

PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN ALGORITMA C4.5

(Studi Kasus : Fakultas Teknologi Informasi Universitas Bale Bandung)

SKRIPSI

Karya Tulis sebagai syarat memperoleh
Gelar Sarjana Komputer dari Fakultas Teknologi Informasi
Universitas Bale Bandung

Disusun oleh :

MUHAMAD GHASA SAPUTRA

NIM. C1A150002



PROGRAM STRATA 1
PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS TEKNOLOGI INFORMASI
UNIVERSITAS BALE BANDUNG
BANDUNG

2019

LEMBAR PERNYATAAN KEASLIAN

Saya yang bertanda tangan dibawah ini:

Nama : MUHAMAD GHASA SAPUTRA

NIM : C1A150002

Judul Skripsi : **PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN ALGORITMA C4.5 (Studi Kasus: Fakultas Teknologi Informasi Universitas Bale Bandung)**

Menyatakan dengan sebenarnya bahwa penulisan skripsi ini berdasarkan hasil penelitian, pemikiran dan pemaparan asli dari saya sendiri, baik untuk naskah laporan maupun kegiatan perhitungan yang tercantum sebagai bagian dari skripsi ini. Jika terdapat karya orang lain, saya mencantumkan sumber yang jelas.

Pernyataan ini saya buat dengan sesungguhnya dan apabila dikemudian hari terdapat penyimpangan dan ketidak benaran dalam pernyataan ini, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh karena karya tulis ini dan sanksi lain sesuai dengan peraturan yang berlaku di FAKULTAS TEKNOLOGI INFORMASI UNIVERSITAS BALE BANDUNG.

Demikian surat pernyataan ini saya buat dalam keadaan sadar tanpa paksaan dari pihak manapun.

Baleendah, Agustus 2019

Yang membuat pernyataan,

MUHAMAD GHASA SAPUTRA
NIM. C1A150002

ABSTRACT

In everyday life humans can produce a lot of data, one of them is in the world of education where a university can produce a lot of data about students such as student GPA data, student origin school data, alumni data, data on the number of male students and the number of female students, and data from the area of students. From the amount of data generated, the data can be utilized using data mining techniques and Algoritma C4.5. With Data Mining, each data set or data warehouse can provide important knowledge that can be used as valuable information for the benefit of the university. By using C4.5 Algorithm, the data can be processed, one of them is by developing a system that has the ability to see the pattern of student graduation rates, so that it can later become a strategy in the lecture process. In research conducted at the Information Technology Faculty of Bale Bandung University, the data to be used is Student data in the form of Year, Semester, Student's Number, Name, Study Program, GPA, and Student TOEFL Value. Where the data will be processed or extracted into a decision tree to predict student graduation rates on time and not on time. After the formation of the decision tree has been made then measuring the performance of the decision tree using a confusion matrix to improve the accuracy of the decision tree. From the measurement of the performance of the decision tree that has been made, will produce values of precision, recall, and accuracy, then each value of the confusion matrix will be averaged according to the results that have been made using a confusion matrix. Next, an application is made to predict the level of graduation of students who in the source code represent the rules in the c4.5 algorithm.

Keywords: C4.5 Algorithm, Students

ABSTRAK

Dalam kehidupan sehari-hari manusia dapat menghasilkan banyak sekali data, salah satunya dalam dunia pendidikan dimana suatu universitas dapat menghasilkan banyak sekali data tentang mahasiswa seperti, data IPK mahasiswa, data sekolah asal mahasiswa, data alumni, data jumlah mahasiswa berjenis kelamin laki-laki dan jumlah mahasiswa berjenis kelamin perempuan, dan data asal daerah mahasiswa. Dari banyaknya data yang dihasilkan, data tersebut dapat dimanfaatkan dengan menggunakan teknik data mining dan Algoritma C4.5. Dengan Data Mining, setiap kumpulan data atau gudang data dapat memberikan pengetahuan penting yang dapat dijadikan sebuah informasi yang berharga bagi kepentingan universitas. Dengan menggunakan Algoritma C4.5 data tersebut dapat diolah, salah satunya dengan menyusun sistem yang mempunyai kemampuan melihat pola tingkat kelulusan mahasiswa, untuk selanjutnya bisa menjadi strategi dalam proses perkuliahan. Dalam penelitian yang dilakukan di Fakultas Teknologi Informasi Universitas Bale Bandung data yang akan digunakan berupa data Mahasiswa berupa Tahun, Semester, Nim, Nama, Prodi, Nilai IPK, dan Nilai TOEFL mahasiswa. Dimana data tersebut akan diolah atau di *extract* untuk dijadikan suatu pohon keputusan untuk memprediksi tingkat kelulusan mahasiswa tepat waktu dan tidak tepat waktu. Setelah pembentukan pohon keputusan telah dibuat kemudian dilakukan pengukuran kinerja dari pohon keputusan menggunakan *confusion matrix* untuk meningkatkan akurasi dari pohon keputusan. Dari pengukuran kinerja pohon keputusan yang telah dibuat, akan menghasilkan nilai *precision*, *recall*, dan akurasi, kemudian tiap-tiap nilai dari *confusion matrix* akan dirata-ratakan sesuai dengan hasil yang telah dilakukan menggunakan *confusion matrix*. Selanjutnya dibuatkan suatu aplikasi untuk memprediksi tingkat kelulusan mahasiswa yang didalam *source code* nya merepresentasikan aturan-aturan dalam algoritma c4.5.

Kata Kunci: Algoritma C4.5, Mahasiswa

KATA PENGANTAR

Bismillahirrahmanirrahim

Assalamu'alaikum Warahmatullahi Wabarakatuh.

Puji dan syukur penulis panjatkan kehadirat Allah SWT serta Nabi Besar Muhammad SAW, yang mana berkat rahmat dan hidayah-Nya penulis dapat menyelesaikan Laporan Penelitian ini yang berjudul “PREDIKSI KELULUSAN MAHASISWA MENGGUNAKAN ALGORITMA C4.5 (Studi Kasus: Fakultas Teknologi Informasi Universitas Bale Bandung)” sebagai salah satu syarat untuk dapat melakukan penelitian dan mendapat gelar Sarjana di Fakultas Teknologi Informasi Universitas Bale Bandung.

Laporan Penelitian disusun dengan maksimal dan mendapatkan bantuan dari berbagai pihak sehingga dapat memperlancar dalam proses pembuatan. Untuk itu kami menyampaikan banyak terimakasih kepada semua pihak yang telah berkontribusi dalam pembuatan laporan ini.

Terlepas dari semua itu, kami menyadari sepenuhnya bahwa masih ada kekurangan baik dari segi susunan kalimat maupun tata bahasanya. Oleh karena itu dengan tangan terbuka kami menerima segala saran dan kritik dari pembaca agar kami dapat memperbaiki penelitian ini.

Penulis juga mengucapkan terima kasih kepada semua pihak yang terlibat dalam pembuatan laporan ini. Demikian, semoga laporan ini dapat bermanfaat. Terima kasih

Wassalamualaikum Warahmatullahi Wabarakatuh.

Bandung, Agustus 2019

Penyusun

DAFTAR ISI

<i>ABSTRACT</i>	v
ABSTRAK	vi
KATA PENGANTAR	vii
DAFTAR ISI.....	viii
DAFTAR GAMBAR	xi
DAFTAR TABEL.....	xii
BAB I PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah.....	2
1.4 Tujuan.....	3
1.5 Metode Penelitian.....	3
1.6 Sistematika Penulisan.....	4
BAB II LANDASAN TEORI	6
2.1 Landasan Teori	6
2.1.1 Jurnal 1.....	6
2.1.2 Jurnal 2.....	7
2.1.3 Jurnal 3.....	7
2.1.4 Jurnal 4.....	8
2.2 Dasar Teori	8

2.2.1 Prediksi	8
2.2.2 Kelulusan	10
2.2.3 Mahasiswa Lulus	15
2.2.4 Data Mining	15
2.2.5 Algoritma C4.5	25
2.2.6 Aplikasi	27
2.2.7 Web	27
2.2.8 <i>My Structured Query Language (MySQL)</i>	28
2.2.9 <i>Hypertext Preprocessor (PHP)</i>	30
2.2.10 <i>Unified Modeling Language (UML)</i>	30
2.2.11 <i>Model Driven Development (MDD)</i>	34
2.2.12 <i>Flowmap</i>	37
BAB III METODOLOGI PENELITIAN	39
3.1 Kerangka Berpikir	39
3.2 Penjelasan Skema Penelitian	40
3.2.1 Investigasi Pendahuluan	40
3.2.2 Analisa Masalah	41
3.2.3 Analisa Kebutuhan	42
3.2.4 Desain	43
3.2.5 Konstruksi	45
3.2.6 Implementasi	45
3.2.7 Pembuatan Laporan	46
BAB IV ANALISIS DAN PERANCANGAN	48
4.1 Analisis	48
4.1.1 Analisa Masalah	48

4.1.2 Analisa Sistem	48
4.1.3 Analisa Proses.....	50
4.1.4 Analisa Kebutuhan.....	50
4.1.5 Analisa Pengguna	52
4.1.6 <i>User Interface</i>	52
4.1.7 Fitur Pada Aplikasi	52
4.2 Perancangan Sistem.....	53
4.2.1 Perancangan.....	53
4.2.2 <i>Use Case Diagram</i>	63
4.2.3 <i>Class Diagram</i>	65
4.2.4 <i>Activity Diagram</i>	66
4.3 Perancangan <i>User Interface</i>	71
4.4 Perancangan <i>Database</i>	76
BAB V IMPLEMENTASI DAN PENGUJIAN	78
5.1 Implementasi	78
5.1.1 Implementasi Perangkat Keras dan Perangkat Lunak	78
5.2 Pengujian Sistem	79
BAB VI KESIMPULAN DAN SARAN	84
6.1 Kesimpulan.....	84
6.2 Saran.....	84
DAFTAR PUSTAKA	85
LAMPIRAN	86

DAFTAR GAMBAR

Gambar 2.1 Data Mining.....	15
Gambar 2.2 model driven development (mdd)	37
Gambar 3.3 Kerangka Berpikir	39
Gambar 4.4 Sistem yang diusulkan.....	49
Gambar 4.5 Konsep data decision tree.....	50
Gambar 4.6 pembentukan decision tree	60
Gambar 4.7 Use Diagram.....	63
Gambar 4.8 Class Diagram Aplikasi.....	65
Gambar 4.9 Activity Diagram Login	67
Gambar 4.10 Activity Diagram Kelola Data.....	68
Gambar 4.11 Activity Diagram Proses Perhitungan	69
Gambar 4.12 Activity diagram import data	70
Gambar 4.13 Activity Diagram Logout	71
Gambar 4.14 Perancangan form login	72
Gambar 4.15 Perancangan halaman user interface Dataset	73
Gambar 4.16 perancangan halaman nilai atribut.....	74
Gambar 4.17 Perancangan halaman perhitungan	75
Gambar 5.18 Tampilan halaman login.....	79
Gambar 5.19 Tampilan Halaman Dataset Mahasiswa	80
Gambar 5.20 Tampilan Halaman Nilai Atribut.....	82
Gambar 5.21 Tampilan halaman perhitungan	83

DAFTAR TABEL

Tabel 2.1 Data Matriks.....	24
Tabel 2.2 Data Transaksi.....	25
Tabel 2.3 Simbol Flowmap	38
Tabel 3.4 Spesifikasi sistem.....	43
Tabel 4.5 Kebutuhan Sistem	50
Tabel 4.6 Spesifikasi Perangkat Keras.....	51
Tabel 4.7 Spesifikasi Perangkat Lunak	51
Tabel 4.8 Klasifikasi Nilai IPK Tinggi	56
Tabel 4.9 Klasifikasi IPK Rendah.....	56
Tabel 4.10 Klasifikasi Nilai TOEFL menengah atas	57
Tabel 4.11 Klasifikasi Nilai TOEFL menengah bawah	57
Tabel 4.12 Entrophy total.....	58
Tabel 4.13 Pengukuran Kinerja Pohon Keputusan	61
Tabel 4.14 Pengukuran confusion matrix	62
Tabel 4.15 Hasil Pengukuran Kinerja Pohon Keputusan	62
Tabel 4.16 Hasil rata-rata kinerja pohon keputusan.....	62
Tabel 4.17 Use case aplikasi	64
Tabel 4.18 Keterangan form login	72
Tabel 4.19 keterangan halaman dataset	73
Tabel 4.20 Keterangan halaman nilai atribut	74
Tabel 4.21 Halaman Perhitungan.....	75
Tabel 4.22 Perancangan Database.....	76
Tabel 5.23 Spesifikasi Perangkat Keras.....	78
Tabel 5.24 Pengujian halaman login	80
Tabel 5.25 Pengujian Halaman Dataset	81
Tabel 5.26 Pengujian Halaman Nilai Atribut.....	82
Tabel 5.27 Pengujian Halaman Perhitungan	83

BAB I

PENDAHULUAN

1.1 Latar Belakang

Perkembangan teknologi informasi yang begitu maju saat ini, menyebabkan tingkat akurasi suatu data sangat dibutuhkan dalam kehidupan sehari-hari. Setiap informasi yang ada menjadi suatu hal penting untuk menentukan setiap keputusan dalam situasi tertentu. Hal ini menyebabkan penyediaan informasi menjadi sarana untuk dianalisa dan diringkas menjadi suatu pengetahuan dari data yang bermanfaat ketika pengambilan suatu keputusan dilakukan. (Nuqson Masyqur Huda, 2010).

Fakultas Teknologi Informasi Universitas Bale Bandung merupakan salah satu organisasi pendidikan yang bergerak di bidang pendidikan teknologi dan informasi. Informasi tingkat kelulusan mahasiswa Fakultas Teknologi Informasi Universitas Bale Bandung sangatlah penting untuk menentukan kelulusan mahasiswa sehingga dapat meningkatkan pelayanan yang dapat membuat mahasiswa nyaman dalam proses perkuliahan dan bisa lulus tepat waktu. Penggunaan Algoritma C4.5 dapat dijadikan pertimbangan dalam mengambil keputusan lebih lanjut tentang faktor yang mempengaruhi kelulusan mahasiswa.

Dikutip dari <https://forlap.ristekdikti.go.id> jumlah mahasiswa Fakultas Teknologi Informasi Universitas Bale Bandung berjumlah 220 mahasiswa untuk Program Studi Teknik Informatika pada semester genap 2018, Sedangkan untuk Program Studi Sistem Informasi berjumlah 52 mahasiswa untuk semester genap 2018 sehingga dari keseluruhan mahasiswa Fakultas Teknologi Universitas Bale Bandung pada semester genap 2018 berjumlah 272 mahasiswa.

Pada penelitian (Mujib Ridwan, Hadi Suyono, M Sarosa, 2013) menjelaskan bahwa faktor yang paling berpengaruh dalam penentuan klasifikasi kinerja akademik mahasiswa adalah Indeks Prestasi Kumulatif (IPK), Indeks Prestasi Semester (IPS) semester 1, semester 4 dan jenis kelamin. Pada penelitian berdasarkan *attribute* jenis kelamin, asal sekolah SMA dan IPS semester 1 sampai dengan IPS semester 6. Algoritma C4.5 merupakan algoritma klasifikasi pohon

keputusan yang banyak digunakan karena memiliki kelebihan utama dari algoritma yang lainnya. Kelebihan algoritma C4.5 dapat menghasilkan pohon keputusan yang mudah diinterpretasikan, memiliki tingkat akurasi yang dapat diterima, efisien dalam menangani *attribute* bertipe *diskret* dan *numerik*. Dalam mengkonstruksi pohon keputusan, algoritma C4.5 membaca sample data *training* dari *storage* dan memuatnya ke memori. Hal inilah yang menjadi salah satu kelemahan algoritma C4.5 dalam kategori ‘skalabilitas’ adalah algoritma ini dapat digunakan jika data *training* dapat disimpan secara keseluruhan dan pada waktu yang bersamaan pada memori.

Oleh karena itu penyusun pada penelitian ini mengembangkan teknik algoritma C4.5 untuk memprediksi tingkat kelulusan mahasiswa dengan menggunakan data mahasiswa Fakultas Teknologi Informasi Universitas Bale Bandung sebagai objek penelitian. Dimana pada hasil akhir penelitian akan menampilkan tahapan-tahapan dalam memanfaatkan data dan mengolah data menjadi sebuah informasi yang berguna.

1.2 Rumusan Masalah

Dari latar belakang yang telah disampaikan pada penelitian ini, Adapun beberapa rumusan masalah dalam penelitian ini adalah sebagai berikut :

1. Bagaimana pimpinan dapat memperbaiki lulusan agar bisa tepat waktu ?
2. Bagaimana lulusan agar bisa selesai tepat waktu ?

1.3 Batasan Masalah

Dari permasalahan yang telah disampaikan, adapun beberapa batasan-batasan masalah pada penelitian ini adalah sebagai berikut:

1. Penelitian ini menggunakan data mahasiswa berupa nama mahasiswa, NIM mahasiswa, tahun, semester, IPK dan nilai TOEFL.
2. Aplikasi yang dibuat dalam penelitian ini berbasis *web*
3. Untuk mencari nilai *gain* dan *entropy* penyusun melakukan perhitungan manual menggunakan *software* microsoft excel 2010

4. Aplikasi dibuat menggunakan bahasa pemrograman php dan mysql sebagai database nya.
5. Dalam merancang aplikasi, penyusun menggunakan metode pemograman terstruktur dengan *Model Driven Developement* (MDD) dan UML Diagram untuk membuat perancangan sistem.

1.4 Tujuan

Dari rumusan masalah yang telah disampaikan adapun tujuan dari penelitian ini adalah sebagai berikut :

1. Untuk menentukan lulusan tepat waktu maka perlu aplikasi yang dapat memprediksi kelulusan.
2. Dengan algortima C4.5 maka aplikasi dapat menentukan kelulusan.

1.5 Metode Penelitian

Metode yang digunakan untuk menggambarkan aplikasi ini menggunakan *Model Driven Development* yang terdiri dari tahap-tahap dibawah ini :

a. Investigasi Pendahuluan

Penulis memulai penelitian ini dengan mengumpulkan data-data yang dibutuhkan untuk penelitian ini dengan mengunjungi bagian tata usaha Fakultas Teknologi Informasi Universitas Bale Bandung. Data tersebut merupakan data mahasiswa yang meliputi NIM, tahun, semester, Nilai IPK dan Nilai TOEFL.

b. Analisa Masalah

Penulis memulai menganalisa masalah-masalah yang ada pada Fakultas Teknologi Informasi Universitas Bale Bandung yang berhubungan dengan penelitian ini.

c. Analisa Kebutuhan

Penulis menganalisa kebutuhan-kebutuhan yang akan digunaka dalam penelitian ini dimulai dari kebutuhan perangkat lunak maupun kebutuhan perangkat keras.

d. Desain

Pada tahap ini penulis melakukan perancangan desain aplikasi yang akan dibangun dalam penelitian ini.

e. Konstruksi

Pada tahap ini penulis mulai membangun aplikasi yang akan dibuat pada penelitian ini.

f. Implementasi

Pada tahap ini penulis memulai mengimplementasikan aplikasi yang telah dibuat.

g. Pelaporan

Setelah aplikasi telah dibuat dan diimplementasikan, penulis mulai membuat laporan penelitian.

1.6 Sistematika Penulisan

Untuk mempermudah dalam hal penyusunan dan dapat dipahami lebih jelas, laporan ini dibagi atas enam bab yang berisi urutan secara garis besar dan kemudian dibagi lagi dalam sub-sub bab yang akan membahas dan menguraikan masalah secara terperinci dan sub-sub bab yang menjelaskan tahapan-tahapan dalam membangun aplikasi, dengan susunan sebagai berikut :

BAB I PENDAHULUAN

Bab ini membahas tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan, metodologi penelitian dan sistematika penulisan yang akan dibuat pada penelitian ini.

BAB II LANDASAN TEORI

Bab ini membahas tentang teori dan studi literatur yang sesuai dan akurat sehingga bisa mendukung penelitian dalam penulisan sehingga menghasilkan karya tulis yang bernilai ilmiah.

BAB III METODOLOGI PENELITIAN

Bab ini berisi uraian tentang kerangka berpikir, Investigasi pendahuluan, analisa masalah, analisa kebutuhan, desain, konstruksi, implementasi dan pembuatan laporan.

BAB IV ANALISIS DAN PERANCANGAN

Bab ini membahas tentang analisis lebih lanjut terhadap masalah yang ada, perangkat lunak, perangkat keras, dan berbagai macam analisis yang dibutuhkan untuk pembuatan aplikasi. Juga menjelaskan tentang perancangan di mulai dari flowmap diagram, *activity diagram* dan lain sebagainya. .

