20$ "Jumlah data bbm pada data pelatihan dibagi dengan jumlah keseluruhan data"
 $P(P=\text{Jenis Kelamin}=\text{Laki-laki}|C=\text{Tidak Mendapatkan})=5/20$ "Jumlah data tidak mendapatkan pada data pelatihan dibagi dengan jumlah keseluruhan data"
- Menghitung jumlah class/label dari IPK
 $P(P=\text{IPK}=3.00|C=\text{BBM})=10/20$ "Jumlah data BBM pada data pelatihan dibagi dengan jumlah keseluruhan data"
- Menghitung jumlah class/label dari SKS
 $P(P=\text{SKS}=92|C=\text{BBM})=13/20$ "Jumlah data BBM pada data pelatihan dibagi dengan jumlah keseluruhan data"

- Menghitung jumlah class/label dari Gaji Orang Tua

$P(P=SKS=3.000.000|C=BBM)=11/20$ "Jumlah data BBM pada data pelatihan dibagi dengan jumlah keseluruhan data"

$P(P=SKS=3.000.000|C=Tidak Mendapatkan)=9/20$ "Jumlah data tidak mendapatkan pada data pelatihan dibagi dengan jumlah keseluruhan data"

BAB VI

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Dari hasil penelitian dengan program untuk Pemberian Beasiswa BBM di Universitas Ichsan Gorontalo menggunakan metode Naive Bayes, maka pada akhir laporan penelitian ini penulis menyimpulkan bahwa :

1. Pemodelan algoritma Naive Bayes untuk klasifikasi penerima beasiswa memiliki kinerja yang baik dan dapat diterapkan untuk penentuan/klasifikasi penerima beasiswa BBM;
2. Aplikasi klasifikasi penerima beasiswa BBM yang diuji kinerjanya dengan metode White Box Testing menghasilkan $V(G) = CC = 18$, di mana setiap jalurnya dapat dieksekusi, sehingga dinyatakan bahwa sistem ini telah memenuhi syarat logika pemrograman dan tidak kompleks. Sedangkan dengan pengujian Black Box menyatakan bahwa sistem telah bebas dari berbagai kesalahan komponen-komponennya. Dengan demikian, dapat ditarik kesimpulan bahwa diperoleh aplikasi klasifikasi penerima beasiswa BBM ns yang handal sehingga dapat diimplementasikan.

6.2 Saran

Berdasarkan kesimpulan laporan tersebut, peneliti dapat memberikan saran untuk peneliti selanjutnya, yaitu:

1. Peneliti selanjutnya dapat mengoptimalkan metode naive bayes dengan menambah jumlah data agar supaya menghasilkan hasil yang lebih baik.
2. Dapat dikembangkan lagi dengan menambahkan beberapa variabel untuk penerima beasiswa BBM di Universitas Ichsan Gorontalo.

DAFTAR PUSKATA

- [1] E. Hasmin and S. Aisa, "Penerapan Algoritma C4.5 Untuk Penentuan Penerima Beasiswa Mahasiswa," *CogITO Smart J.*, vol. 5, no. 2, p. 308, 2019, doi: 10.31154/cogito.v5i2.219.308-320.
- [2] D. P. dan Kemahasiswaan, *Pedoman Beasiswa dan Bantuan Biaya Pendidikan PPA 2014*. 2014.
- [3] D. Dahri, F. Agus, and D. M. Khairina, "Metode Naive Bayes Untuk Penentuan Penerima Beasiswa Bidikmisi Universitas Mulawarman," *Inform. Mulawarman J. Ilm. Ilmu Komput.*, vol. 11, no. 2, p. 29, 2016, doi: 10.30872/jim.v11i2.211.
- [4] M. Tabrani, "Klasifikasi Penerima Beasiswa Kopertis Dengan Menggunakan Algoritma C4.5," *Pilar Nusa Mandiri*, vol. XII, no. 1, pp. 72–80, 2016.
- [5] S. L. Ting, W. H. Ip, and A. H. C. Tsang, "Is Naïve bayes a good classifier for document classification?," *Int. J. Softw. Eng. its Appl.*, vol. 5, no. 3, pp. 37–46, 2011.
- [6] A. Saleh, "Penerapan Data Mining Dengan Metode Klasifikasi Naive Bayes Untuk Memprediksi Kelulusan Mahasiswa Dalam Mengikuti English Proficiency Test (Studi Kasus : Universitas Potensi Utama)," *Konf. Nas. Sist. Informasi, Univ. Klabat, Manado, Indonesia, Vol. 2015*, no. June, pp. 1–6, 2015.
- [7] K. Neighbor, C. R. Raharya, N. Hidayat, and E. Santoso, "Penentuan Penerimaan Beasiswa Menggunakan Metode Modified," *J. Pengemb. Teknol. Inf. dan Ilmu Komput. Univ. Brawijaya*, vol. 2, no. 11, pp. 4984–4990, 2018.
- [8] D. Salmu, S. and A. Solichin, "Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu Menggunakan Naïve Bayes :," *Pros. Semin. Nas. Multidisiplin Ilmu*, no. April, pp. 701–709, 2017.
- [9] I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques*, 4th ed. San

Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2016.

- [10] J. Han and M. Kamber, *Data Mining : Concepts and Techniques Second Edition*, 2nd ed. Morgan Kauffman Publisher, 2006.
- [11] D. Soria, J. M. Garibaldi, F. Ambrogi, E. M. Biganzoli, and I. O. Ellis, “A ‘non-parametric’ version of the naive Bayes classifier,” *Knowledge-Based Syst.*, vol. 24, no. 6, pp. 775–784, 2011, doi: <https://doi.org/10.1016/j.knosys.2011.02.014>.
- [12] I. Sommerville, *Software Engineering*, 9th ed. USA: Addison-Wesley Publishing Company, 2010.
- [13] R. Pressman, *Software Engineering: A Practitioner’s Approach*, 7th ed. USA: McGraw-Hill, Inc., 2009.
- [14] A. Dennis, B. H. Wixom, and D. Tegarden, *Systems Analysis and Design: An Object-Oriented Approach with UML*, 5th ed. Wiley Publishing, 2015.



LAMPIRAN 1 : KODE PROGRAM NAÏVE BAYES

```

<?php
require_once '../naivebayes/vendor/autoload.php';
include '../libs/db_connection.php';
use Phpml\Classification\NaiveBayes;

$sql1 = "SELECT ipk,sks_lulus,penghasilan_orangtua,tanggungan FROM training ";
$result1 = $connect->query($sql1);
$sql2 = "SELECT status FROM training ";
$result2 = $connect->query($sql2);
$training = [];
$training_status = [];
if ($result1->num_rows > 0) {
    while ($row = $result1->fetch_assoc()) {
        $training[] = $row;
    }
}
if ($result2->num_rows > 0) {
    while ($row = $result2->fetch_assoc()) {
        $training_status[] = $row["status"];
    }
}
$samples = [];
foreach ($training as $t) {
    array_push($samples, [$t["ipk"], $t["sks_lulus"], $t["penghasilan_orangtua"], $t["tanggunan"]]);
}
$classifier = new NaiveBayes();
$classifier->train($samples, $training_status);
$sql3 = "SELECT ipk,sks_lulus,penghasilan_orangtua,tanggungan FROM testing ";
$result3 = $connect->query($sql3);
$testing = [];
if ($result3->num_rows > 0) {
    while ($row = $result3->fetch_assoc()) {
        $testing[] = $row;
    }
}
$test = [];
foreach ($testing as $ta) {
    array_push($test, sprintf('%s', $classifier->predict([$ta["ipk"], $ta["sks_lulus"], $ta["penghasilan_orangtua"], $ta["tanggunan"]])));
}
$result = mysqli_query($connect, "SELECT COUNT(id) AS asd FROM testing");
$row = mysqli_fetch_array($result);
$total = $row[0];
$n = 1;

```



```

for ($i = 0; $i < $total; $i++) {
    $query = "INSERT INTO klasifikasi (npm, nama, jenis_kelamin, tgl_lahir, angkatan, fakultas, prodi, periode,
ipk, sks_lulus, penghasilan_orangtua, tanggungan, actual, predict) VALUES ((SELECT npm FROM testing
WHERE id='$n'),(SELECT nama FROM testing WHERE id='$n'),(SELECT jenis_kelamin FROM testing
WHERE id='$n'),(SELECT tgl_lahir FROM testing WHERE id='$n'),(SELECT angkatan FROM testing WHERE
id='$n'),(SELECT fakultas FROM testing WHERE id='$n'),(SELECT prodi FROM testing WHERE
id='$n'),(SELECT periode FROM testing WHERE id='$n'),(SELECT ipk FROM testing WHERE
id='$n'),(SELECT sks_lulus FROM testing WHERE id='$n'),(SELECT penghasilan_orangtua FROM testing
WHERE id='$n'),(SELECT tanggungan FROM testing WHERE id='$n'),(SELECT status FROM testing WHERE
id='$n'),'test[$i]')";

    $result = mysqli_query($connect, $query);
    $n++;
}
header("location:../template.php?page=klasifikasi")