BAB V IMPLEMENTASI DAN PENGUJIAN

Bab ini membahas tentang implementasi aplikasi yang telah dibuat.

BAB VI KESIMPULAN DAN SARAN

Bab ini membahas tentang kesimpulan yang diperoleh dari pembuatan sistem ini serta saran yang bertujuan untuk pengembangan sistem dimasa yang akan datang.

BAB II

LANDASAN TEORI

2.1 Landasan Teori

Landasan Teori yang berisi referensi dari jurnal yang berkaitan dengan judul dan objek penelitian, berikut beberapa judul jurnal yang digunakan dalam proses penelitian ini :

2.1.1 Jurnal 1

Judul : Prediksi Kelulusan Mahasiswa Tepat Waktu Berdasarkan Usia, Jenis Kelamin, dan Indeks Prestasi Menggunakan Algoritma *Decision Tree* (Agus Romadhona, Suprapedi, M Himawan, 2017)

Dari penelitian yang dilakukan oleh Agus Romadhona Dkk. Untuk menentukan kelulusan mahasiswa tepat waktu dan tidak tepat waktu dengan mengambil data dari indeks prestasi mahasiswa yang diperoleh dari pengambilan mata kuliah yang ditawarkan pada setiap semesternya sesuai dengan kurikulum yang berlaku. Dari penelitian tersebut menjelaskan bahwa kelulusan tepat waktu merupakan indikator keberhasilan mahasiswa dalam proses pendidikan dilihat dari nilai yang diperoleh indeks prestasi serta faktor lain yang dapat menyebabkan kegagalan masa studi mahasiswa sehingga perlu adanya *tool* yang mampu memprediksi. Data yang digunakan dalam penelitian tersebut adalah mahasiswa angkatan 2009 sebagai *training set* yang akan diprediksi kelulusannya ditinjau dari usia, jenis kelamin, indeks prestasi selama 4 semester pertama yaitu semester 1, semester 2, semester 3 dan semester 4. Selanjutnya dilakukan analisis komparasi algoritma klasifikasi yaitu *Algoritma Decision Tree C4.5*, *ID3* dan *Chaid* sehingga dapat diketahui algoritma yang paling akurat untuk memprediksi kelulusan mahasiswa lulus tepat waktu.

2.1.2 Jurnal 2

Judul : Penerapan Algoritma C4.5 Dalam Memprediksi Kelulusan Tepat Waktu Pada Perguruan Tinggi (Studi Kasus : STMIK Royal Kisaran)

(Nanda Dimas Prayoga, 2017)

Dari penelitian yang dilakukan oleh Nanda Dimas Prayoga data yang diambilkan dan dikumpulkan bersumber dari data kelulusan mahasiswa yang diberikan oleh perguruan tinggi STMIK Royal Kisaran. Penelitian tersebut mengumpulkan data sample yang memiliki atribut nilai IPK, Prestasi, Etika, Sks dan Kelayakan. Setelah mengumpulkan data sample kemudian melakukan perhitungan jumlah data, *entropy* dan *gain*. Kemudian, dalam melakukan pengujian terhadap data sample menggunakan *tools* RapidMiner 5 dimulai dari proses koneksi antar basis data sample. Setelah dilakukan perhitungan dan pengujian keluarlah hasil pengukuran akurasi terhadap kinerja algoritma C4.5 yaitu sebesar 92,60%+-1.60%.

2.1.3 Jurnal 3

Judul : Prediksi Kelulusan Mahasiswa Menggunakan Algoritma *Decision Tree* C4.5 Dengan Teknik *Pruning* (Isa Iskandar, Leli Hiryanto, Janson Hendryli)

Dari penelitian yang dilakukan oleh Isa Iskandar Dkk. Sistem prediksi yang dibuat menggunakan metode *decision tree* dengan algoritma C4.5 , dimana masa studi mahasiswa akan dibagi menjadi 8 kelas, yaitu 7 semester, 8 semester, 9 semester, 10 semester, 11 semester, 12 semester, 13 semester, dan 14 semester. Pengujian terhadap data dilakukan terhadap hasil prediksi lama masa studi mahasiswa. Data yang digunakan untuk pengujian yaitu menggunakan data mahasiswa angkatan 2008 sampai angkatan 2012 yang telah menyelesaikan studinya. Data tersebut dibagi dengan ketentuan pembagian adalah 80% sebagai data *training* dan 20% sebagai data *testing*. Pengujian dilakukan menjadi 2 tahap yaitu pengujian terhadap model *tree* yang terbentuk dan pengujian terhadap hasil-

prediksi. Hasil dari pengujian tersebut terbagi menjadi 6 percobaan dimana percobaan 1 sampai dengan percobaan enam memiliki hasil yang berbeda setelah dilakukan *prune*.

2.1.4 Jurnal 4

Judul : Prediksi Ketepatan Waktu Kelulusan Dengan Algoritma Data Mining C4.5 (Indah Puji Astuti, 2017)

Dari penelitian yang dilakukan oleh Indah Puji Astuti data yang digunakan adalah data mahasiswa angkatan 2012/2013 dengan penggunaan atribut berupa NIM, jenis sekolah asal, asal daerah, pekerjaan orang tua, dan kelas. Data mahasiswa tersebut dibagi menjadi 2 atribut yaitu atribut *input* dan atribut *output*. Dimana atribut *input* berisi jenis sekolah asal, asal daerah, pekerjaan orang tua, dan kelas. Sedangkan atribut *output* berisi nilai berupa kelulusan “tepat waktu” dan “tidak tepat waktu”. Metode yang digunakan oleh penelitian tersebut yaitu metode CRISP-DM (*Cross Industry Standar Process For Data Mining*). Hasil dari penelitian ini adalah berupa analisa yang mengarah ke pengambilan keputusan. Hasil analisa diharapkan dapat dimanfaatkan oleh institusi perguruan tinggi sebagai bahan pertimbangan dalam menentukan langkah guna mengatasi permasalahan mahasiswa yang lulus tidak tepat waktu dimasa yang akan datang.

2.2 Dasar Teori

2.2.1 Prediksi

Prediksi adalah suatu proses memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa depan berdasarkan informasi masa lalu dan sekarang yang dimiliki, agar kesalahannya (selisih antara sesuatu yang terjadi dengan hasil perkiraan) dapat diperkecil. Prediksi tidak harus memberikan jawaban secara pasti kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin yang akan terjadi. (Herdianto, 2013)

Pengertian Prediksi sama dengan ramalan atau perkiraan. Menurut kamus besar bahasa Indonesia, prediksi adalah hasil dari kegiatan memprediksi atau meramal atau memperkirakan nilai pada masa yang akan datang dengan menggunakan data masa lalu. Prediksi menunjukkan apa yang akan terjadi pada suatu keadaan tertentu dan merupakan input bagi proses perencanaan dan pengambilan keputusan. Prediksi bisa berdasarkan metode ilmiah ataupun subjektif belaka. Ambil contoh, prediksi cuaca selalu berdasarkan data dan informasi terbaru yang didasarkan pengamatan termasuk oleh satelit. Begitupun prediksi gempa, gunung meletus ataupun bencana secara umum. Namun, prediksi seperti pertandingan sepakbola, olahraga, dll umumnya berdasarkan pandangan subjektif dengan sudut pandang sendiri yang memprediksinya. Permulaan awal, walaupun pengkajian yang mendalam mengenai alternatif masa depan adalah suatu disiplin baru, barangkali orang telah menaruh perhatian besar tentang apa yang akan terjadi kemudian semenjak manusia mulai mengetahui sesuatu. Populasi tukang ramal dan tukang nujum pada zaman kuno dan abad pertengahan merupakan satu manifestasi dari keinginan tahu orang tentang masa depannya. Perhatian tentang masa depan ini berlangsung terus bahkan berkembang menjadi kolom astrologi yang disindikatkan pada tahun Secara Eksplisit, pembahasan mengenai teori peramalan kebijakan sangatlah sedikit.

2.2.1.1 Teknik Prediksi

Berdasarkan teknik yang digunakan untuk memprediksi maka prediksi dapat dibagi menjadi dua bagian yaitu prediksi kualitatif dan prediksi kuantitatif (Herdianto, 2013)

1. Prediksi Kualitatif

Prediksi Kualitatif didasarkan atas data kualitatif pada masa lalu. Metode kualitatif digunakan jika data masa lalu dari variabel yang akan diprediksi tidak ada, tidak cukup atau kurang dipercaya. Hasil prediksi yang dibuat sangat tergantung pada individu yang menyusunnya. Hal ini penting karena hasil prediksi tersebut ditentukan-

berdasarkan pemikiran yang bersifat *judgement* atau opini, pengetahuan dan pengalaman dari penyusunnya. Oleh karena itu metode kualitatif ini disebut juga *judgemental, subjective, intuitive*.

2. Prediksi Kuantitatif

Prediksi kuantitatif didasarkan atas data kuantitatif pada masa lalu. Hasil prediksi yang dibuat sangat tergantung pada metode yang dipergunakan dalam prediksi tersebut. Dengan metode yang berbeda akan diperoleh hasil prediksi yang berbeda. Hal yang perlu diperhatikan dari penggunaan metode tersebut adalah baik tidaknya metode yang digunakan dan sangat ditentukan dari penyimpangan antara hasil prediksi dengan kenyataan yang terjadi. Metode yang baik adalah metode yang memberikan nilai-nilai perbedaan atau penyimpangan yang mungkin. Prediksi kuantitatif hanya dapat digunakan apabila terdapat tiga kondisi sebagai berikut :

- a. Adanya informasi tentang keadaan yang lain
- b. Informasi tersebut dapat dikuantifikasikan dalam bentuk data
- c. Dapat diasumsikan bahwa pola yang lalu akan berkelanjutan pada masa yang akan datang

2.2.2 Kelulusan

Proses kelulusan adalah kegiatan paling akhir dari manajemen peserta didik. Kelulusan adalah pernyataan dari lembaga pendidikan (sekolah) tentang telah diselesaikannya program pendidikan yang harus diikuti oleh peserta didik. Setelah peserta didik selesai mengikuti seluruh program pendidikan disuatu lembaga pendidikan dan berhasil lulus dan ujian akhir, maka kepada peserta didik tersebut diberikan surat keterangan lulus atau sertifikat. Umumnya surat keterangan tersebut sering disebut ijazah atau surat tanda tamat belajar (STTB). Kompetensi lulusan-yaitu kemampuan minimal yang harus dicapai oleh peserta didik setelah tamat mengikuti pendidikan pada jenjang atau satuan pendidikan tertentu.

Ketika peserta didik sudah lulus, maka secara formal hubungan antara peserta didik dan lembaga telah selesai. Namun demikian, diharapkan hubungan antara para alumni dan sekolah tetap terjalin. Dari hubungan sekolah dan alumni ini, lembaga pendidikan (sekolah) bisa memanfaatkan hasil-hasilnya. Lembaga pendidikan (sekolah) bisa menjaring berbagai informasi. Misalnya informasi tentang materi pelajaran mana yang sangat membantu untuk studi selanjutnya. Mungkin juga informasi tentang lapangan kerja yang bisa dijangkau bagi alumni lainnya.

Hubungan antara sekolah dengan para alumni dapat dipelihara lewat pertemuan-pertemuan yang diselenggarakan oleh para alumni, yang biasa disebut “reuni”. Bahkan saat ini setiap lembaga pendidikan (sekolah) ada organisasi alumni, misalnya IKA (Ikatan Alumni). Prestasi yang dicapai oleh alumni dari lembaga pendidikan (sekolah) ini perlu didata atau dicatat oleh lembaga. Sebab catatan tersebut sangat berguna bagi lembaga dalam mempromosikan lembaga pendidikannya. (Tim Dosen UPI, 2012)

2.2.2.1 Menentukan Kelulusan

untuk mengetahui cara kelulusan peserta didik kita perlu mengetahui arti dari standar penilaian pendidik yakni : (Sanjaya Wina, 2007)

1. Standar umum penilaian, merupakan aturan main dari aspek-aspek umum dalam pelaksanaan penilaian. Untuk melakukan penilaian pendidik harus selalu mengacu pada standar umum penilaian. Adapun prinsip-prinsip dari standar umum penilaian adalah sebagai berikut:
 - a. Pemilihan teknik penilaian disesuaikan dengan karakteristik mata pelajaran serta jenis informasi yang ingin diperoleh dari peserta didik.
 - b. Informasi yang dihimpun mencakup ranah-ranah yang sesuai dengan standar isi dan standar kompetensi kelulusan.
 - c. Informasi mengenai perkembangan perilaku peserta didik dilakukan secara berkala pada kelompok mata pelajaran masing-masing.

- d. Pendidik harus selalu mencatat perilaku peserta didik yang menonjol, baik yang bersifat positif maupun yang negatif dalam buku catatab perilaku.
 - e. Melakukan sekurang-kurangnya tiga kali ulangan harian menjelang ulangan tengah semester, dan tiga kali menjelang ulangan akhir semester.
 - f. Pendidikan harus menggunakan teknik penilaian yang divariasi sesuai dengan kebutuhan.
 - g. Pendidik harus selalu memeriksa dan memberi balikan kepada peserta didik atas hasil kerjanya sebelum memberikan tugas lanjutan.
 - h. Pendidik harus memiliki catatan kumulatif tentang hasil penilaian untuk setiap peserta didik yang berada dibawah tanggung jawabnya. Pendidik harus pula mencatat semua kinerja peserta didik untuk menentukan pencapaian kompetensi peserta didik.
 - i. Pendidik melakukan ulangan tengah dan akhir semester untuk menilai penguasaan kompetensi sesuai dengan tuntutan dalam standar kompetensi (SK) dan standar lulusan (SL).
 - j. Pendidik yang diberi tugas menangani pengembangan diri harus melaporkan kegiatan pengembangan pada buku laporan pendidikan.
 - k. Pendidik menjaga kerahasiaan pribadi peserta didik dan tidak disampaikan kepada pihak lain tanpa seizing yang bersangkutan maupun orang tua/wali murid.
2. Standar Perencanaan Penilaian, merupakan prinsip-prinsip yang harus dipedomani bagi pendidik dalam melakukan perencanaan penilaian. Adapun prinsip-prinsipnya sebagai berikut:
- a. Pendidik harus membuat rencana penilaian secara terpadu dengan silabus dan rencana pembelajarannya. Perencanaan penilaian setidak-tidaknya meliputi komponen yang akan dinilai, teknik yang akan digunakan serta kriteria pecapaian kompetensi.
 - b. Pendidikan harus mengembangkan kriteria pencapaian kompetensi dasar (KD) sebagai dasar untuk penilaian.

- c. Pendidik menentukan teknik penilaian dan instrument penilaiannya sesuai dengan indikator pencapaian KD.
 - d. Pendidik harus menginformasikan seawal mungkin kepada peserta didik tentang aspek-aspek yang dinilai dan criteria pencapaiannya.
 - e. Pendidik menuangkan seluruh komponen penilaian ke dalam kisi-kisi penilaian.
 - f. Pendidik membuat instrumen berdasarkan kisi-kisi yang telah dibuat dan dilengkapi dengan pedoman penskoran sesuai dengan teknik penilaian yang digunakan.
 - g. Pendidik menggunakan acuan kriteria dalam menentukan nilai peserta didik.
3. Standar Pelaksanaan Penilaian Dalam pedoman umum penilaian yang disusun oleh BSNP, standar pelaksanaan penilaian oleh pendidik meliputi:
- a. Pendidik melakukan kegiatan penilaian sesuai dengan rencana penilaian yang telah disusun di awal kegiatan pembelajaran.
 - b. Pendidik menganalisis kualitas instrumen dengan mengacu pada persyaratan instrument serta menggunakan acuan criteria.
 - c. Pendidik menjamin pelaksanaan ulangan dan ujian yang bebas dari kemungkinan terjadinya tindak kecurangan.
 - d. Pendidik memeriksa pekerjaan peserta didik dan memberikan umpan balik dan komentar yang bersifat mendidik.
4. Standar Penilaian oleh satuan Pendidikan Menurut BSNP ada dua standar pokok yang harus diperhatikan dalam penilaian hasil belajar dalam satuan pendidikan, yaitu;
- a. Standar penentuan kenaikan kelas. Standar ini terdiri atas tiga hal pokok, yaitu: (a) pada akhir tahun pelajaran, satuan pendidikan menyelenggarakan ulangan kenaikan kelas. (b) satuan pendidikan menetapkan SKBM (standar ketuntasan belajar minimal) pada setiap mata pelajaran. SKBM tersebut harus ditingkatkan secara berencana dan berkala. (c) satuan pendidikan menyelenggarakan rapat dewan pendidikan untuk menentukan kenaikan kelas setiap peserta didik.

- b. Standar penentuan kelulusan: (a) pada akhir jenjang pendidikan, satuan pendidikan menyelenggarakan ujian sekolah pada kelompok mata pelajaran IPTEKS. (b) satuan pendidikan menyelenggarakan rapat dewan pendidikan untuk menentukan nilai akhir peserta didik pada:
- a. Kelompok mata pelajaran agama dan mulia.
 - b. Kelompok mata pelajaran kewarganegaraan dan kepribadian.
 - c. Kelompok mata pelajaran estetika.
 - d. Kelompok mata pelajaran jasmani, olahraga dan kesehatan untuk menentukan kelulusan.
- c. satuan pendidikan menentukan kelulusan peserta didik berdasarkan kriteria kelulusan yang telah ditetapkan dalam peraturan pemerintah No. 19/2005 pasal 72 ayat (1) yang menyatakan bahwa peserta didik dinyatakan lulus dari satuan pendidikan pada pendidikan dasar dan menengah setelah:
1. Menyelesaikan seluruh program pembelajaran.
 2. Memperoleh nilai minimal baik pada penilaian akhir untuk seluruh mata pelajaran kelompok mata pelajaran agama dan akhlak mulia, kelompok mata pelajaran kewarganegaraan dan kepribadian, kelompok mata pelajaran estetika dan kelompok mata pelajaran jasmani, olahraga dan kesehatan. (Zainal Arifin, 2012)

Standart kompetensi kelulusan dalam kurikulum 2013 adalah sikap, keterampilan, dan pengetahuan berikut skemanya:

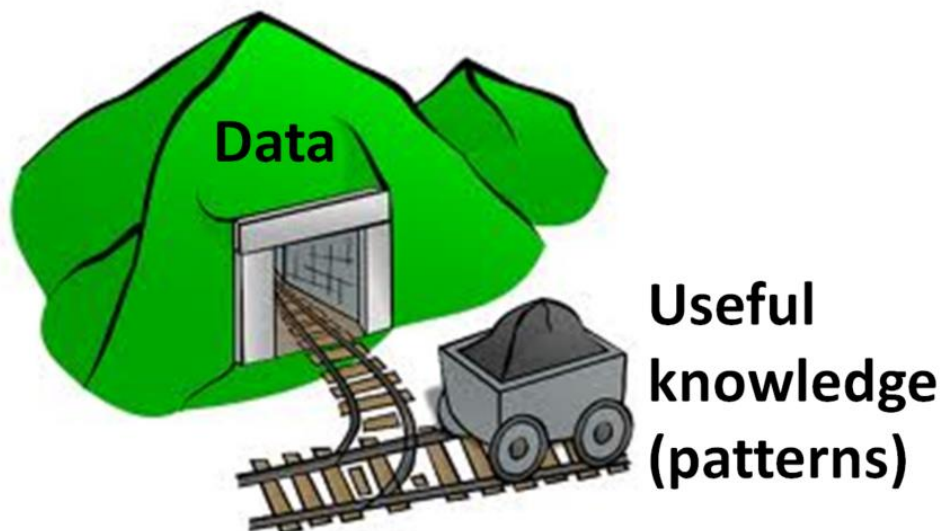
1. Penilaian berbasis kompetensi
2. Pergeseran dari penilaian yang mengukur pengetahuan dan berdasarkan hasil saja, menuju penilaian otentik (mengukur sikap, ketrampilan, dan pengetahuan berdasarkan proses dan hasil)
3. Memperkuat PAP (penilaian acuan patokan) yaitu pencapaian hasil belajar didasarkan pada posisi skor yang diperolehnya terhadap skor ideal (maksimal)

2.2.3 Mahasiswa Lulus

Mahasiswa yang lulus merupakan mahasiswa yang telah dinyatakan lulus program sarjana apabila memenuhi kriteria yang telah ditentukan, misalnya mahasiswa tersebut telah menyelesaikan sekurang-kurangnya sejumlah 144-148 sks yang diwajibkan, IPK minimal 2,00, Nilai mata kuliah wajib minimal C, Jumlah SKS dengan nilai D tidak lebih dari 10% jumlah SKS total, tidak ada nilai E dan lulus ujian skripsi

2.2.4 Data Mining

Nama data mining sebenarnya mulai dikenal sejak tahun 1990 (Gambar 2.1), ketika pekerjaan pemanfaatan data menjadi sesuatu yang penting dalam berbagai bidang, mulai dari bidang akademik, bisnis hingga medis (Goronescu, 2011). Data mining dapat diterapkan pada berbagai bidang yang mempunyai sejumlah data, tetapi karena wilayah penelitian dengan sejarah yang belum lama dan belum melewati masa ‘remaja’, maka data mining masih diperdebatkan posisi bidang pengetahuan yang memilikinya. Maka, Daryl Pregibon menyatakan bahwa “data mining adalah campuran dari statistik kecerdasan buatan, dan riset basis data” yang masih berkembang (Goronescu, 2011).



Gambar 2.1 Data Mining

(Sumber : Philipe Fournier Viger)

Terlepas dari ‘remaja-nya’ data mining, ternyata data mining diproyeksikan menjadi jutaan dollar di dunia industri pada tahun 2000, sedangkan pada saat yang sama, ternyata data mining dipandang sebelah mata oleh sejumlah peneliti sebagai *dirty word in statics* (Goronescu, 2011). Mereka adalah orang-orang yang tidak memandang data mining sebagai sesuatu yang menarik bagi mereka pada saat itu.

Munculnya data mining didasarkan pada jumlah data yang tersimpan dalam basis data semakin besar. Misalnya dalam sebuah supermarket, ada beberapa-transaksi pelanggan yang terjadi dalam sehari dan ada berapa juta data yang sudah tersimpan dalam sebulan. Dalam perusahaan, ada berapa juta data yang sudah tersimpan dari setiap kegiatan produksi untuk setiap produk yang dibuat dalam beberapa tahun. Contoh lain, jika anda mempunyai kartu kredit, mungkin anda sering menerima surat penawaran barang atau jasa. Jika bank mempunyai 1.000.000 nasabah dan biaya pengiriman surat per nasabah adalah 500 rupiah, maka biaya yang harus dikeluarkan bank adalah 500 juta rupiah, padahal nasabah yang mungkin benar-benar hanya membeli sekitar 15%. Akibatnya ada pembuangan biaya sekitar 85% dari 500 juta atau sekitar 425 juta; sungguh sia-sia. Jika perusahaan dapat memanfaatkan data-data yang ada sehingga hanya nasabah yang berpotensi untuk membeli saja yang dikirim surat maka biaya pengiriman tersebut dapat ditekan.

Yang menjadi pertanyaan untuk data-data dalam perusahaan yang semakin menggenung jumlahnya dari waktu ke waktu adalah mau diapakan data tersebut, apakah hanya dibuat untuk laporan akhir tahun kemudian dibuang? Apakah hanya akan dikubur dalam gudang data dan tidak diapa-apakan? Tentu sayang sekali jika data-data tersebut tidak dimanfaatkan untuk kepentingan perusahaan atau instansi-instansi yang berkepentingan.

Ada istilah lain yang mempunyai makna yang sama dengan data mining yaitu *knowledge-discovery database* (KDD). Memang data mining atau KDD bertujuan untuk memanfaatkan data dalam basis data dengan mengolahnya sehingga menghasilkan informasi baru yang berguna. *Knowledge discovery database* adalah keseluruhan proses *non-trivial* Seperti diilustrasikan pada Gambar

2.1, jika dilacak akar keilmuannya, ternyata data mining mempunyai empat akar bidang ilmu sebagai berikut:

1. Statistik

Bidang ini merupakan akar paling tua, tanpa ada statistik maka data mining mungkin tidak ada. Dengan menggunakan statistik klasik ternyata data yang diolah dapat diringkas dalam apa yang umum dikenal sebagai *exploratory data analysis* (EDA). EDA berguna untuk mengidentifikasi hubungan sistematis antar variable/fitur ketika tidak ada cukup informasi alami yang dibawanya. Teknik EDA klasik yang digunakan dalam data mining, diantaranya :

- a. Metode komputasional: statistik deskriptif (distribusi, parameter statistik klasik (mean, median, rata-rata, varian dan sebagainya), korelasi, tabel frekuensi, teknik eksplorasi multivariat (analisis cluster, analisis faktor, analisis komponen utama dan klasifikasi, analisis kanonik, analisis diskriminan, *classification tree*, analisis korespondensi), model linear/non linear lanjutan (regresi linear/non linear, *time series/forecasting*, dan sebagainya).
- b. Visualisasi Data: mengarah pada representasi informasi dalam bentuk visual dan dapat dipandang sebagai satu yang paling berguna. Pada saat yang sama, visualisasi data merupakan metode eksplorasi data yang atraktif. Teknik visualisasi yang paling umum yang dikenal adalah histogram semua jenis (kolom, silinder, kerucut, piramida, lingkaran, batang, dan sebagainya), kotak, scatter, kontur, matriks, ikon, dan sebagainya.

1. Kecerdasan buatan atau *artificial intelligence* (AI)

Bidang ilmu ini berbeda dengan statistik. Teorinya dibangun berdasarkan teknik heuristik sehingga AI berkontribusi terhadap teknik pengolahan informasi berdasarkan pada model penalaran manusia. Salah satu cabang dari AI, yaitu pembelajaran mesin atau *machine learning*, merupakan disiplin ilmu yang paling penting yang direpresentasikan dalam pembangunan

data mining, menggunakan teknik dimana sistem komputer berjalan dengan ‘pelatihan’.

2. Pengenalan Pola

Sebenarnya data mining juga menjadi turunan bidang pengenalan pola, tetapi hanya mengolah data dari basis data. Data yang diambil dari basis data untuk diolah bukan dalam bentuk relasi, melainkan dalam bentuk normal pertama sehingga set data dibentuk menjadi bentuk normal pertama. Akan tetapi, data mining mempunyai ciri khas yaitu pencarian pola asosiasi dan pola sekuensial.

3. Sistem Basis Data

Akar bidang ilmu keempat dari data mining yang menyediakan informasi berupa data yang akan ‘digali’ menggunakan metode-metode yang disebutkan sebelumnya.

Kebutuhan ‘penggalan’ informasi dari dalam data dapat dilihat pada kasus dunia nyata, diantaranya sebagai berikut:

- a. Ekonomi: Ada jumlah data yang sangat besar yang dikumpulkan dari berbagai bidang, seperti data web, *e-commerce*, supermarket, transaksi keuangan dan perbankan, dan sebagainya yang siap dianalisis dengan tujuan untuk mendapatkan keputusan yang optimal terkait tujuan lembaga.
- b. Pelayanan kesehatan: Saat ini ada banyak basis data berbeda dalam bidang pelayanan kesehatan (medis dan farmasi), yang dianalisis secara parsial, khususnya dengan cara medis sendiri, padahal sebelumnya dalam data tersembunyi banyak informasi yang belum dibuka secara tepat.
- c. Riset pengetahuan: Ada basis data besar yang dikumpulkan bertahun-tahun dalam bermacam-macam bidang (astronomi, meteorologi, biologi, linguistik, dan sebagainya) yang tidak dapat dieksplorasi menggunakan cara tradisional.

Dari penjelasan diatas bahwa disatu sisi ada sejumlah data dalam jumlah besar yang secara sistematis belum dieksplorasi, dan di sisi lain, kekuatan teknik

komputasi dan ilmu komputer sudah tumbuh secara eksponensial sehingga menyebabkan tekanan pada kebutuhan untuk membuka informasi yang ‘tersembunyi’ dari data menjadi meningkat. Bidang data mining menjadi jawaban untuk menyelesaikan persoalan diatas yang pada awalnya tidak mungkin untuk dideteksi dengan cara tradisional dan hanya menggunakan kemampuan analisis manusia.

Pengertian data mining cukup sulit dijelaskan dengan gambar 2.1 jika mengingat data mining juga merupakan gabungan dari beberapa bidang ilmu. Berikut beberapa pengertian data mining yang secara naratif mempunyai beberapa maksud yang mirip:

- a. Pencarian otomatis pola dalam basis data besar, menggunakan teknik komputasional campuran dari statistik, pembelajaran mesin dan pengenalan pola.
- b. Pengekstrakan implisit non-trivial, yang sebelumnya belum diketahui secara potensial adalah informasi berguna dari data.
- c. Ilmu pengekstrakan informasi yang berguna dari set data atau basis data besar.
- d. Eksplorasi otomatis atau semiotomatis dan analisis data dalam jumlah besar, dengan tujuan untuk menemukan pola yang bermakna.
- e. Proses penemuan informasi otomatis dengan mengidentifikasi pola dan hubungan ‘tersembunyi’ dalam data.

Meskipun data mining juga diartikan penemuan informasi, tidak semua penemuan informasi/data disebut data mining. Berikut ini beberapa contoh apa yang disebut dengan data mining dan tidak disebut dengan data mining:

- a. Bukan data mining: Pencarian informasi tertentu di internet (misalnya mencari info masakan menggunakan Google).
Data mining: Pengelompokan informasi yang mirip dalam konteks tertentu pada hasil pencarian (misalnya masakan Indonesia, masakan belanda, dan sebagainya yang ditemukan Google).
- b. Bukan data mining: Petugas medis mencari data medis untuk menganalisis catatan pasien dengan penyakit tertentu.

Data mining: Peneliti medis mencari cara pengelompokan data penyakit pasien berdasarkan data diagnosis, umur, alamat.

- c. Bukan data mining: Analisis gambar laporan keuangan penjualan perusahaan.

Data mining: Menggunakan basis data transaksi perusahaan dengan fokus ke data sales untuk mengidentifikasi profil utama pelanggan.

- d. Bukan data mining: Pembuatan laporan tahunan penjualan perusahaan dengan merekap semua data selama setahun.
- e. Data mining: Pemanfaatan data penjualan perusahaan untuk mendapatkan pola prediksi stok yang sebaiknya disediakan pada tahun berikutnya.

Contoh lebih jelas mengenai perbedaan antara pencarian dalam basis data dengan data mining adalah seseorang dapat tertarik pada perbedaan antara jumlah pembelian dari jenis tertentu dari supermarket dibandingkan dengan hipermarket, atau dari dua supermarket dari dua wilayah berbeda. Dalam kasus ini, biasanya sudah ada asumsi informasi awal bahwa memang ada perbedaan besar antara supermarket dan hipermarket dalam hal tempat, atau sales supermarket antara wilayah satu dengan lainnya. Sebaliknya, dalam data mining, masalah bisa berisi objek dalam pengidentifikasian faktor yang mempengaruhi volume sales tanpa mengandalkan dugaan awal apapun. Jadi kesimpulannya, metode dalam data mining melakukan identifikasi pola dan hubungan tersembunyi yang tidak selalu jelas (dan dengan mudah dapat mengidentifikasinya) dibawah keadaan asumsi tertentu.

Jadi, kita tidak bisa menyamakan pencarian tertentu dari objek individu (jenis apa pun) dengan pencarian data mining. Pencarian dalam data mining tidak mencari secara individualis, tetapi set individualis, atau dengan kata lain dikelompokkan dengan kriteria tertentu. Perbedaan antara pencarian biasa dengan data mining dianalogikan dengan pencarian pohon dengan pencarian hutan; “tidak bisa menilai hutan hanya dengan menilai pohon”.

2.2.4.1 Pekerjaan Dalam Data Mining

Pekerjaan yang berkaitan dengan data mining dapat dibagi menjadi empat kelompok: model prediksi (*prediction modelling*), analisis cluster (*cluster-analysis*), analisis asosiasi (*association analysis*), dan deteksi anomaly (*anomaly detection*).

Model Prediksi (*Prediction Modelling*)

Pekerjaan ini berkaitan dengan pembuatan sebuah model yang dapat melakukan pemetaan dari setiap himpunan variable ke setiap targetnya, kemudian menggunakan model tersebut untuk memberikan nilai target pada himpunan baru yang didapat. Ada 2 jenis model prediksi, yaitu klasifikasi dan regresi. Klasifikasi digunakan untuk variabel target diskret, sedangkan regresi digunakan untuk variabel target kontinu.

Contoh pekerjaan yang menggunakan jenis klasifikasi adalah melakukan deteksi jenis penyakit pasien berdasarkan sejumlah nilai-nilai parameter penyakit yang diderita masuk. Pekerjaan ini termasuk jenis klasifikasi karena target yang diharapkan adalah diskret, hanya beberapa jenis kemungkinan nilai target yang didapatkan dan tidak ada nilai seri waktu (*time series*) yang harus didapatkan untuk mendapat target nilai akhir. Sementara melakukan prediksi jumlah penjualan yang didapatkan pada 3 bulan kedepan itu termasuk regresi karena untuk mendapatkan nilai penjualan bulan ketiga harus mendapatkan nilai penjualan bulan ke dua dan untuk mendapatkan nilai penjualan ke dua harus mendapatkan nilai penjualan bulan pertama. Dalam hal ini ada nilai seri waktu yang harus dihitung untuk sampai pada target akhir yang diinginkan dan ada nilai kontinu yang harus dihitung untuk mendapatkan nilai target akhir yang diinginkan.

1. Analisis Cluster (*Cluster Analysis*)

Contoh pekerjaan yang berkaitan dengan analisis cluster adalah bagaimana bisa mengetahui pola pembelian barang oleh konsumen pada waktu-waktu tertentu. Dengan mengetahui pola kelompok pembelian tersebut, maka perusahaan/retailer dapat menentukan jadwal promosi yang dapat diberikan sehingga dapat membantu meningkatkan omzet penjualan.

Analisis kelompok melakukan pengelompokan data ke dalam sejumlah kelompok berdasarkan kesamaan karakteristik masing-masing data pada kelompok-kelompok yang ada. Data-data yang masuk dalam batas kesamaan dengan kelompoknya yang akan bergabung dalam kelompok tersebut, dan akan terpisah dalam kelompok yang berbeda jika keluar dari batas kesamaan kelompok tersebut.

2. Analisis Asosiasi (*Association Analysis*)

Analisis asosiasi digunakan untuk menemukan pola yang menggambarkan kekuatan hubungan fitur dalam data. Pola yang ditemukan biasanya mempresentasikan bentuk aturan implikasi atau subset fitur. Tujuannya adalah untuk menemukan pola yang menarik dengan cara yang efisien.

Penerapan yang paling dekat dalam kehidupan sehari-hari adalah analisis data keranjang belanja. Jika ibu rumah tangga akan membeli barang kebutuhan rumah tangga (misalnya beras) di sebuah supermarket, maka sangat besar kemungkinan ibu rumah tangga tersebut juga akan membeli kebutuhan rumah tangga yang lain, misalnya minyak atau telur, dan tidak mungkin (atau jarang) membeli barang rumah tangga lain seperti topi atau buku. Dengan mengetahui hubungan yang lebih kuat antara beras dengan telur daripada beras dengan topi, maka retailer dapat menentukan barang-barang yang sebaiknya disediakan dalam jumlah yang cukup banyak.

3. Deteksi Anomali (*Anomaly Detection*)

Pekerjaan deteksi anomali berkaitan dengan pengamatan sebuah data dari sejumlah data yang secara signifikan mempunyai karakteristik yang berbeda dari sisi data yang lain. Data-data yang karakteristiknya menyimpang (berbeda) dari data yang lain disebut sebagai *outliner*. Algoritma deteksi anomaly yang baik harus mempunyai laju deteksi yang tinggi dan laju kesalahan yang rendah. Deteksi anomali dapat diterapkan pada sistem jaringan untuk mengetahui pola data yang memasuki jaringan sehingga dapat diketahui adanya penyusupan jika pola kerja

data yang datang berbeda ataupun perilaku kondisi cuaca yang mengalami anomali juga dapat dideteksi dengan algoritma ini.

2.2.4.2 Proses Data Mining

Secara sistematis, ada tiga langkah utama dalam data mining (Goronescu, 2011):

1. Explorasi/pemrosesan awal data

Explorasi/pemrosesan awal data terdiri dari ‘pembersihan’ data, normalisasi data, transformasi data, penanganan data yang salah, reduksi dimensi, pemilihan subset fitur, dan sebagainya.

2. Membangun model dan melakukan validasi terhadapnya

Membangun model dan melakukan validasi terhadapnya berarti melakukan analisis berbagai model dan memilih model dengan kinerja prediksi yang terbaik. Dalam langkah ini digunakan metode-metode seperti klasifikasi, regresi, analisis cluster, deteksi anomali, analisis asosiasi, analisis pola sekuensial, dan sebagainya. Dalam beberapa referensi, deteksi anomaly juga masuk dalam langkah eksplorasi. Akan tetapi, deteksi anomali juga dapat digunakan sebagai algoritma utama, terutama untuk mencari data-data yang spesial.

3. Penerapan

Penerapan berarti menerapkan model pada data yang baru untuk menghasilkan perkiraan/prediksi masalah yang diinvestigasi.

2.2.4.1 Set Data

Bukan data mining namanya jika tidak ada set data yang diolah di dalamnya. Kata ‘data’ dalam terminologi statistik adalah kumpulan objek dengan atribut-atribut tertentu, dimana objek tersebut adalah individu berupa data dimana setiap data memilih jumlah atribut. Atribut tersebut berpengaruh pada dimensi dari data, semakin banyak atribut/fitur maka semakin besar dimensi data. Kumpulan data-data membentuk set data.

Dalam set data berbentuk record data, tidak ada hubungan antara baris data dengan baris data yang lain dan juga tidak punya hubungan dengan set data yang lain. Setiap baris data berdiri sendiri sebagai sebuah data individu. Dalam sistem basis data, umumnya ada sejumlah tabel yang saling berhubungan menggunakan suatu kunci, tetapi dalam set berbentuk record data, diasumsikan bahwa hanya ada satu tabel yang berisi sejumlah baris data. Maka dari itu, biasanya set data yang diolah dalam data mining adalah keluaran dari sistem *data warehouse* yang-menggunakan *query* untuk melakukan pengambilan data dari sejumlah tabel dalam sistem basis data. Ada tiga jenis record data, yaitu matriks, transaksi, dan dokumen.

Set data yang berisi kumpulan data, dengan semua data mempunyai sejumlah atribut (fitur) numerik yang sama, dapat dipandang sebagai vector (data) dalam wilayah multidimensi, dimana setiap dimensi (fitur) mempresentasikan atribut berbeda yang menggambarkan objek/data. Set data dengan model seperti itu disebut juga dengan data matrixs karena semua nilainya berisi nilai numerik. Sekumpulan objek/data matriks dapat diintegrasikan sebagai matriks $M \times N$. M adalah jumlah baris, satu baris menyatakan satu data, dan N adalah jumlah kolom, satu kolom menyatakan satu atribut/fitur. Data matriks merupakan jenis record data yang paling umum dan banyak digunakan pada aplikasi-aplikasi statistik. Contoh data matriks ditunjukkan pada tabel 2.1.

Tabel 2.1 Data Matriks

Tinggi	Berat	Sepatu	Celana
168	60	38	30
175	85	42	35
170	77	39	39
165	55	38	28
171	66	41	27
169	70	40	29

Data keranjang belanja adalah contoh data transaksi. Setiap record (transaksi) berisi sejumlah item untuk sebuah transaksi bisa berbeda dengan transaksi yang lain. Contohnya bisa dilihat pada kasus keranjang belanja di pasar

atau supermarket. Setiap pembeli melakukan pembelian barang yang jumlahnya dan jenisnya bisa berbeda dengan pembeli yang lain. Contoh data transaksi dapat dilihat pada Tabel 2.2.

Tabel 2.2 Data Transaksi

IDT	Item
1	Susu, Bedak, Sabun
2	Roti, Susu, Mentega
3	Gula, Roti, Terigu
4	Bedak, Gula, Sabun, Roti
5	Roti, Terigu
6	Gula, Susu, Sabun, Mentega

2.2.5 Algoritma C4.5

Algoritma C4.5 merupakan salah satu algoritma yang digunakan untuk membentuk *decision tree* berdasarkan *training data*. Algoritma ini merupakan metode klasifikasi dan prediksi yang sangat kuat dan terkenal (Wenefrida Tulit Ina, 2013). *Decision tree* adalah model prediksi menggunakan struktur *tree* atau struktur berhirarki. Konsep dari *decision tree* adalah mengubah data menjadi *decision tree* dan aturan-aturan keputusan (*decision rules*) (eko prasetyo, 2014).

Beberapa pengembangan yang dilakukan algoritma C4.5 adalah bisa mengatasi *missing value*, *continue data*, dan *pruning*. Algoritma C4.5 mempunyai *input* berupa *training samples* dan *samples*, *training samples* berupa data contoh yang akan digunakan untuk membangun sebuah *tree* yang telah diuji kebenarannya, sedangkan untuk *samples* merupakan *field-field* data yang nantinya akan kita gunakan sebagai parameter dalam melakukan klasifikasi data (eko prasetyo, 2014).