```

LAMPIRAN 2 : SURAT KETERANGAN PENELITIAN



**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
LEMBAGA PENELITIAN (LEMLIT)
UNIVERSITAS ICHSAN GORONTALO**

Jl. Raden Saleh No. 17 Kota Gorontalo
Telp: (0435) 8724466, 829975; Fax: (0435) 82997;
E-mail: lembagapenelitian@unisan.ac.id

SURAT KETERANGAN

Nomor : 2335/SK/LEMLIT-UNISAN/GTO/XI/2020

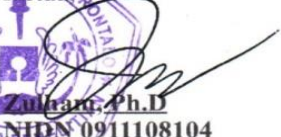
Yang bertanda tangan di bawah ini :

Nama : Zulham, Ph.D
NIDN : 0911108104
Jabatan : Ketua Lembaga Penelitian

Menerangkan bahwa :

Nama Mahasiswa : Imam Noralim Ansar
NIM : T3115162
Fakultas : Fakultas Ilmu Komputer
Program Studi : Teknik Informatika
Judul Penelitian : KLASIFIKASI PENERIMA BEASISWA
MENGGUNAKAN ALGORITMA NAIVE BAYES
CLASSIFIER

Adalah benar telah melakukan pengambilan data penelitian dalam rangka Penyusunan Proposal/Skripsi pada KEPALA BAAK UNIVERSITAS ICHSAN GORONTALO.

Gorontalo, 30 November 2020
Ketua

Zulham, Ph.D
NIDN 0911108104

+

LAMPIRAN 3 : SURAT REKOMENDASI BEBAS PLAGIASI



**KEMENTERIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS ICHSAN
(UNISAN) GORONTALO**

SURAT KEPUTUSAN MENDIKNAS RI NOMOR 84/D/O/2001
Jl. Achmad Nadjamuddin No. 17 Telp (0435) 829975 Fax (0435) 829976 Gorontalo

SURAT REKOMENDASI BEBAS PLAGIASI

No. 0706/UNISAN-G/S-BP/XII/2020

Yang bertanda tangan di bawah ini :

Nama : Sunarto Taliki, M.Kom
NIDN : 0906058301
Unit Kerja : Pustikom, Universitas Ichsan Gorontalo

Dengan ini Menyatakan bahwa :

Nama Mahasisw : IMAM NORALIM ANSAR
NIM : T3115162
Program Studi : Teknik Informatika (S1)
Fakultas : Fakultas Ilmu Komputer
Judul Skripsi : KLASIFIKASI PENERIMA BEASISWA
MENGUNAKAN ALGORITMA NAIVE BAYES
CLASSIFIER

Sesuai dengan hasil pengecekan tingkat kemiripan skripsi melalui aplikasi Turnitin untuk judul skripsi di atas diperoleh hasil Similarity sebesar 34%, berdasarkan SK Rektor No. 237/UNISAN-G/SK/IX/2019 tentang Panduan Pencegahan dan Penanggulangan Plagiarisme, bahwa batas kemiripan skripsi maksimal 35% dan sesuai dengan Surat Pernyataan dari kedua Pembimbing yang bersangkutan menyatakan bahwa isi softcopy skripsi yang diolah di Turnitin SAMA ISINYA dengan Skripsi Aslinya serta format penulisannya sudah sesuai dengan Buku Panduan Penulisan Skripsi, untuk itu skripsi tersebut di atas dinyatakan BEBAS PLAGIASI dan layak untuk diujikan.

Demikian surat rekomendasi ini dibuat untuk digunakan sebagaimana mestinya.

Gorontalo, 05 Desember 2020
Tim Verifikasi,



Sunarto Taliki, M.Kom
NIDN. 0906058301

Tembusan :

1. Dekan
2. Ketua Program Studi
3. Pembimbing I dan Pembimbing II
4. Yang bersangkutan
5. Arsip

LAMPIRAN 4 : HASIL TURNITIN



SKRIPSI_1_T3115162_IMAM NORALIM ANSAR.docx

Dec 5, 2020

6905 words / 43200 characters

T3115162 IMAM NORALIM ANSAR

KLASIFIKASI PENERIMA BEASISWA MENGGUNAKAN ALGORI...