Algoritma C4.5 dan *decision tree* merupakan dua model yang tidak terpisahkan. Karena untuk membangun *decision tree* dibutuhkan algoritma C4.5 (Anik Andriani, 2013). Di akhir tahun 1970 hingga di awal tahun 1980, J Ross Quinlan seorang peneliti di bidang mesin pembelajaran membuat pengembangan

sebuah *decision tree* yang dinamakan *ID3 (Iterative Dichotomizer)*, walaupun sebenarnya proyek ini telah dibuat sebelumnya oleh E.B. Hunt, J. Marin dan PT. Stone, kemudian Quinlan membuat algoritma dari pengembangan *ID3* yang dinamakan C4.5 yang berbasis *supervised learning* (Anik Andriani, 2013).

Serangkaian perbaikan yang dilakukan pada *ID3* mencapai puncaknya dengan menghasilkan sebuah sistem praktis dan berpengaruh untuk *decision tree*-yaitu C4.5. Adapun beberapa perbaikannya adalah sebagai berikut (eko prasetyo, 2014) :

1. Algoritma C4.5 menghitung *gain ratio* untuk masing-masing atribut, dan atribut yang memiliki nilai tertinggi yang akan dipilih sebagai simpul atau *node*. Penggunaan *gain ratio* ini memperbaiki kelemahan dari *ID3* yang menggunakan *information gain*.
2. Pemangkasan (*pruning*) dapat dilakukan pada saat pembangunan pohon (*tree*) ataupun pada saat proses pembangunan *tree* selesai.
3. Mampu menangani *continues attribute*.
4. Mampu menangani *missing data*.
5. Mampu membangkitkan *rule* dari sebuah *tree*.

Ada beberapa tahapan dalam membuat sebuah *decision tree* dengan algoritma C4.5 yaitu (eko prasetyo, 2014) :

1. Mempersiapkan *data training*. *Data Training* biasanya diambil dari data histori yang pernah terjadi sebelumnya atau disebut data masa lalu dan sudah dikelompokkan dalam kelas-kelas tertentu.
2. Menghitung akar dari pohon. Akar akan diambil dari atribut yang dipilih, dengan cara menghitung nilai *gain* yang paling tinggi yang akan menjadi akar pertama. Sebelum menghitung nilai *gain* dari atribut, hitung dulu nilai *entropy*. Untuk menghitung nilai *entropy* digunakan rumus :

$$Entropy(S) = \sum_{i=1}^n -P_i * \log_2 P_i$$

Dimana :

S = Himpunan kasus

N = Jumlah partisi S

P_i = Proporsi S_i terhadap S

3. Menghitung nilai *Gain* menggunakan rumus :

$$Gain(S,A) = Entropy(S) - \sum_{i=1}^n \frac{|S_i|}{S} * Entropy(S_i)$$

Dimana :

S = Himpunan kasus

A = fitur

N = Jumlah partisi atribut

$|S_i|$ = proporsi S_i terhadap S

$|S|$ = Jumlah kasus dalam S

4. Ulangi langkah ke-2 dan langkah ke-3 hingga semua *record* terpartisi
5. Proses partisi *decision tree* akan berhenti pada saat:
- Semua *record* dalam simpul N mendapat kelas yang sama.
 - Tidak ada atribut didalam *record* yang dipartisi lagi.
 - Tidak ada *record* didalam cabang yang kosong.

2.2.6 Aplikasi

Aplikasi adalah program siap pakai yang dapat di gunakan untuk menjalankan perintah-perintah dari pengguna aplikasi tersebut dengan tujuan mendapatkan hasil yang lebih akurat sesuai dengan pembuatan aplikasi tersebut, aplikasi adalah program komputer yang di tulis dalam suatu bahasa pemograman dan di pergunakan untuk menyelesaikan masalah tertentu (Robi Muhamad, 2009).

2.2.7 Web

Secara *basic*, *website* di gunakan untuk publikasi informasi. Adapun informasi yang akan di sediakan adalah beraneka ragam dari *profil* pribadi hingga *company profil*. *Website* sering juga di sebut *web*, dapat diartikan suatu kumpulan halaman-halaman yang menampilkan berbagai macam informasi teks, data gambar diam ataupun bergerak, data dinamis, suara maupun gabungan dari semuanya, baik itu yang bersifat statis maupun dinamis yang dimana membentuk suatu rangkaian bangunan yang saling berkaitan dimana masing-masing dihubungkan dengan jaringan halaman atau *hyperlink* (Nurhadi, 2017).

Definisi lainnya dari *website* adalah kumpulan dari berbagai macam halaman situs, yang terangkum dari sebuah *domain* atau juga *subdomain* yang lebih tepatnya berada di dalam WWW (*World Wide Web*) yang tentunya terdapat di dalam internet.

Halaman website bisanya berupa dokumen yang ditulis dalam format *Hyper Text Markup Language* (HTML), yang bisa di akses menggunakan HTTP, HTTP adalah suatu protokol yang menyampaikan berbagai informasi dari *server website* untuk ditampilkan kepada para user atau pemakai melalui *web browser*. Pada dasarnya *website* dibagi menjadi dua bagian, yaitu :

1. *Website Statis*

Merupakan *web* yang halamannya tidak berubah, biasanya untuk melakukan perubahan di lakukan secara manual dengan mengubah kode. *Website* statis informasinya merupakan informasi satu arah, yakni hanya berasal dari pemilik *software*nya saja, hanya bisa di *update* oleh pemiliknya saja, contoh dari pengertian *website* statis ini adalah *profil* perusahaan.

2. *Website Dinamis*

Merupakan *web* yang halamannya selalu *update*, biasanya terdapat halaman *backend* (halaman administrator) yang di gunakan untuk menambah atau merubah konten. *Web* dinamis membutuhkan *database* untuk menyimpan data. Web dinamis mempunyai arus informasi dua arah, yakni berasal dari pengguna dan pemilik, sehingga peng-*update*-an dapat di lakukan oleh pengguna dan pemilik *website*. Contoh dari pengertian *website* dinamis ini yaitu *friendster*, *Multipl*y, *Facebook* dan lain-lain.

2.2.8 *My Structured Query Language (MySQL)*

MySQL adalah Sebuah program *database server* yang mampu menerima dan mengirimkan datanya sangat cepat, *multi user* serta menggunakan perintah dasar SQL (*Structured Query Language*). MySQL merupakan dua bentuk lisensi, yaitu *FreeSoftware* dan *Shareware*. MySQL yang biasa kita gunakan adalah MySQL

FreeSoftware yang berada dibawah Lisensi GNU/GPL (*General Public License*), (Haris Spuro, 2012).

MySQL Merupakan sebuah *database server* yang free, artinya kita bebas menggunakan database ini untuk keperluan pribadi atau usaha tanpa harus membeli atau membayar lisensinya. MySQL pertama kali dirintis oleh seorang *programmer-database* bernama Michael Widenius, Selain *database server*, MySQL juga merupakan program yang dapat mengakses suatu *database* MySQL yang berposisi sebagai *Server*, yang berarti program kita berposisi sebagai *Client*. Jadi MySQL adalah sebuah *database* yang dapat digunakan sebagai *Client* maupun *server*.

Database MySQL merupakan suatu perangkat lunak *database* yang berbentuk *database relasional* atau disebut *Relational Database Management System* (RDBMS) yang menggunakan suatu bahasa permintaan yang bernama SQL (*Structured Query Language*).

Kelebihan MySQL

Database MySQL memiliki beberapa kelebihan dibanding database lain, diantaranya :

- a. MySQL merupakan *Database Management System* (DBMS)
- b. MySQL sebagai *Relation Database Management System* (RDBMS) atau disebut dengan *database Relational*
- c. MySQL Merupakan sebuah *database server* yang *free*, artinya kita bebas menggunakan database ini untuk keperluan pribadi atau usaha tanpa harus membeli atau membayar lisensinya
- d. MySQL merupakan sebuah *database client*
- e. MySQL mampu menerima *query* yang bertupuk dalam satu permintaan atau *Multi Threading*.
- f. MySQL merupakan *Database* yang mampu menyimpan data berkapasitas sangat besar hingga berukuran *Giga Byte* sekalipun.
- g. MySQL di dukung oleh *driver* ODBC, artinya *database* MySQL dapat diakses menggunakan aplikasi apa saja termasuk berupa visual seperti *visual Basic* dan *Delphi*.

- h. MySQL adalah *database* menggunakan *enkripsi password*, jadi database ini cukup aman karena memiliki *password* untuk mengakses • MySQL merupakan *Database Server* yang *multi user*, artinya database ini tidak hanya digunakan oleh satu pihak orang akan tetapi dapat digunakan oleh banyak pengguna.
- i. MySQL mendukung *field* yang dijadikan sebagai kunci *premier* dan kunci unqi (*Unique*).
- j. MySQL memiliki kecepatan dalam pembuatan table maupun *update-an* table.

2.2.9 Hypertext Preprocessor (PHP)

Bahasa pemograman PHP merupakan bahasa pemograman yang cukup populer dan banyak di gunakan oleh para *programmer* di dunia. PHP atau yang memiliki kepanjangan *Hypertext Preprocessor*, merupakan suatu bahasa pemograman yang di fungsikan untuk membangun suatu website dinamis, PHP menyatu degan kode HTML. HTML digunakan seagai pembangun atau pondasi dari kerangka *layout web*, sedangkan PHP difungsikan sebagai prosesnya, sehingga dengan adanya PHP tersebut sebuah web akan mudah di *maintenance*, (Agus Saputra, 2013)

PHP berjalan pada sisi *server* sehingga PHP disebut juga sebagai bahasa *Sever Side Scripting*, artinya bahwa dalam setiap/unuk menjalankan PHP, akan mebutuhka web *server* untuk menjalankannya. PHP ini besifat *open source* sehingga dapat di pakai secara gratis, dan mampu lintas *platform*, yaitu dapat berjalan pada sistem operasi *widow* maupun *linux*.

2.2.10 Unifed Modeling Language (UML)

Menurut (Ivar Jacobson, 2010). *Unified Modeling Language* (UML) adalah bahasa standar untuk menulis cetak biru perangkat lunak. UML dapat digunakan untuk memvisualisasikan, menentukan, membuat, dan mendokumentasikan artefak dari sistem intensif perangkat lunak. UML hanya bahasa dan hanya satu bagian dari

metode pengembangan perangkat lunak. Secara khusus, *Unified Modeling Language* (UML) menspesifikasikan langkah-langkah dalam pengambilan keputusan analisis, perancangan serta implementasi dalam sistem yang sangat bernuansa perangkat lunak (*Software intensiv sistem*). Dalam hal ini, *unified Modeling Language* bukanlah bahasa pemrograman tetapi model-model yang tercipta berhubungan langsung dengan bahasa pemrograman.

Kosakata UML mencakup tiga jenis blok bangunan:

1. *Things*

Ada empat macam *thing* dalam UML:

a. *Structural things*

Structural things adalah kata benda dari model UML

b. *Behavioral things*

Behavioral things bagian dinamis dari model UML. Ini adalah kata kerja dari model, mewakili perilaku dari waktu ke waktu dan ruang.

c. *Grouping things*

Grouping things adalah bagian organisasi dari model UML. Ini adalah kotak di mana model dapat diuraikan.

d. *Annotational things*

Annotational things adalah bagian penjelasan dari model UML. Ini adalah komentar yang dapat Anda terapkan untuk menggambarkan, menerangi, dan berkomentar tentang elemen apa pun dalam model *Relationships*

Ada empat macam *relationships* dalam UML:

a. *Dependency*

Dependency adalah hubungan semantik antara dua hal di mana perubahan ke satu hal (hal yang independen) dapat mempengaruhi semantik dari hal lain (hal yang tergantung).

b. *Association*

Association adalah hubungan struktural yang menggambarkan sekumpulan tautan, tautan yang menjadi koneksi antar objek.

c. *Generalization*

Generalization adalah hubungan spesialisasi / generalisasi di mana objek-objek elemen khusus (anak) dapat disubstitusikan untuk objek-objek elemen generalisasi (induk).

d. *Realization*

Realization adalah hubungan semantik antara pengklasifikasi, di mana satu pengklasifikasi menentukan kontrak yang dijamin untuk diklasifikasi oleh pengklasifikasi lain

2. *Diagrams*

Diagram adalah presentasi grafis dari sekumpulan elemen, paling sering disajikan sebagai grafik terhubung dari simpul (benda) dan busur (hubungan).

Ada 3 *diagram* UML yang di gunakan:

a. *Class diagram*

Class adalah sebuah spesifikasi yang jika diinstansiasi akan menghasilkan sebuah objek dan merupakan inti dari pengembangan dan desain berorientasi objek. *Class* menggambarkan keadaan (atribut/properti) suatu sistem, sekaligus menawarkan layanan untuk memanipulasi keadaan tersebut (metoda/fungsi).

Class memiliki tiga area pokok :

- a. Nama (dan *stereo type*)
- b. Atribut
- c. Metoda

Atribut dan metoda dapat memiliki salah satu sifat berikut :

- a. *Private*, tidak dapat dipanggil dari luar class yang bersangkutan
- b. *Protected*, hanya dapat dipanggil oleh class yang bersangkutan dan anak-anak yang mewarisinya

c. *Public*, dapat dipanggil oleh siapa saja

b. *Activity diagram*

Activity Diagram adalah diagram yang menggambarkan *workflow* (aliran kerja) atau aktivitas dari sebuah sistem atau proses bisnis.

Komponen yang ada pada activity diagram antara lain:

- a. *Activity* atau *state*: Menunjukkan aktivitas yang dilakukan.
- b. *Initial activity* atau *initial state*: Menunjukkan awal aktivitas dimulai.
- c. *Final Activity* atau *final state*: Menunjukkan bagian akhir dari aktivitas.
- d. *Decission*: Digunakan untuk menggambarkan test kondisi untuk memastikan bahwa control flow atau object flow mengalir lebih ke satu jalur. Jumlah jalur sesuai yang diinginkan.
- e. *Merge*: Berfungsi menggabungkan flow yang dipecah oleh decission.
- f. *Synchronization*: dibagi menjadi 2 yaitu fork dan join. Fork digunakan untuk memecah behaviour menjadi activity atau action yang paralel, sedangkan join untuk menggabungkan kembali activity atau action yang paralel.
- g. *Swimlanes*: Memecah activity diagram menjadi baris dan kolom untuk membagi tanggung jawab obyek-obyek yang melakukan aktivitas.
- h. *Transition*: Menunjukkan aktivitas selanjutnya setelah aktivitas sebelumnya.

2.2.11 *Model Driven Development (MDD)*

Teknik pengembangan berbasis model (MDD) menekankan gambar model untuk membantu memvisualisasikan dan menganalisis masalah, mendefinisikan kebutuhan bisnis, dan merancang sistem informasi. Analisis dan desain sistem terstruktur - berpusat pada proses Teknik informasi (IE) - berpusat pada data Analisis dan desain berorientasi obyek (OOAD) - terpusat pada objek (integrasi data dan masalah proses) Rute model driven development.

Tahapan yang dipakai pada penelitian adalah sebagai berikut:

1. *Preliminary investigation* (investigasi awal)
2. *Problem analysis* (Analisis masalah)
3. *Requirements analysis* (Analisis Kebutuhan)
4. *Design* (Desain)
5. *Construction* (Kontruksi)
6. *Implementation* (implementasi)

1. *Preliminary Investigation Phase*

Tahap ini merupakan tahap awal dari pengembangan sistem. Fase ini berisikan investigasi awal ketika ingin merancang sebuah sistem, seperti wawancara, tinjauan langsung dan mempelajari dokumen perusahaan.

Tujuan dari tahap ini ialah menjawab pertanyaan mengenai apakah proyek ini cukup berharga untuk diperhatikan. Untuk menjawab pertanyaan ini perlu didefinisikan terlebih dahulu masalah, kesempatan, dan resiko-resiko dalam melanjutkan proyek. Kerangka kerja PIECES dapat digunakan untuk menjawab pertanyaan ini namun hasilnya bukanlah solusi permasalahan melainkan kategori-kategori masalah (dengan asumsi bahwa proyek ini berharga untuk diperhatikan) menetapkan rincian proyek yang akan menetapkan lingkup, kebutuhan dan hambatan proyek, anggota proyek, biaya, dan jadwal.

2. *Problem Analysis Phase*

Problem Analysis ialah menganalisa masalah-masalah yang terdapat di lapangan. Tahap ini merupakan pengembangan dari tahap pertama. Pada tahap ini dilakukan analisis terhadap sistem yang telah ada saat itu. Tahap ini memberikan pemahaman yang lebih dalam bagi tim proyek mengenai permasalahan yang dihadapi. Analisis ini dilakukan untuk menjawab pertanyaan apakah keuntungan yang diperoleh setelah pemecahan masalah lebih besar daripada biaya yang dikeluarkan.

Input utama dari tahap ini adalah project charter dari tahap sebelumnya. Informasi yang digunakan dalam mempelajari permasalahan yang dihadapi adalah fakta-fakta yang terdapat dalam sistem, masalah, akibat, penyebab dari permasalahan, dan spesialis IT yang merancang sistem yang telah ada.

Output yang dihasilkan adalah system improvement objectives yang menyatakan kriteria bisnis yang akan digunakan untuk mengevaluasi sistem. Kadang-kadang dilakukan representasi pada tahap ini.

3. *Requirement Analysis Phase*

Requirement Analysis ialah melakukan analisa terhadap kebutuhan perusahaan. Pekerjaan pada tahap ini adalah mendefinisikan apa saja yang perlu dilakukan oleh sistem, apa yang dibutuhkan dan diinginkan oleh pengguna dari sistem baru.

Tahap ini memerlukan perhatian yang besar karena jika terjadi kesalahan dalam menerjemahkan kebutuhan dan keinginan pengguna sistem maka dapat mengakibatkan adanya rasa tidak puas pada sistem final dan perlu diadakan modifikasi yang tentunya akan kembali mengeluarkan biaya.

- a. Input dari tahap ini adalah system improvement objectives yang dihasilkan pada tahap sebelumnya. Pada tahap ini, tim akan

mengumpulkan dan mendiskusikan kebutuhan dan prioritas berdasarkan informasi yang diperoleh dari kuesioner, wawancara, dan rapat-rapat. Tantangannya adalah untuk memvalidasi semua kebutuhan informasi.

- b. Output yang dihasilkan dari tahap ini adalah business requirement statement. Tahap ini pun merupakan tahap yang penting karena dapat menimbulkan ketidakpuasan dari pengguna sistem yang merasa kebutuhannya tidak terpenuhi. Tim proyek harus dapat membedakan antara apa yang dibutuhkan oleh pengguna dan bagaimana sebaiknya sistem yang baru bekerja.

4. *Desain Phase*

Setelah diperoleh proposal sistem yang disetujui, maka dapat mulai dilakukan proses desain dari sistem target. Tujuan dari tahap ini adalah untuk mentransformasikan business requirement statement menjadi spesifikasi desain untuk proses konstruksi. Dengan kata lain, tahap desain menyatakan bagaimana teknologi akan digunakan dalam sistem yang baru. Tahap ini memerlukan ide dan opini dari pengguna, vendor, dan spesialis IT.

Pada akhir tahap ini masih terdapat beberapa alternatif keputusan mengenai proyek walaupun pembatalan proyek jarang dilakukan pada tahap ini (kecuali benar-benar over budget atau sangat terlambat dari jadwal). Perubahan lingkup menjadi lebih kecil masih dapat terjadi. Selain itu, mungkin juga terjadi perubahan ulang jadwal untuk menghasilkan solusi yang lebih lengkap.

5. *Construction Phase*

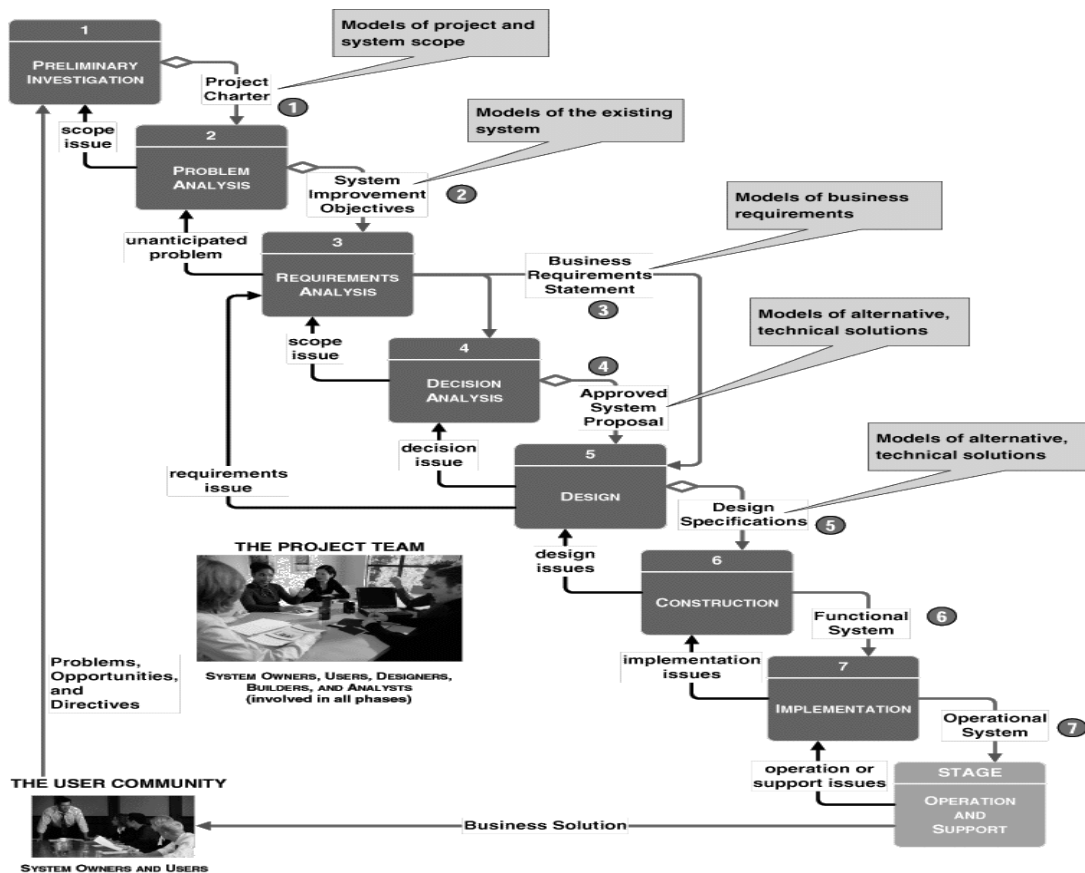
Construction Phase ialah tahapan melaksanakan pengujian pada komponen sistem secara individu dan sistem secara keseluruhan.

Tujuan dari tahap ini adalah :

- a. Membangun dan menguji sistem yang memenuhi business requirement dan spesifikasi desain

- b. Mengimplementasikan penghubung antara sistem baru dan sistem lama, termasuk instalasi dari software yang dibeli atau disewa
- c. Pada tahap ini dilakukan konstruksi basis data, program aplikasi, dan penghubung antara sistem dan pengguna. Beberapa dari komponen ini telah ada sebelumnya.

Setelah dilakukan pengujian, maka sistem dapat mulai diimplementasikan



Gambar 2.2 model driven development (mdd)

2..2.12 Flowmap

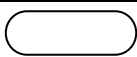
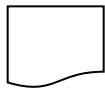

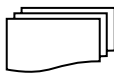
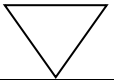


Flowmap atau bagan alir adalah bagan yang menunjukkan aliran di dalam program atau prosedur sistem secara logika. Flowmap ini berfungsi untuk memodelkan masukan, keluaran, proses maupun transaksi dengan menggunakan simbol-simbol tertentu. Pembuatan flowmap ini harus dapat memudahkan bagi pemakai dalam memahami alur dari sistem atau transaksi.

Adapun pedoman-pedoman dalam pembuatan flowmap adalah sebagai berikut:

1. *Flowmap* sebaiknya digambarkan dari atas ke bawah dan mulai dari bagian kiri dari suatu halaman.
2. Kegiatan di dalam *flowmap* harus ditunjukkan dengan jelas.
3. Harus ditunjukkan dari mana kegiatan akan dimulai dan dimana akan berakhir.
4. Masing-masing kegiatan didalam *flowmap* sebaiknya digunakan suatu kata yang mewakili suatu pekerjaan.
5. Masing-masing kegiatan didalam flow map harus didalam urutan yang semestinya.
6. Kegiatan yang terpotong dan akan disambung ditempat lain harus ditunjukkan dengan jelas menggunakan simbol penghubung.
7. Gunakan simbol simbol *flowmap* yang standar.

Adapun simbol-simbol yang sering digunakan dalam *flowmap* dapat dilihat pada tabel 2.3 berikut ini:

Tabel 2.3 Simbol Flowmap

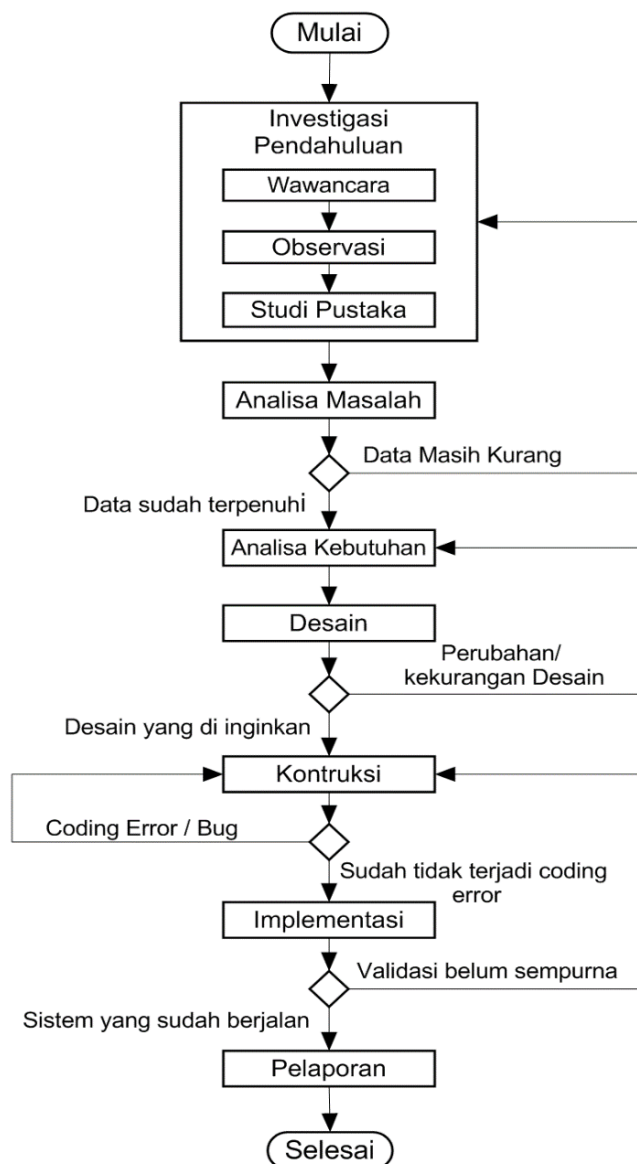
Simbol	Deskripsi
	Simbol yang digunakan untuk menunjukkan awal atau akhir dari suatu proses
	Menunjukkan dokumen input dan output baik untuk proses manual mekanik atau komputer
	Menunjukkan pekerjaan manual
	Menunjukkan multi dokumen
	Pengarsipan Dat
	Menunjukkan Proses
	Simbol input/output digunakan untuk mewakili data input/output

BAB III

METODOLOGI PENELITIAN

3.1 Kerangka Berpikir

Bab ini berisi gambaran kerangka berpikir dan deskripsinya. Berikut adalah langkah-langkah yang dilakukan untuk mencapai tujuan dari penelitian ini adalah sebagai berikut :



Gambar 3.3 Kerangka Berpikir

3.2 Penjelasan Skema Penelitian

3.2.1 Investigasi Pendahuluan

Penulis memulai pekerjaan dengan mengumpulkan semua data-data yang dibutuhkan untuk penelitian ini dengan mengunjungi Fakultas Teknologi Informasi Universitas Bale Bandung (UNIBBA) pada hari Sabtu tanggal 21 Mei 2019 untuk bagian tata usaha yang mengelola dan menangani data mahasiswa yaitu Bapak Yusuf Muharom, S.Kom. Adapun langkah –langkah dalam mengumpulkan data adalah sebagai berikut:

1. Wawancara

Wawancara dilakukan langsung dengan pihak-pihak terkait dalam menganalisa dan mengumpulkan data mahasiswa dan yang dimulai pada yaitu Bapak Yusuf Muharom S.Kom, sebagai bagian tata usaha di Fakultas Teknologi Informasi yang mengatur data-data mahasiswa Fakultas Teknologi Informasi Universitas Bale Bandung. Dalam melakukan wawancara penulis mengajukan beberapa pertanyaan diantaranya sebagai berikut:

1. Apakah Fakultas Teknologi Informasi Universitas Bale Bandung memiliki model untuk memprediksi tingkat kelulusan mahasiswa ?
2. Apa yang menjadi penyebab utama mahasiswa tidak lulus tepat waktu ?
3. Apakah bagian tata usaha Fakultas Teknologi Informasi Universitas Bale Bandung menangani juga data Nilai IPK dan Nilai TOEFL mahasiswa ?
4. Apakah akan berpengaruh pada proses akreditasi jika Fakultas Teknologi Informasi Universitas Bale Bandung memiliki suatu model untuk memprediksi tingkat kelulusan mahasiswa ?

5. Apakah Fakultas Teknologi Informasi Universitas Bale Bandung membutuhkan suatu model untuk memprediksi mahasiswa ?

2. Observasi

Observasi dilakukan langsung ke Fakultas Teknologi Informasi Universitas Bale Bandung (UNIBBA) divisi akademik guna mengetahui model seperti apa yang digunakan Fakultas Teknologi Informasi Universitas Bale Bandung dalam memprediksi tingkat kelulusan mahasiswa.

3. Studi Pustaka

Mempelajari teori-teori dan hasil penelitian sebelumnya yang mendukung pemecahan masalah bagi penelitian ini yang terdiri dari buku-buku digital (*ebook*), jurnal dan beberapa referensi dari website yang berkaitan dengan masalah-masalah dalam pemodelan untuk memprediksi tingkat kelulusan mahasiswa Fakultas Teknologi Informasi Universitas Bale Bandung.

3.2.2 Analisa Masalah

Hasil dari analisis masalah yang ada pada pembuatan model keputusan untuk memprediksi tingkat kelulusan mahasiswa di Fakultas Teknologi Informasi. Fakta-fakta yang sudah ditemukan diantaranya :

1. Fakultas Teknologi Informasi Universitas Bale Bandung tidak memiliki model keputusan untuk memprediksi tingkat kelulusan mahasiswa-nya.

Berdasarkan permasalahan tersebut, penulis melakukan perbandingan dengan data yang telah didapat pada proses pengumpulan data dan melakukan studi banding dengan penelitian-penelitian sebelumnya, maka dapat disimpulkan bahwa-

dalam melakukan proses memprediksi tingkat kelulusan mahasiswa dibutuhkan algoritma yang dapat memprediksi suatu keputusan dengan tingkat akurasi yang tinggi, untuk itu penulis menggunakan algoritma C4.5 sebagai solusi untuk permasalahan yang ada pada sistem yang sedang berjalan.

3.2.3 Analisa Kebutuhan

Adapun analisis kebutuhan yang diperlukan untuk menjawab pertanyaan dan permasalahan yang ada terdiri dari 3, yaitu:

1. Kebutuhan Prosedur

Prosedur yang dibutuhkan yaitu mengumpulkan data-data yang akan digunakan. Berikut data-data yang dibutuhkan dalam proses memprediksi tingkat kelulusan mahasiswa:

- a. Tahun mahasiswa sedang melakukan studi
- b. NIM mahasiswa
- c. Semester
- d. Nilai IPK
- e. Nilai TOEFL

2. Kebutuhan Perangkat Lunak

Adapun perangkat lunak (*Software*) yang dibutuhkan pada penelitian ini sebagai berikut:

- a. Microsoft Excel 2013, untuk melakukan perhitungan data
- b. *Balsamiq Mockup* untuk merancang tampilan *web*
- c. Bahasa pemrograman PHP
- d. *Database Mysql* untuk menyimpan data
- e. Microsoft Word 2013, untuk pembuatan laporan

3. Kebutuhan Perangkat Keras

Adapun kebutuhan minimum perangkat keras (*hardware*) dalam proses perhitungan data dapat di lihat pada table 3.4.

Tabel 3.4 Spesifikasi sistem

CPU	1 GHz clock speed, IA-32 or x64 architecture with SSE2 support
RAM	IA-32 edition: 1GB X64 edition: 2GB
Hardisk Drive	3.0 GB free disk space
Operating System	Windows 7 Windows server 2008 R2 Windows 8 Windows server 2012

3.2.4 Desain

Pada tahap ini penyusun membuat desain aplikasi sistem prediksi tingkat kelulusan mahasiswa yang akan dibangun dari hasil analisis yang telah dilakukan. Ada beberapa proses tahapan yang akan dilakukan diantaranya sebagai berikut :

1. Perancangan Model

Pada perancangan model dilakukan untuk mengetahui alur yang akan dibuat untuk membangun aplikasi sistem prediksi kelulusan mahasiswa. Perancangan akan dibuat dengan *Unified Modeling Language* (UML) agar dengan mudah dalam proses pengembangan dan visualisasinya. Diagram UML yang digunakan antara lain:

- a. *Use Case Diagram* digunakan untuk menggambarkan keterhubungan *actor* dan *use case* dalam aplikasi yang akan dibuat. *Actor* dan *use case* yang terlibat pada penelitian ini adalah:
 1. Actor : Staf Tu dan Pimpinan,.
 2. *Use case*: *input* data mahasiswa, mengelola data, melakukan perhitungan dan pembentukan pohon keputusan.
- b. *Class Diagram* digunakan untuk menggambarkan atribut, *operation* dan juga *constraint* pada sistem yang akan dibuat. *Class diagram*

yang akan dibuat *class diagram* struktur sistem aplikasi yang diusulkan

- c. *Activity Diagram* digunakan untuk memodelkan alur kerja dari sistem dan aktivitas dari *actor* dalam aplikasi. Adapun diagram alur yang akan dibuat diantaranya:

1. Proses login
2. Proses impor data (Format file .xls microsoft excel)
3. Proses mengelola data
4. Proses perhitungan C4.5
5. Proses *Logout*

2. *Database*

Database adalah salah satu bagian yang paling penting dalam membangun aplikasi prediksi kelulusan. Pengelolaan data juga menjadi bagian paling utama dalam penelitian ini. Dalam perancangan *database*, data yang dibutuhkan didapatkan dari pengumpulan data sebelumnya. Beberapa data yang digunakan dalam penelitian ini adalah:

- a. Data tahun mahasiswa melakukan studi
- b. Data semester mahasiswa
- c. Data NIM mahasiswa
- d. Data Nilai TOEFL
- e. Data Nilai IPK

3. Perancangan Antar Muka (*User Interface*)

Perancangan antar muka adalah perancangan tahap akhir yang akan dibuat. Adapun perancangan antar muka yang akan dibuat sebagai berikut:

- a. Form login
- b. Form halaman menu utama
- c. Form dataset
- d. Form halaman atribut
- e. Form halaman nilai atribut
- f. Form halaman perhitungan

3.2.5 Konstruksi

Input sistem yang akan dibangun yaitu berupa data mahasiswa yang meliputi Nilai IPK, NIM, Tahum, Semester, Nilai TOEFL tersebut akan dibuat tabel perhitungan dan dihitung menggunakan algoritma c4.5 dimana setiap atribut akan dicari nilai entropy dan informasi gain. Setelah dilakukan perhitungan maka akan menghasilkan suatu kondisi atau rule yang digunakan dalam proses memprediksi. Output untuk sistem ini berupa hasil persenan tingkat akurasi kelulusan mahasiswa yang diprediksi. Setelah tahap perancangan selesai maka tahap selanjutnya adalah pembuatan aplikasi.

Pada pembuatan aplikasi digunakan perangkat lunak dan bahasa pemrograman sebagai berikut:

1. *Sublime Text*, digunakan untuk menulis *source code* aplikasi yang akan dibangun.
2. *Database MYSQL*, digunakan untuk menyimpan data yang terdiri dari data mahasiswa.
3. *PHP*, bahasa pemrograman yang digunakan untuk membuat aplikasi yang akan dibangun.
4. *XAMPP*, digunakan untuk membuat database dan menjalankan bahasa pemrograman php.
5. *Microsoft excel 2010*, digunakan untuk melakukan perhitungan manual pada algoritma C4.5

3.2.6 Implementasi

Setelah aplikasi dibuat pada tahap sebelumnya, tahapan selanjutnya yaitu membuat pengujian terhadap aplikasi yang sudah dibuat, yaitu dilakukan pengujian dengan menggunakan metode *black box* yaitu untuk menguji fungsionalitas dari suatu aplikasi.

Jika di dalam tahap ini masih ada kekurangan maka akan kembali ke tahap investigasi awal sampai mendapat hasil yang tepat, kemudian di lanjutkan ketahap berikutnya.

3.2.7 Pembuatan Laporan

Tahap ini dibuat laporan untuk mempertanggung jawabkan hasil dari penelitian yang telah dilakukan di Fakultas Teknologi informasi Universitas Bale Bandung (UNIBBA). Laporan yang dibuat dituangkan ke dalam bentuk skripsi yang nantinya akan di uji dalam sidang skripsi sebagai syarat memperoleh gelar sarjana Strata 1 (S1) Teknik Informatika Pada tahap ini merupakan tahap terakhir dalam melakukan penelitian yang terdiri dari 6 bab. Berikut adalah sistematika penulisan:

BAB I PENDAHULUAN

Bab ini membahas tentang latar belakang masalah, rumusan masalah, batasan masalah, tujuan, metodologi penelitian dan sistematika penulisan yang akan dibuat pada penelitian ini.

BAB II LANDASAN TEORI

Bab ini membahas tentang teori dan studi literatur yang sesuai dan akurat sehingga bisa mendukung penelitian dalam penulisan sehingga menghasilkan karya tulis yang bernilai ilmiah.

BAB III METODOLOGI PENELITIAN

Bab ini berisi uraian tentang kerangka berpikir, Investigasi pendahuluan, analisa masalah, analisa kebutuhan, desain, konstruksi, implementasu dan pembuatan laporan.

BAB IV ANALISIS DAN PERANCANGAN

Bab ini membahas tentang analisis lebih lanjut terhadap masalah yang ada, perangkat lunak, perangkat keras, dan berbagai macam analisis yang di butuhkan untuk pembuatan aplikasi. Juga menjelaskan tentang perancangan di mulai dari flowmap diagram, *activity diagram* dan lain sebagainya.

BAB V IMPLEMENTASI DAN PENGUJIAN

Bab ini membahas tentang implementasi dan pengujian dari aplikasi yang telah dibuat.

BAB VI KESIMPULAN DAN SARAN

Bab ini membahas tentang kesimpulan yang diperoleh dari pembuatan sistem ini serta saran yang bertujuan untuk pengembangan sistem dimasa yang akan datang.

BAB IV

ANALISIS DAN PERANCANGAN

4.1 Analisis

4.1.1 Analisa Masalah

Langkah awal dalam pembuatan sistem adalah meng-identifikasi permasalahan yang ada sebagai dasar untuk membuat sebuah solusi yang akan disajikan dalam bentuk sebuah aplikasi. Langkah identifikasi dilakukan dengan cara observasi dan wawancara ke bagian tata usaha Fakultas Teknologi Informasi Universitas Bale Bandung, seperti yang telah disebutkan pada bab sebelumnya. Sehingga dapat dilakukan tindakan pengambilan solusi yang tepat untuk menyelesaikan permasalahan tersebut.

Menurut proses analisis yang dilakukan dalam operasional tentang hal yang paling berpengaruh dalam menentukan kelulusan mahasiswa Fakultas Teknologi Informasi Universitas Bale Bandung terdapat dua *indikator*, yaitu Nilai IPK dan Nilai TOEFL.