Sources Overview

34%

OVERALL SIMILARITY

1	www.scribd.com	8%
	INTERNET	
2	docplayer.info	7%
	INTERNET	
3	informatikalogi.com	3%
	INTERNET	
4	srs30.blogspot.com	2%
	INTERNET	
5	repository.upi.edu	2%
	INTERNET	
6	Mukhamad Masrur. "ANALISIS DAN PENGEMBANGAN SISTEM INFORMASI AKADEMIK (SIKAD) DI UNIPDU JOMBANG", Teknologi, 2012	1%
	CROSSREF	
7	repository.its.ac.id	1%
	INTERNET	
8	pt.scribd.com	<1%
	INTERNET	
9	www.studineews.co.id	<1%
	INTERNET	
10	qdoc.tips	<1%
	INTERNET	
11	www.warse.org	<1%
	INTERNET	
12	zombiedoc.com	<1%
	INTERNET	
13	docobook.com	<1%
	INTERNET	
14	ejurnal.stmik-budidarma.ac.id	<1%
	INTERNET	
15	www.slideshare.net	<1%
	INTERNET	
16	andikhalilgibran.blogspot.com	<1%
	INTERNET	

17	ejournal.uin-suka.ac.id	INTERNET	<1%
18	jurnalsaintek.uinsby.ac.id	INTERNET	<1%
19	moam.info	INTERNET	<1%
20	ncunk.blogspot.com	INTERNET	<1%
21	www.neliti.com	INTERNET	<1%
22	Akhmad Pandhu Wijaya, Heru Agus Santoso. "Improving The Accuracy of Naïve Bayes Algorithm for Hoax Classification Using Particl...	CROSSREF	<1%
23	dome.umm.ac.id	INTERNET	<1%
24	repository.uinjkt.ac.id	INTERNET	<1%
25	cogito.unklab.ac.id	INTERNET	<1%
26	media.neliti.com	INTERNET	<1%
27	de.scribd.com	INTERNET	<1%
28	Herlawati Herlawati, Rahmadya Trias Handayanto. "Penggunaan Matlab dan Python dalam Klasterisasi Data", Jurnal Kajian Ilmiah, 20...	CROSSREF	<1%
29	www.niagahoster.co.id	INTERNET	<1%

Excluded search repositories:

- Submitted Works

Excluded from Similarity Report:

- Small Matches (less than 25 words).

Excluded sources:

- None

LAMPIRAN 5 : SURAT KETERANGAN BEBAS PUSTAKA



**KEMENTRIAN PENDIDIKAN DAN KEBUDAYAAN
UNIVERSITAS ICHSAN GORONTALO**

UPT. PERPUSTAKAAN PUSAT

SURAT KEPUTUSAN MENDIKS RI NO. 84/D/0/2001

Jln. Achmad Nadjamuddin No.17 Telp. (0435) 829975 Fax. (0435) 829976 Gorontalo

SURAT KETERANGAN BEBAS PUSTAKA

No.002/perpus_fikom/XI/2020

Perpustakaan Selasa, 24 November 2020 Fakultas Ilmu komputer (FIKOM) Universitas Ichsan Gorontalo dengan ini menerangkan bahwa :

Nama : Imam Noralim Ansar

Nim : T3115162

No anggota : M202002

Terhitung sejak tanggal 24 November 2020, dinyatakan telah bebas dari pinjaman buku dan koleksi lainnya dipergustakaan Fakultas Ilmu komputer.

Demikian keterangan ini di buat untuk di pergunakan sebagaimana mestinya.

Gorontalo, 24 November 2020

Kepala Perpustakaan
Fakultas Ilmu Komputer



Apriyanto Alhamad, M.Kom
NIDN 09240486

LAMPIRAN 6 : RIWAYAT HIDUP PENELITI

Nama : Imam Noralim Ansar
NIM : T3115162
Tempat/Tgl Lahir : Bone, 18 Agustus 1993
Pekerjaan : Mahasiswa
Agama : Islam



Riwayat Pendidikan :

1. Tahun 2006, menyelesaikan Pendidikan di Madrasah Ibtidaiyah Muhammadiyah Unggulan Kota Gorontalo
2. Tahun 2009, menyelesaikan Pendidikan di Madrasah Tsanawiyah Negeri Gorontalo
3. Tahun 2012, menyelesaikan Pendidikan di Sekolah Menengah Kejuruan Negeri 1 Kota Gorontalo
4. Tahun 2015, telah diterima menjadi Mahasiswa di Perguruan Tinggi Swasta Universitas Ichsan Gorontalo