4.1.2 Analisa Sistem

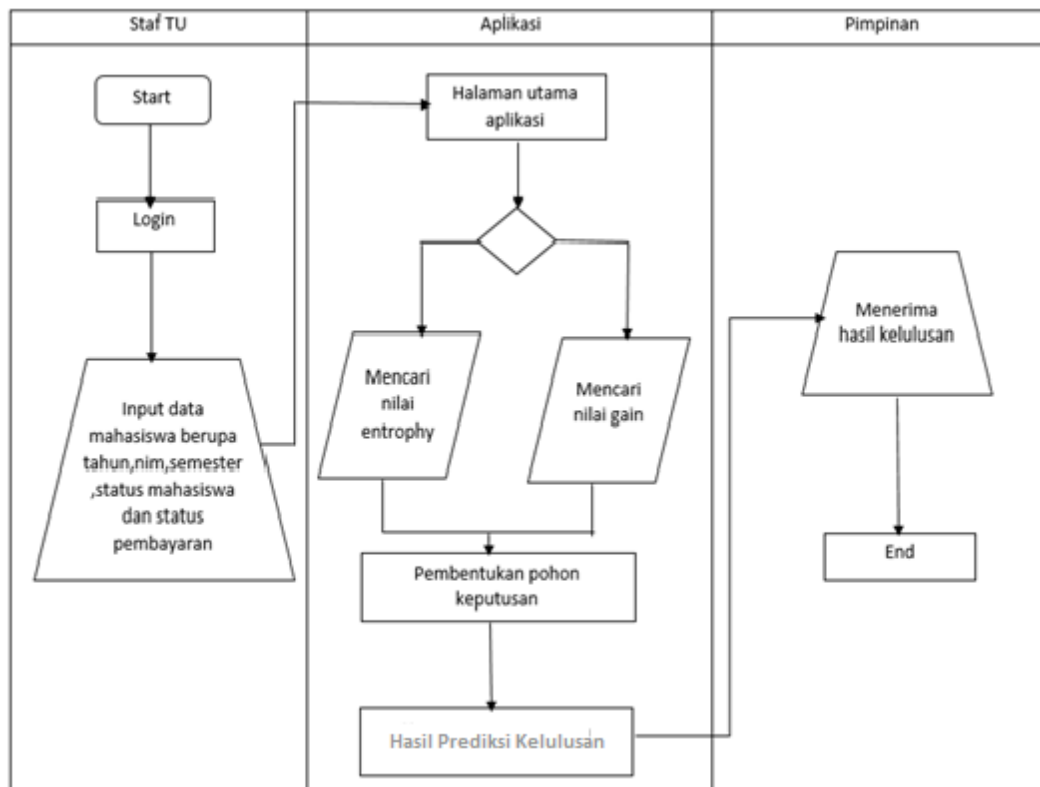
Pada perancangan dan perhitungan yang akan dibuat, menggunakan perangkat lunak dan perangkat keras untuk membuatnya maka dapat di analisa kebutuhan apa saja dari perangkat lunak dan perangkat keras yang dibutuhkan untuk membuat aplikasi ini.

Lalu orang atau kelompok yang melakukan perbaikan atau perancangan suatu sistem di namakan dengan sistem analis. Sistem analis adalah orang atau kelompok yang melaksanakan pengembangan sistem. Sistem analis menekuni permasalahan ataupun kebutuhan dari suatu sistem dan sistem analis juga merupakan orang yang bertanggung jawab atas terjadinya proses analisa ataupun perancangan pada sistem informasi.

Berdasarkan uraian diatas maka dapat di simpulkan bahwa analisis sistem adalah tahap yang bertujuan untuk memahami sistem, mengetahui kekurangan sistem dan menentukan kebutuhan hasil proses pada perangkat lunak dan perangkat keras yang di gunakan. Dengan menganalisis prosedur sistem yang di gunakan dan melakukan pengujian hasil, maka sistem dapat di evaluasi sehingga dapat di jadikan acuan dalam proses pembentukan kesimpulan.

Pada perancangan aplikasi prediksi kelulusan mahasiswa ini karena menggunakan perangkat lunak dan perangkat keras untuk membuatnya maka dapat di analisa kebutuhan apa saja dari perangkat lunak dan perangkan keras yang di butuhkan untuk membuat aplikasi ini.

Pada poin ini juga penyusun berperan sebagai sistem analis yang mana penyusun akan melakukan analisis terhadap sistem yang diusulkan guna mengetahui hal apa saja yang di butuhkan ketika hendak membuat aplikasi, berikut adalah sistem yang diusulkan oleh penyusun :




Gambar 4.4 Sistem yang diusulkan

4.1.3 Analisa Proses

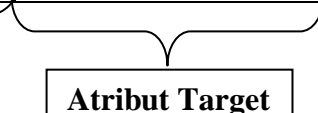
Dalam tahap ini penulis mengambil data di Fakultas Teknologi Informasi Universitas Bale Bandung. Data tersebut kemudian penulis tampung pada aplikasi Microsoft Excel. *Decision tree* memiliki dua jenis *atribut* dari data yang terdiri dari beberapa *atribut* input dari *atribut* target dan tentunya mendukung masalah yang ada, fungsinya sebagai pembanding dalam perhitungan *Gain* dan *Ratio*.

Dalam data *set atribut sample* dan minimal harus memiliki satu atribut target yang nilainya merupakan kesimpulan sementara permasalahan dari setiap *instance* (*record*), dalam penelitian ini nilai dari *atribut* target adalah Lulus dan Tidak Lulus. *Atribut input* yang memiliki *gain ratio* yang terbesar adalah *atribut* yang menjadi akar. Contohnya seperti gambar di bawah ini :

No	IPK	TOEFL	Keterangan
1	≥ 3.00	≥ 400	Lulus
2	Tidak Aktif	Belum Lunas	Tidak Lulus



Atribut Input



Atribut Target

Gambar 4.5 Konsep data decision tree

4.1.4 Analisa Kebutuhan

Analisis Kebutuhan adalah sebuah proses untuk mendapatkan informasi, model dan spesifikasi tentang kebutuhan perangkat lunak dan perangkat keras yang di butuhkan dalam membuat aplikasi pengendalian *inventory* ini. Berikut adalah beberapa software dan hardware yang di butuhkan :

1. Kebutuhan Sistem

Tabel 4.5 Kebutuhan Sistem

No	Perangkat Keras	Perangkat Lunak
1	Seperangkat PC / Laptop	Windows 7 32 bit
2	-	Xampp

3	-	<i>Sublime text</i>
4	-	<i>Web Browser</i>
5	-	<i>Microsoft word 2010</i>
6		<i>Microsoft excel 2010</i>
6	-	<i>Star UML</i>
7	-	<i>Balsamiq mockups 3 V.3.5.17</i>
8	-	<i>Framework CSS</i>

2. Spesifikasi Perangkat Keras

Tabel 4.6 Spesifikasi Perangkat Keras

No	Spesifikasi	Komputer
1	<i>Processor</i>	Intel Celeron 1.80 GHz (2CPUs)
2	<i>Ram (Randaom Acces Memory)</i>	2 GB
3	<i>Hard Disk</i>	320

3. Spesifikasi Perangkat Lunak

Tabel 4.7 Spesifikasi Perangkat Lunak

No	Spesifikasi	Komputer
1	Sistem Operasi	<i>Windows 7 32 bit</i>
2	Bahasa Pemograman	Bahasa Pemograman PHP

4.1.5 Analisa Pengguna

Penganalisaan pengguna adalah yang berkaitan dengan yang akan memakai aplikasi yang akan di buat, pengguna aplikasi ini adalah ;

a. Staf TU

agar aplikasi ini dapat di gunakan dengan baik dan mudah untuk di operasikan penyusun akan menyediakan berbagai menu pada *user interface* yang akan di buat.

4.1.6 User Interface

User interface dari aplikasi sangat berpengaruh terhadap minat dari *user* dalam mengoperasikan-nya. Hal ini mencakup perangkat yang di gunakan sebagai piranti masukan dan keluaran dari aplikasi yang akan di buat. Karena aplikasi yang di buat di tujuan untuk melakukan suatu perhitungan terhadap data yang di *input*-kan kedalam aplikasi.

Untuk perangkat masukan yang di gunakan di antaranya : *Mouse*, agar *user* dapat leluasa menggerakkan petunjuk *cursor* untuk memilih menu atau sub menu yang ada, dan *keyboard* yang akan di gunakan sebagai sarana *input* data yang di perlukan.

4.1.7 Fitur Pada Aplikasi

Fitur yang di gunakan dalam aplikasi ini di maksudkan agar *user* dapat dengan mudah mengoperasikan-nya, di tunjang dengan metode C4.5 atau pembentukan *decision tree* untuk menampilkan hasil perhitungan yang bersifat *prediktif*, bagian terpenting dari fitur-fitur yang ada adalah sebagai berikut :

- a. *Login user*
- b. Menampilkan dataset
- c. Menampilkan atribut
- d. Menampilkan nilai atribut
- e. Melakukan proses perhitungan

Dari hasil yang di dapat dari hal di atas panduan antara user *interface* dengan fitur aplikasi perlu di bentuk untuk menciptakan aplikasi yang dapat memberikan manfaat luas dan nyata bagi pengguna. *Fitur* pada aplikasi berfungsi agar *user* dapat dengan mudah mengoperasikan aplikasi.

4.2 Perancangan Sistem

Perancangan sistem perangkat lunak dideskripsikan dengan model analisis menggunakan diagram *Use Case*. Analisis digunakan untuk pemetaan awal mengenai perilaku yang diisyaratkan sistem aplikasi kedalam elemen-elemen pemodelan. Untuk membantu perancangan dan melengkapi dokumentasi perancangan.

4.2.1 Perancangan

Perancangan data mining ini penulis menggunakan algoritma C4.5 proses pada pohon keputusan adalah mengubah bentuk data (table) menjadi model pohon, mengubah model pohon menjadi rule, dan menyederhanakan rule.

Data akan dibuat pohon keputusan untuk menentukan mahasiswa yang lulus tepat waktu atau tidak dengan melihat nilai IPK mahasiswa dan nilai TOEFL mahasiswa. . Secara umum algoritma C4.5 untuk membangun pohon keputusan adalah sebagai berikut :

- a. Pilih atribut akar.
- b. Buat cabang untuk tiap-tiap nilai.
- b. Bagi kasus dalam cabang.

Ulangi proses untuk setiap cabang sampai semua kasus pada cabang memiliki kelas yang sama. Untuk memilih atribut sebagai akar, didasarkan pada nilai *gain* tertinggi dari atribut-atribut yang ada. Untuk menghitung *gain*, seharusnya mencari nilai *entropy* terlebih dahulu.

Sebelum melakukan pembentukan *decision tree*, penulis menggunakan teknik data *mining* terlebih dahulu untuk memproses data seperti langkah-langkah sebagai berikut :

1. Pemilihan dan Penambahan Data

Data yang akan digunakan dalam proses ini adalah data mahasiswa pada tahun 2018 pada ajaran Semester 8 Genap. Data mahasiswa tersebut diambil dari Fakultas Teknologi Informasi Universitas Bale Bandung. Tidak semua data akan diambil untuk melakukan proses perhitungan, Data yang akan digunakan meliputi nilai IPK mahasiswa dan nilai TOEFL mahasiswa akan digunakan sebagai atribut dan keterangan lulus akan digunakan sebagai atribut target.

Dalam pembangunan aplikasi prediksi kelulusan mahasiswa, IPK menggambarkan peforma akademik mahasiswa sedangkan untuk nilai TOEFL menggambarkan pemahaman mahasiswa dalam memahami *literatur* pembelajaran bahasa inggris.

2. Pembersihan dan Pemrosesan Awal Data

Pembersihan dilakukan untuk membersihkan *noise* pada data. Dalam penelitian ini pembersihan pada data dilakukan menggunakan aplikasi *microsoft excel* 2010. Pada pembersihan tersebut telah diatur sedemikian rupa sehingga setiap data tentang mahasiswa, termasuk data nilai IPK mahasiswa dan nilai TOEFL mahasiswa yang disimpan tidak boleh kosong.

3. Transformasi pada Data

Transformasi data digunakan untuk mengolah data menjadi nilai dengan format tertentu. Dalam proses ini digunakan data yang bersifat *Diskrit*, sehingga data yang bersifat *kontinu* akan diubah menjadi data yang bersifat *diskrit*. Pengelompokan IPK akan dibagi menjadi dua kelas dimana $IPK \geq 3.00$ akan dikelompokkan menjadi IPK tinggi sedangkan $IPK < 3.00$ akan dikelompokkan menjadi IPK rendah. Sedangkan untuk nilai TOEFL akan dibagi menjadi dua kelas yaitu dimana nilai $IPK \geq 400$ dikelompokkan menjadi menengah atas sedangkan $TOEFL < 400$ akan dikelompokkan menjadi menengah bawah.

4. Data Mining

Tahap data *mining* dimulai dengan pembagian pada data tentang mahasiswa yang memiliki atribut target lulus dan tidak lulus yang telah di *transformasikan* sebagai data latih dan data uji. Data latih akan digunakan sebagai pembentukan *decision tree* sedangkan data uji akan digunakan sebagai pengukuran kinerja pada *decision tree*. Setelah data latih telah dibagi akan digunakan dalam proses pembentukan *decision tree*.

5. Analisis Data

Analisis data merupakan tahap untuk menganalisis dan mengidentifikasi data untuk kebutuhan perancangan sistem yang akan dibuat. Dalam hal ini, penulis menggunakan data mahasiswa Fakultas Teknologi Informasi Universitas Bale Bandung sebagai objek penelitian ini dimana data tersebut dapat dilihat pada lampiran 1 data mahasiswa.

6. Pembentukan Decision Tree

Proses klasifikasi tiap-tiap *field* pada tabel mahasiswa

1. Klasifikasi total seluruh kasus

Setelah seluruh kasus telah di klasifikasikan kemudian dicari nilai *entropi* dari keseluruhan kasus tersebut dengan perhitungan sebagai berikut:

Diketahui :

$$\text{Jumlah kasus (a)} = 56$$

$$\text{Jumlah kasus mahasiswa yang tidak lulus (b)} = 11$$

$$\text{Jumlah kasus mahasiswa yang lulus (c)} = 45$$

$$\left(-\frac{b}{a} * \log_2 \left(\frac{b}{a} \right) \right) + \left(-\frac{c}{a} * \log_2 \left(\frac{c}{a} \right) \right)$$

$$= (0.461199) + (0.253528)$$

$$\text{Entropi total} = 0.714727$$

Telah diketahui bahwa nilai *entropi* dari keseluruhan kasus adalah 0.714727.

2. Klasifikasi Nilai IPK Tinggi

Tabel 4.8 Klasifikasi Nilai IPK Tinggi

Klasifikasi IPK Tinggi	Tidak Lulus	Lulus
50	5	45

Setelah seluruh kasus Status IPK tinggi telah di klasifikasikan kemudian dicari nilai *entropi* dari keseluruhan kasus mahasiswa IPK tinggi tersebut dengan perhitungan sebagai berikut:

Diketahui :

Jumlah kasus (a) = 50

Jumlah kasus mahasiswa IPK tinggi yang tidak lulus (b) = 5

Jumlah kasus mahasiswa IPK tinggi yang lulus (c) = 45

$$\left(-\frac{b}{a} * \log_2 \left(\frac{b}{a} \right) \right) + \left(-\frac{c}{a} * \log_2 \left(\frac{c}{a} \right) \right)$$

$$= (0.332193) + (0.136803)$$

Entropi total = 0.468996

Telah diketahui bahwa nilai *entropi* dari keseluruhan kasus mahasiswa yang IPK tinggi adalah 0.468996.

3. Klasifikasi Status IPK rendah

Tabel 4.9 Klasifikasi IPK Rendah

Klasifikasi mahasiswa IPK rendah	Tidak Lulus	Lulus
6	6	0

Setelah seluruh kasus mahasiswa IPK rendah telah di klasifikasikan kemudian dicari nilai *entropi* dari keseluruhan kasus mahasiswa IPK rendah tersebut dengan perhitungan sebagai berikut:

Diketahui :

Jumlah kasus mahasiswa IPK rendah (a) = 6

Jumlah kasus mahasiswa IPK rendah yang tidak lulus (b) = 6

Jumlah kasus mahasiswa IPK rendah yang lulus (c) = 0

$$\left(-\frac{b}{a} * \log_2 \left(\frac{b}{a} \right) \right) + \left(-\frac{c}{a} * \log_2 \left(\frac{c}{a} \right) \right)$$

$$= (0)+(0)$$

Entropi total = 0

Telah diketahui bahwa nilai *entropi* dari keseluruhan kasus mahasiswa yang aktif adalah 0.

4. Klasifikasi Nilai TOEFL

Tabel 4.10 Klasifikasi Nilai TOEFL menengah atas

Klasifikasi Nilai TOEFL menengah atas	Tidak Lulus	Lulus
50	5	45

Setelah seluruh kasus Status nilai TOEFL menengah atas telah di klasifikasikan kemudian dicari nilai *entropi* dari keseluruhan kasus TOEFL menengah atas tersebut dengan perhitungan sebagai berikut:

Diketahui :

Jumlah kasus nilai TOEFL menengah atas (a) = 50

Jumlah kasus TOEFL menengah atas tidak lulus (b) = 5

Jumlah kasus TOEFL menengah atas yang lulus (c) = 45

$$\left(-\frac{b}{a} * \log_2 \left(\frac{b}{a} \right) \right) + \left(-\frac{c}{a} * \log_2 \left(\frac{c}{a} \right) \right)$$

$$= (0.332193)+(0.136803)$$

Entropi total = 0.468996

Telah diketahui bahwa nilai *entropi* dari keseluruhan kasus TOEFL menengah atas adalah 0.468996.

5. Klasifikasi Status TOEFL menengah bawah

Tabel 4.11 Klasifikasi Nilai TOEFL menengah bawah

Klasifikasi Status TOEFL menengah bawah	Tidak Lulus	Lulus
6	6	0

Setelah seluruh kasus Status TOEFL menengah bawah telah di klasifikasikan kemudian dicari nilai *entropi* dari keseluruhan kasus TOEFL menengah bawah tersebut dengan perhitungan sebagai berikut:

Diketahui :

Jumlah kasus TOEFL menengah bawah (a) = 6

Jumlah kasus TOEFL menengah bawah yang tidak lulus (b) = 6

Jumlah kasus TOEFL menengah bawah yang lulus (c) = 0

$$\left(-\frac{b}{a} * \log_2 \left(\frac{b}{a} \right) \right) + \left(-\frac{c}{a} * \log_2 \left(\frac{c}{a} \right) \right)$$

$$= (0)+(0)$$

Entropi total = 0

Telah diketahui bahwa nilai *entropi* dari keseluruhan kasus TOEFL menengah bawah adalah 0.

Setelah masing-masing atribut telah dicari nilai *entropi* nya (tabel 4.12) kemudian dicari nilai *gain* nya.

Tabel 4.12 Entrophy total

Atribut	Total	Lulus	Tidak Lulus	<i>Entropi</i>
Seluruh Kasus	56	45	11	0.714727
Mahasiswa IPK tinggi	50	5	45	0.468996
Mahasiswa IPK rendah	6	0	6	0
TOEFL menengah atas	50	45	5	0.468996
TOEFL menengah bawah	6	0	6	0

Setelah nilai *entropi* dari masing-masing nilai atribut telah diketahui, selanjutnya mencari nilai *gain* dari setiap atribut.

1. Mencari nilai gain Nilai IPK

Diketahui :

<i>Entropi</i> kasus keseluruhan (a)	= 0.714727
<i>Entropi</i> Kasus Nilai IPK tinggi (b)	= 0.468996
<i>Entropi</i> Kasus Nilai IPK rendah (c)	= 0
Jumlah Kasus Keseluruhan (d)	= 56
Jumlah Kasus Nilai IPK tinggi(e)	= 50
Jumlah Status Nilai IPK rendah (f)	= 6

$$Gain \text{ (Total, Nilai IPK)} = a - \left(\frac{e}{d} * b\right) + \left(\frac{f}{d} * c\right)$$

$$\begin{aligned} Gain \text{ (Total, Nilai IPK)} &= 0.418746429 + 0 \\ &= 0.418746429 \end{aligned}$$

Telah diketahui bahwa nilai *gain* dari nilai IPK yaitu 0.418746429.

2. Mencari Nilai *Gain* Nilai TOEFL

Diketahui :

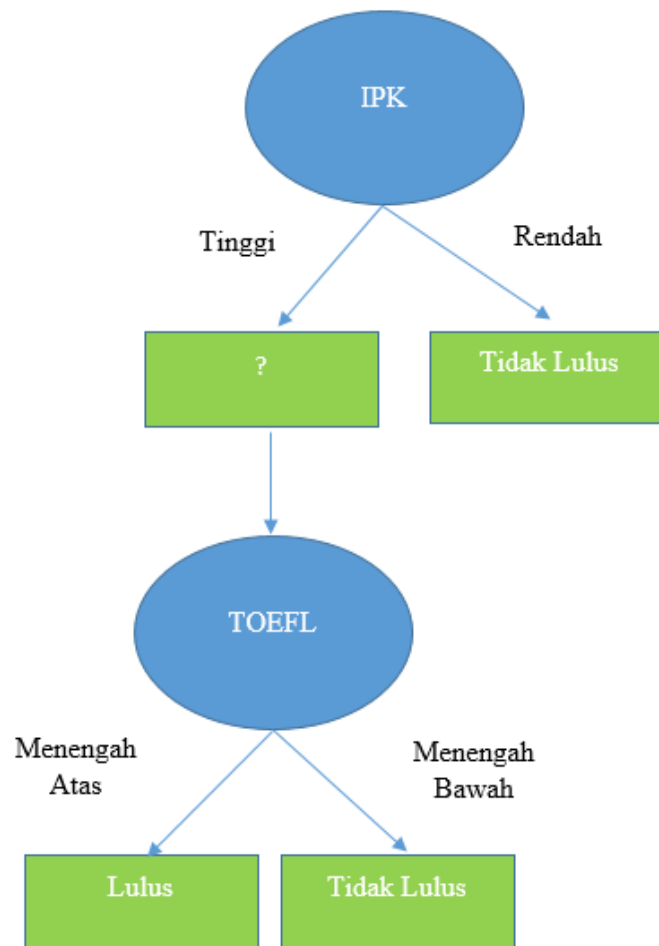
<i>Entropi</i> kasus keseluruhan (a)	= 0.714727
<i>Entropi</i> nilai TOEFL menengah atas (b)	= 0.468996
<i>Entropi</i> nilai TOEFL menengah bawah (c)	= 0
Jumlah Kasus Keseluruhan (d)	= 56
Jumlah Kasus Nilai TOEFL menengah atas (e)	= 50
Jumlah Kasus Nilai TOEFL menengah bawah (f)	= 6

$$Gain \text{ (Total, Nilai TOEFL)} = a - \left(\frac{e}{d} * b\right) + \left(\frac{f}{d} * c\right)$$

$$\begin{aligned} Gain \text{ (Total, Nilai TOEFL)} &= 0.418746429 + 0 \\ &= 0.418746429 \end{aligned}$$

Telah diketahui bahwa nilai *gain* dari Nilai TOEFL yaitu 0.418746429.

Setelah nilai *entropi* dan nilai *gain* dari masing-masing atribut telah diketahui, kemudian melakukan pembentukan *decision tree* dimana nilai gain tertinggi yaitu Nilai IPK akan dijadikan sebagai root/node nya sehingga akan menghasilkan suatu *rule* dari hasil perhitungan dan pembentukan *decision tree* dari data mahasiswa yang digunakan. Untuk pembentukan pohon keputusan dapat dilihat pada gambar 4.6.



Gambar 4.6 pembentukan *decision tree*

Pada pembentukan pohon keputusan, menghasilkan suatu *rule* seperti berikut :

if Nilai IPK = Rendah

then = Tidak Lulus

if Nilai IPK tinggi = Aktif *and*

Nilai TOEFL = Menengah atas

then = Lulus

if Nilai IPK = Tinggi *and*

Nilai TOEFL = Menengah bawah

then = Tidak Lulus

7. Pengukuran Kinerja Pohon Keputusan

Pengukuran kinerja pohon keputusan digunakan untuk mengevaluasi pohon keputusan yang telah dibuat. Pengukuran kinerja pohon keputusan dilakukan dengan menggunakan *confusion matrix*. *Confusion Matrix* didapatkan dengan melakukan perbandingan hasil prediksi dan hasil sebenarnya. Untuk mengetahui kinerja pohon keputusan akan digunakan *precision*, *recall*, dan akurasi. Proses pengukuran kinerja akan dilakukan sebanyak 5 kali terhadap data uji. Kemudian nilai *precision*, *recall*, dan akurasi diubah menjadi nilai *presentasi* seperti pada tabel berikut.

Tabel 4.13 Pengukuran Kinerja Pohon Keputusan

Pengukuran Ke-	Data Uji	Nilai <i>Confidence</i>
1	56	0,20
2	52	0,15
3	56	0,20
4	56	0,20
5	56	0,20

Confusion Matrix Pengukuran ke-1

Tabel 4.14 Pengukuran confusion matrix

Prediksi Lulus	Prediksi Tidak Lulus
<i>True Positive</i> : 45	<i>False Negative</i> : 11
<i>False Positive</i> : 11	<i>True Negative</i> : 45

Hasil Pengukuran Kinerja

Tabel 4.15 Hasil Pengukuran Kinerja Pohon Keputusan

Pengukuran Ke	<i>Precision</i>	<i>Recall</i>	Akurasi
1	65%	61.23%	70%
2	61.5%	20.50%	58%
3	65%	61.23%	70%
4	62.30%	59.20%	60%
5	61.5%	20.50%	65%

Hasil rata-rata pengukuran kinerja

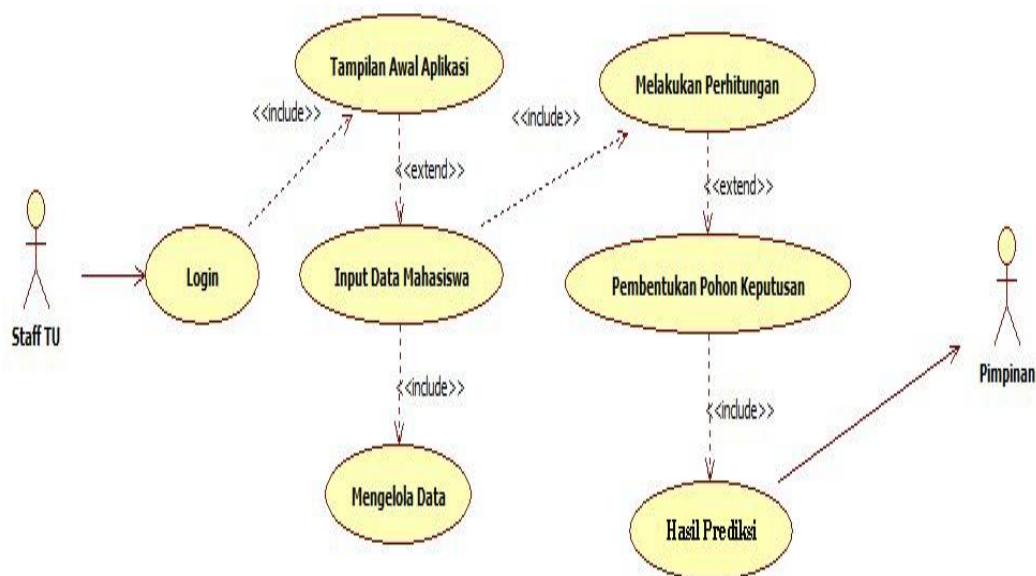
Tabel 4.16 Hasil rata-rata kinerja pohon keputusan

Nilai <i>Confidence</i>	<i>Precision</i>	<i>Recall</i>	Akurasi
0.20	65%	61%	70%

Data yang digunakan dalam pengukuran kinerja adalah sebanyak 56 data mahasiswa dan dilakukan percobaan sebanyak 5 kali dari data yang telah di *transformasi*. Untuk melakukan kinerja pohon keputusan, langkah pertama yaitu hitung *true positive*, *false positive*, dan *false negative* dari setiap pengukuran, kemudian masukan hasilnya kedalam *confusion matrix*. Langkah kedua, hitung

precision, *recall* dan Akurasi. Akurasi didefinisikan sebagai tingkat kedekatan antara nilai hasil prediksi dengan nilai aktual. *Precision* didefinisikan sebagai pengukuran ketepatan. Jika data diprediksi positif, seberapa seringkah data prediksi tersebut benar. Sedangkan *recall* didefinisikan sebagai pengukuran kelengkapan. Dari jumlah data sebenarnya yang bernilai positif, sebanyak apakah data yang diprediksi positif. Kemudian ubah nilai *precision*, *recall*, dan akurasi menjadi nilai persentase, Kemudian di rata-rata kan berdasarkan pohon keputusan.

4.2.2 Use Case Diagram



Gambar 4.7 Use Diagram

Diagram Use Case di atas menunjukkan Actor yang terlibat di dalam Aplikasi perediksi kelulusan ada 2 Actor, yang masing-masing mempunyai hak sebagai berikut :

1. Definisi Aktor

a. Staf TU

Staf Tu adalah yang menggunakan aplikasi, dimulai mengelola data sampai melakukan input data pada aplikasi untuk dilakukan perhitungan dimana hasilnya akan diterima oleh pimpinan.

b. Pimpinan

Pimpinan menerima laporan akhir dalam penentuan kelulusan yang diprediksi menggunakan algoritma c4.5 yang diimplementasikan ke dalam aplikasi

2. *Use Case*

Use case menggambarkan kegiatan yang akan terjadi antara aktor dan aplikasi sistem. Diagram *use case* digunakan untuk memahami fungsi- apa saja yang ada dalam sebuah sistem yang akan digambarkan pada tabel berikut ini :

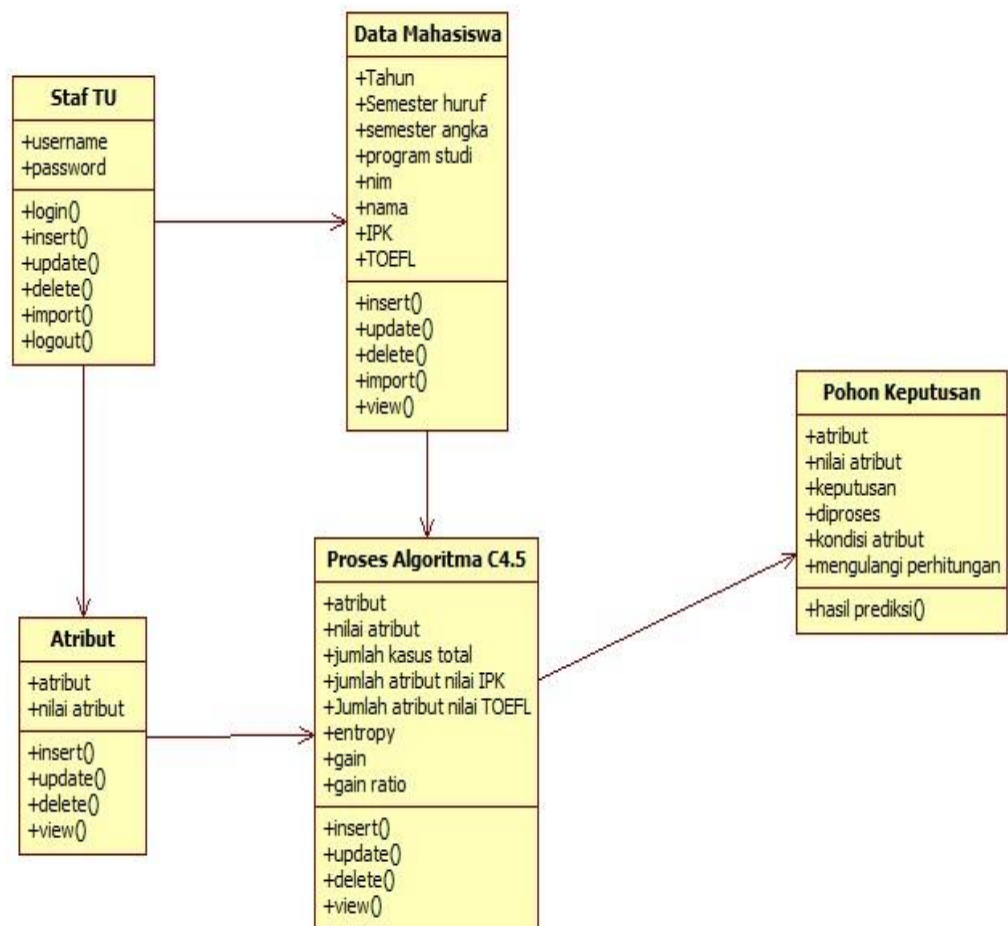
Tabel 4.17 Use case aplikasi

NO	Aktor	<i>Use Case</i>	Deskripsi
1	Staf TU	<i>Login</i>	Melakukan login pada aplikasi prediksi kelulusan mahasiswa
2		Tampilan Awal Aplikasi	Tampilan yang pertama kali muncul saat <i>user</i> melakukan <i>login</i>
3		<i>Input Data</i> Mahasiswa	Digunakan untuk memasukan data mahasiswa kedalam aplikasi untuk diproses sebagai hasil prediksi kelulusan
4		Mengelola Data	Digunakan untuk menambah data, <i>delete</i> data, <i>import</i> data dan <i>update</i> data
5		Melakukan Perhitungan	Aplikasi akan melakukan perhitungan jika data mahasiswa telah dimasukan kedalam aplikasi
6		Pembentukan pohon keputusan	Untuk melakukan pembentukan pohon keputusan jika nilai

			<i>entropy</i> dan nilai <i>gain</i> telah ditemukan
7	Pimpinan	Hasil Perhitungan	Memunculkan hasil perhitungan jika semua tahap telah dilakukan

4.2.3 Class Diagram

Class Diagram menggambarkan kelas-kelas dalam sebuah sistem dan hubungannya antara satu dengan yang lain, serta dimasukan pula atribut dan operasi. *Class diagram* (diagram kelas) ini menggambarkan struktur sistem aplikasi yang diusulkan dari segi pendefinisian kelas-kelas yang akan dibuat untuk membuat aplikasi prediksi kelulusan mahasiswa.



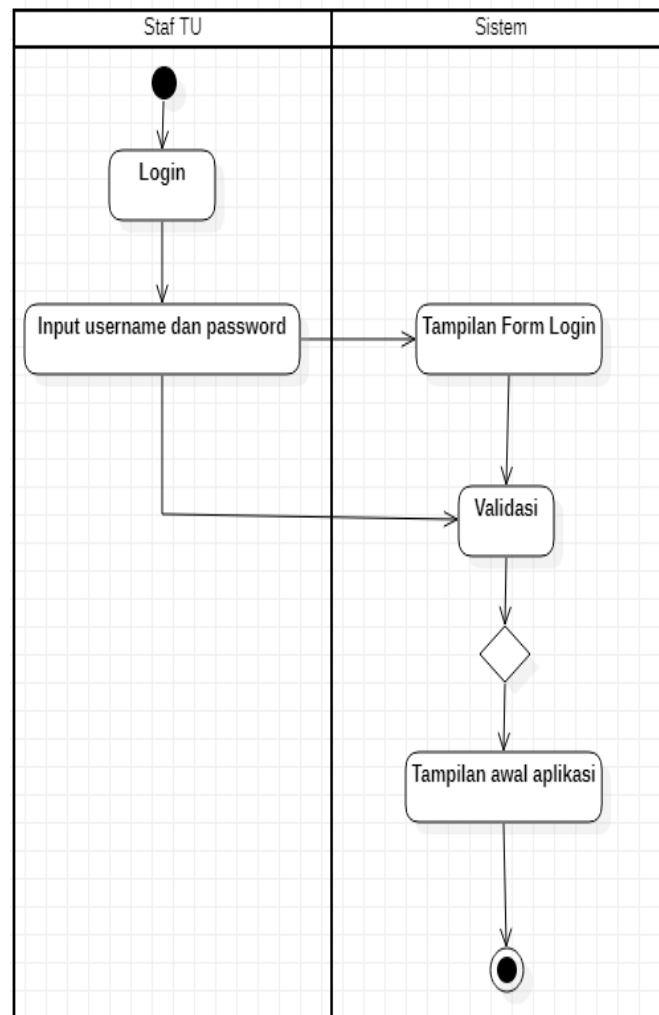
Gambar 4.8 Class Diagram Aplikasi

4.2.4 Activity Diagram

Activity Diagram menggambarkan alur penggunaan aplikasi pada tiap-tiap aplikasi. *Activity diagram* merupakan bentuk *visual* dari alir kerja yang berisi aktivitas dan tindakan, yang juga dapat berisi pilihan, pengulangan, dan *concurrency*. Dalam *unified modeling language*, diagram aktivitas dibuat untuk menjelaskan aktivitas komputer maupun alur aktivitas dalam organisasi.

Diagram aktivitas memiliki komponen dengan bentuk tertentu, dihubungkan dengan tanda panah. Panah tersebut mengarahkan urutan aktivitas yang terjadi, dari awal sampai akhir. Yang perlu diperhatikan yaitu diagram aktivitas bukan menggambarkan aktivitas sistem yang dilakukan aktor, tetapi menggambarkan aktivitas yang dapat dilakukan oleh sistem.

a. *Activity Diagram Login*



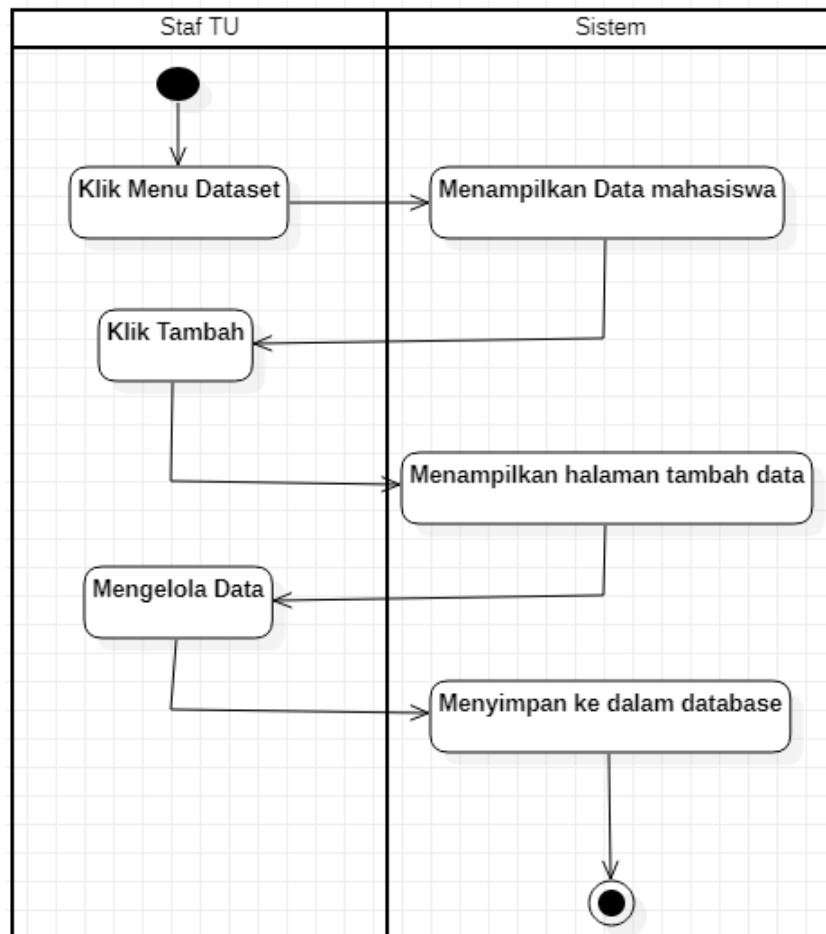
Gambar 4.9 Activity Diagram Login

Pada Gambar 4.9 diatas Staf TU melakukan *login* dengan mengakses menu *login* kemudian masukan *username* dan *password* lalu sistem akan mem-*verifikasi* jika *username* dan *password* salah maka sistem akan memberikan informasi tidak *valid* jika *username* dan *password* benar. maka sistem akan menampilkan tampilan awal.

b. Activity Diagram Kelola Data

Pada gambar 4.10 dijelaskan bahwa Staf TU masuk ke form Dataset dengan melakukan login terlebih dahulu. Lalu Staf TU dapat melihat dan mengelola semua data yang tersedia pada form tersebut seperti memasukan data ataupun

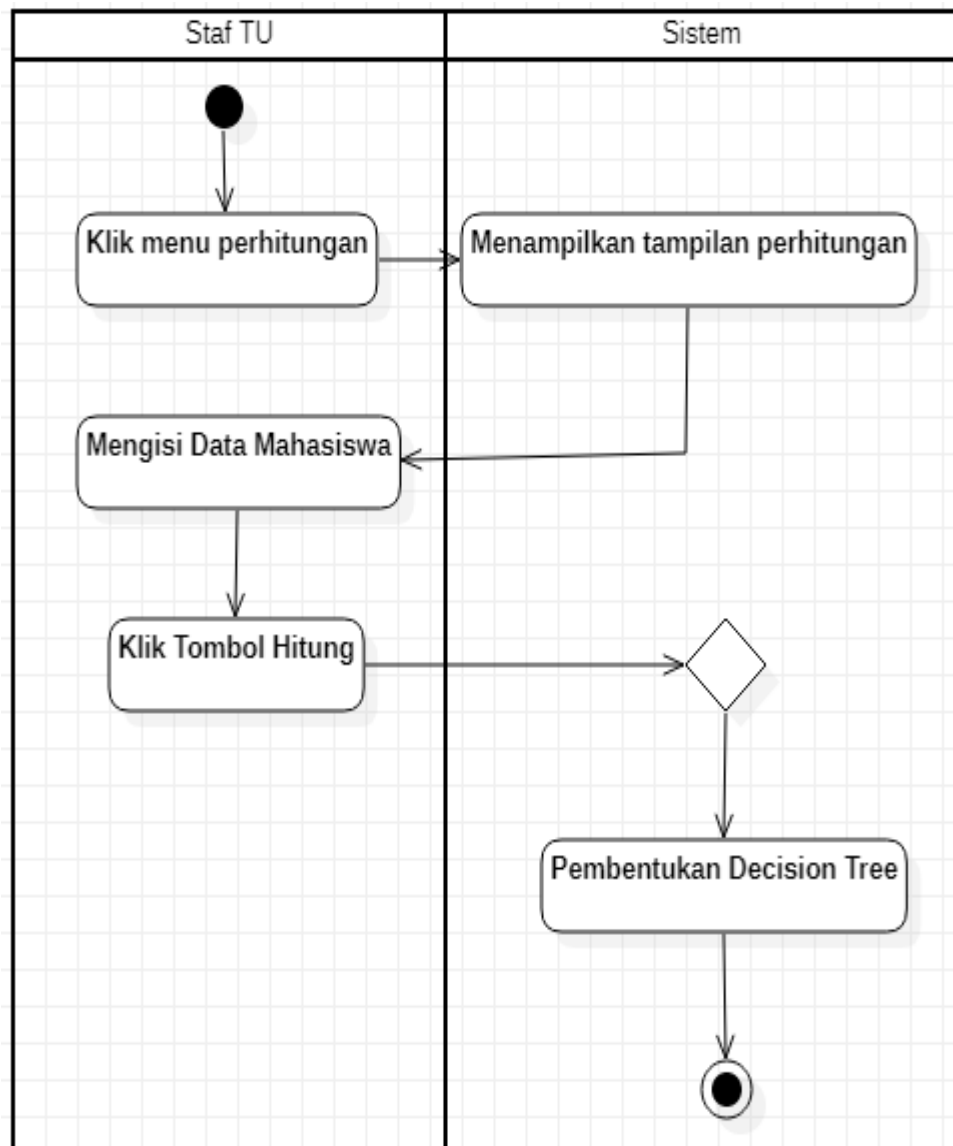
meng-input-kan data dari *microsoft excel* dengan *extensi .xls*. User pun dapat *mengupdate* dan *delete* data.



Gambar 4.10 Activity Diagram Kelola Data

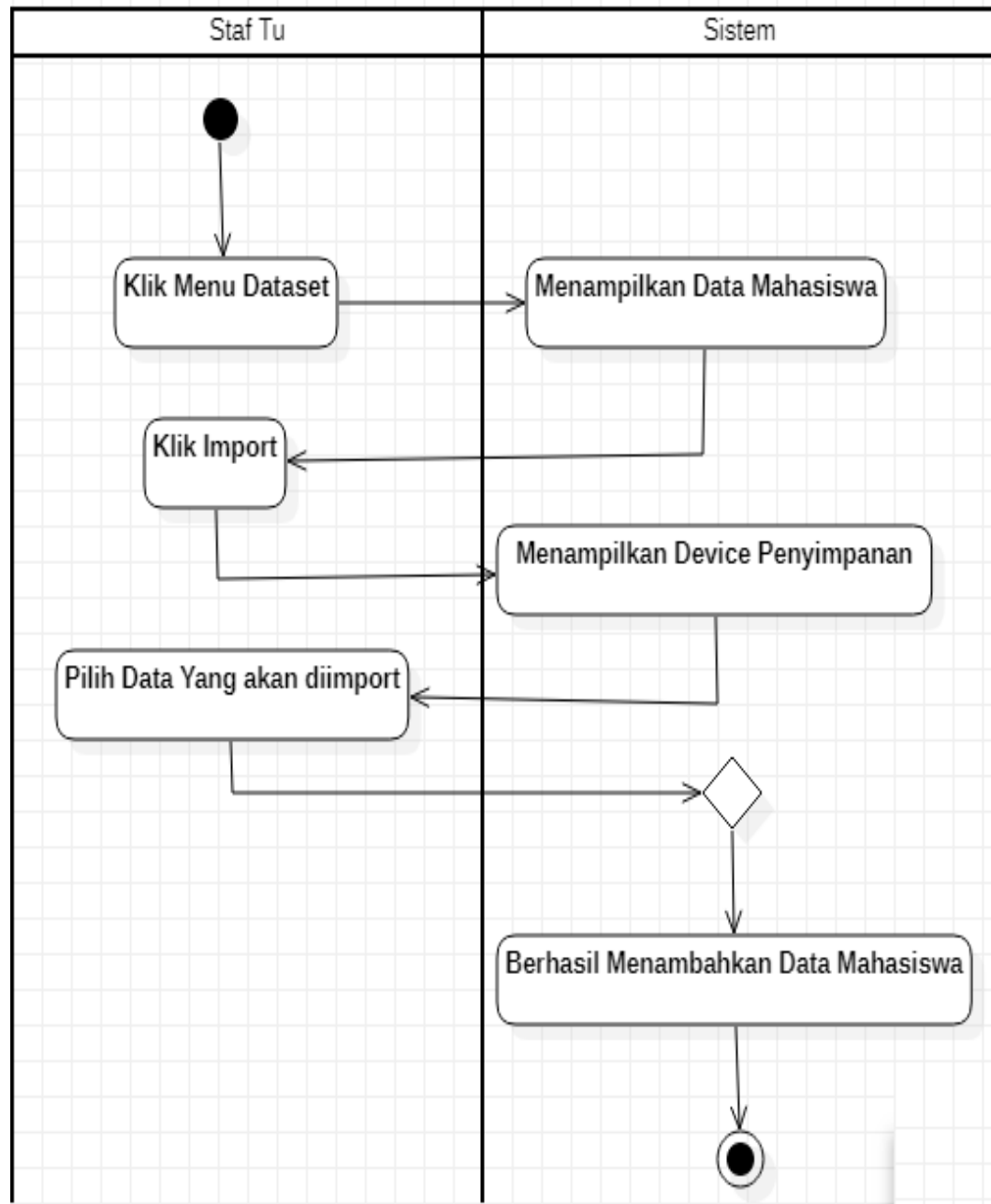
c. Activity Diagram perhitungan Algoritma c4.5

Pada gambar 4.11 dijelaskan bahwa setelah melakukan *login* pada aplikasi Staf TU masu ke menu perhitungan setelah itu aplikasi akan menampilkan halaman perhitungan lalu Staf TU mulai memasukan data mahasiswa setelah data mahasiswa telah dimasukan lalu klik tombol hitung sehingga aplikasi memulai melakukan proses perhitngan.



Gambar 4.11 Activity Diagram Proses Perhitungan

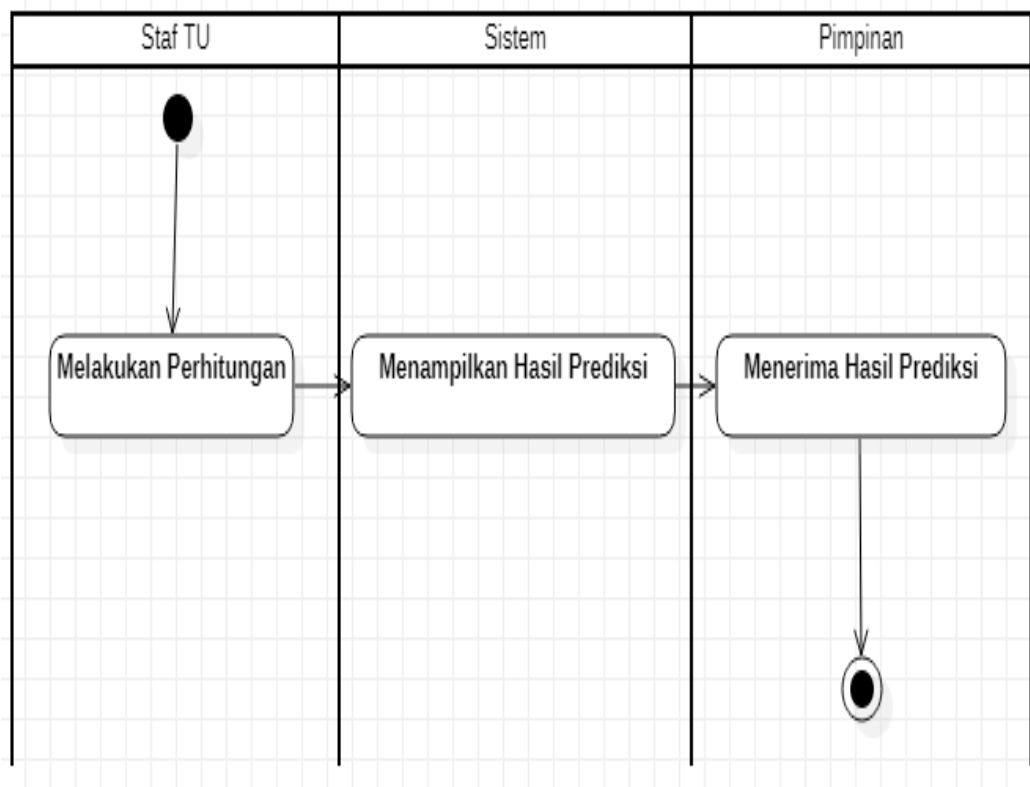
d. Activity Diagram Import Data



Gambar 4.12 Activity diagram import data

Pada Gambar 4.12 dijelaskan bahwa setelah Staf TU melakukan *login* pada aplikasi, Staf TU memilih form dataset dimana form tersebut akan menampilkan data mahasiswa yang telah dimasukan jika ingin meng*import* data baru, Staf TU memilih tombol *import* pada aplikasi sehingga aplikasi akan membuka *device* penyimpanan komputer lalu pilih data mahasiswa yang akan di-*import* lalu sistem akan memproses dan data berhasil dimasukan.

e. *Activity Diagram Pimpinan*



Gambar 4.13 Activity Diagram Logout

Pada Gambar 4.13 dapat dijelaskan Staf TU mulai melakukan proses perhitungan dan sistem akan menampilkan hasil prediksi dan pimpinan akan menerima hasil prediksi.

4.3 Perancangan *User Interface*

Perancangan *user interface* merupakan mekanisme komunikasi antara pengguna (*user*) dengan sistem. Antar muka pemakai (*user interface*) dapat menerima informasi dari pengguna (*user*) dan memberikan informasi kepada pengguna (*user*) untuk membantu mengarahkan alur penelusuran masalah sampai ditemukan suatu solusi.

a. Perancangan Form Login

Form *login* ini berfungsi untuk menjaga keamanan aplikasi agar tidak sembarang orang bisa mengakses aplikasi ini. Untuk mengakses nya- dibutuhkan

username dan *password*. Form login ini bertujuan untuk membatasi akses pengguna aplikasi.

Gambar 4.14 Perancangan form login

Keterangan :

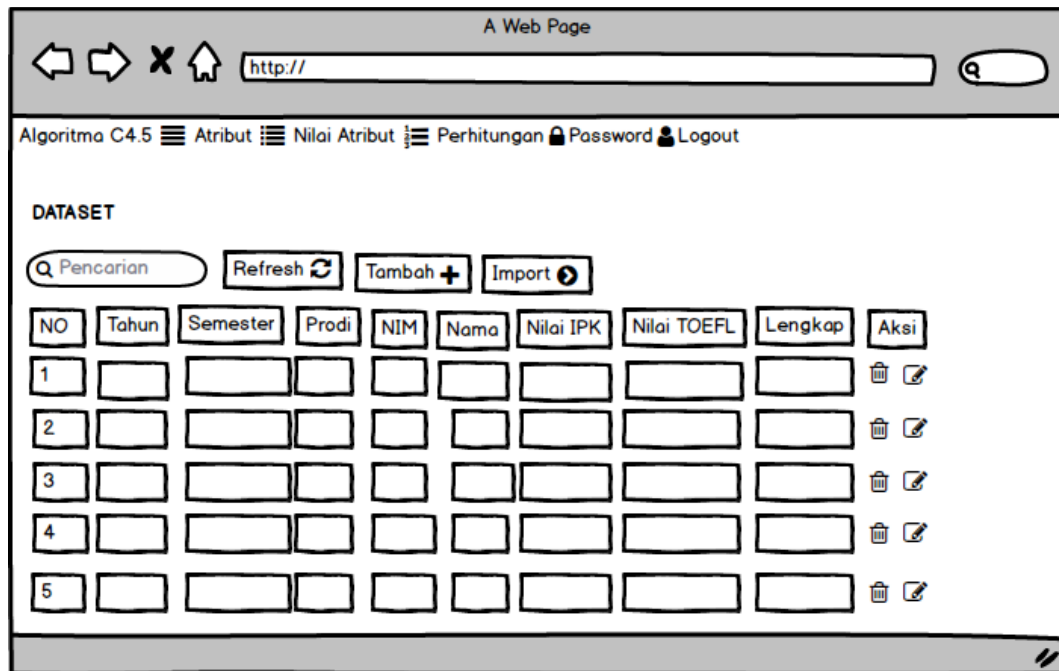
Tabel 4.18 Keterangan form login

<i>Username</i>	Di gunakan untuk mengisi <i>user name user</i>
<i>Password</i>	Di gunakan untuk mengisi <i>Password user</i>

a. Tampilan Dataset mahasiswa

Tampilan ini akan muncul setelah *user* melakukan *login* pada aplikasi dan akan menampilkan data mahasiswa yang telah dimasukan sebelumnya

dan jika sebelumnya belum pernah memasukan data mahasiswa, aplikasi hanya akan menampilkan beberapa form seperti pencarian sampai form *import*.



Gambar 4.15 Perancangan halaman user interface Dataset

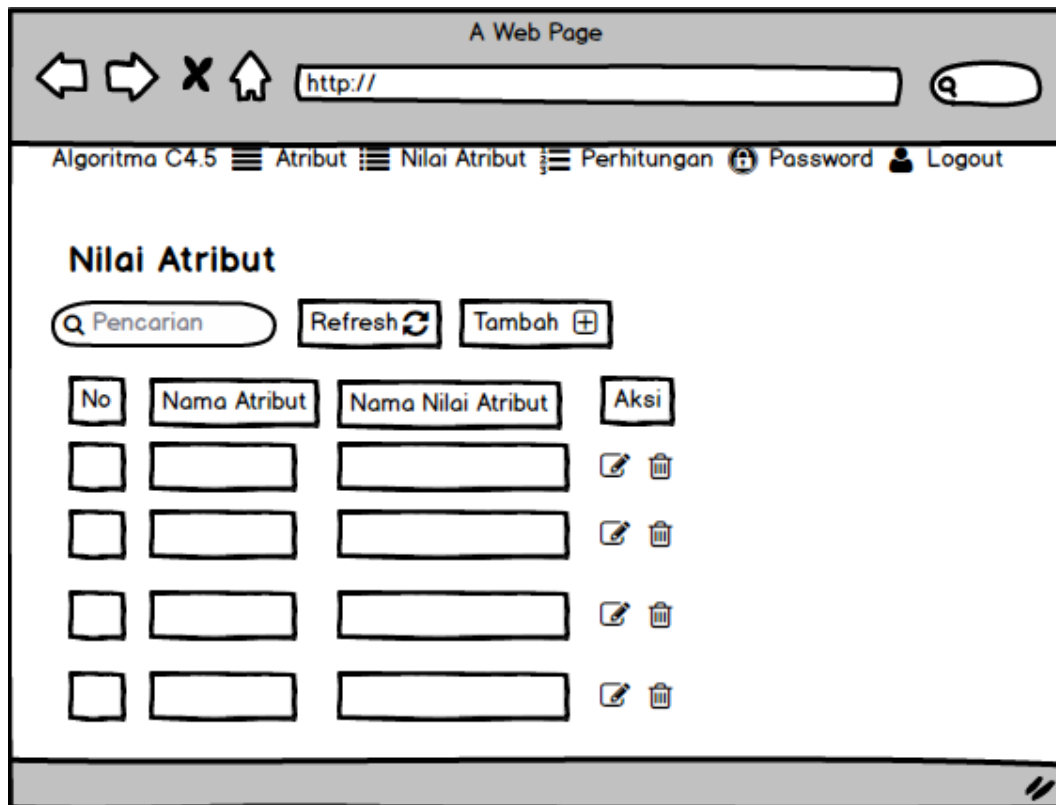
Keterangan :

Tabel 4.19 keterangan halaman dataset

Pencarian	Digunakan untuk mencari data mahasiswa yang akan dicari.
<i>Refresh</i>	Digunakan untuk memuat ulang data mahasiswa
Tambah	Untuk melakukan penambahan data mahasiswa
<i>Import</i>	Untuk melakukan <i>import</i> data mahasiswa pada aplikasi lain

c. Tampilan halaman nilai atribut

Tampilan halaman nilai atribut ini akan menampilkan atribut yang sangat berpengaruh dalam prediksi kelulusan mahasiswa.



Gambar 4.16 perancangan halaman nilai atribut

Keterangan :

Tabel 4.20 Keterangan halaman nilai atribut

Pencarian	Melakukan pencarian atribut
Refresh	Memuat ulang halaman
Tambah	Melakukan penambahan atribut jika dibutuhkan
Aksi	Digunakan menyunting nama atribut dan menghapus nama atribut jika tidak diperlukan

d. Halaman perhitungan

Halaman perhitungan digunakan untuk menguji data mahasiswa yang akan diproses dimana data mahasiswa tersebut akan diambil atribut yang paling

berpengaruh dalam prediksi kelulusan mahasiswa yaitu Nilai IPK dan Nilai TOEFL.

A Web Page

Algoritma C4.5 Atribut Nilai Atribut Perhitungan Password Logout

Perhitungan

Data yang diketahui


Tahun

Semester

NIM

Nilai IPK

Nilai TOEFL

 Hitung

Gambar 4.17 Perancangan halaman perhitungan

Keterangan :

Tabel 4.21 Halaman Perhitungan

Tahun	Mengisi tahun mahasiswa
Semester	Mengisi semester mahasiswa
NIM	Mengisi NIM mahasiswa
Nilai IPK	Untuk mengisi nilai IPK mahasiswa
Nilai TOEFL	Untuk mengisi nilai TOEFL mahasiswa

Hitung	Untuk melakukan proses perhitungan dan jika selesai menekan tombol hitung, aplikasi akan mulai melakukan pembentukan pohon keputusan.
--------	---

4.4 Perancangan Database

Perancangan database adalah proses untuk menemukan isi dan pengaturan data yang di butuhkan untuk mendukung berbagai rencana aplikasi yang akan di bangun pada aplikasi prediksi kelulusan mahasiswa. Dapat di lihat pada tabel-tabel berikut ini :

A. Mahasiswa

Tujuan di buatnya *form input* data mahasiswa adalah untuk meng-*input*-kan data ke dalam *database*, dan rancangan *database* nya dapat di lihat pada tabel di bawah ini :

Tabel 4.22 Perancangan Database

No	Field	Type	Panjang	Keterangan
1	Nama_mahasiswa	Varchar	30	Menyimpan nama mahasiswa
2	Tahun	INT	4	Menyimpan tahun mahasiswa yang sedang melakukan study
3	Semester_angka	INT	2	Memasukan semester mahasiswa yang sedang melakukan study menggunakan angka
4	Semester_huruf	Varchar	7	Memasukan semester mahasiswa yang sedang melakukan study menggunakan huruf
5	Nim_mahasiswa	Primary Key	13	Menyimpan NIM mahasiswa

6		<i>Varchar</i>	10	Menyimpan semester mahasiswa
7	Nilai_IPK	<i>INT</i>	10	Menyimpan Nilai IPK mahasiswa
8	Nilai_TOEFL	<i>INT</i>	10	Menyimpan Nilai TOEFL mahasiswa
9	Lengkap	<i>INT</i>	1	Menyimpan nilai kelengkapan mahasiswa

BAB V

IMPLEMENTASI DAN PENGUJIAN

5.1 Implementasi

Implementasi ini dilakukan sesuai dengan perancangan yang dilakukan pada bab sebelumnya. Terdiri dari pembentukan pohon keputusan dimana nilai tiap cabang yang telah ditemukan pada bab sebelumnya.

5.1.1 Implementasi Perangkat Keras dan Perangkat Lunak

Perangkat keras yang penyusun gunakan dalam tahap perhitungan dan pembentukan pohon keputusan tidak harus yang berspesifikasi tinggi. Adapun spesifikasi perangkat keras dan perangkat lunak yang digunakan adalah sebagai berikut :

a. Spesifikasi Perangkat Keras

Tabel 5.23 Spesifikasi Perangkat Keras

Nama Perangkat	Spesifikasi
<i>Processor</i>	Intel pentium dual CPU E2160 @ 1,8 Ghz
RAM	2 x 2 GB DDR 2
<i>Hard disk</i>	232.88 GB

b. Spesifikasi Perangkat Lunak

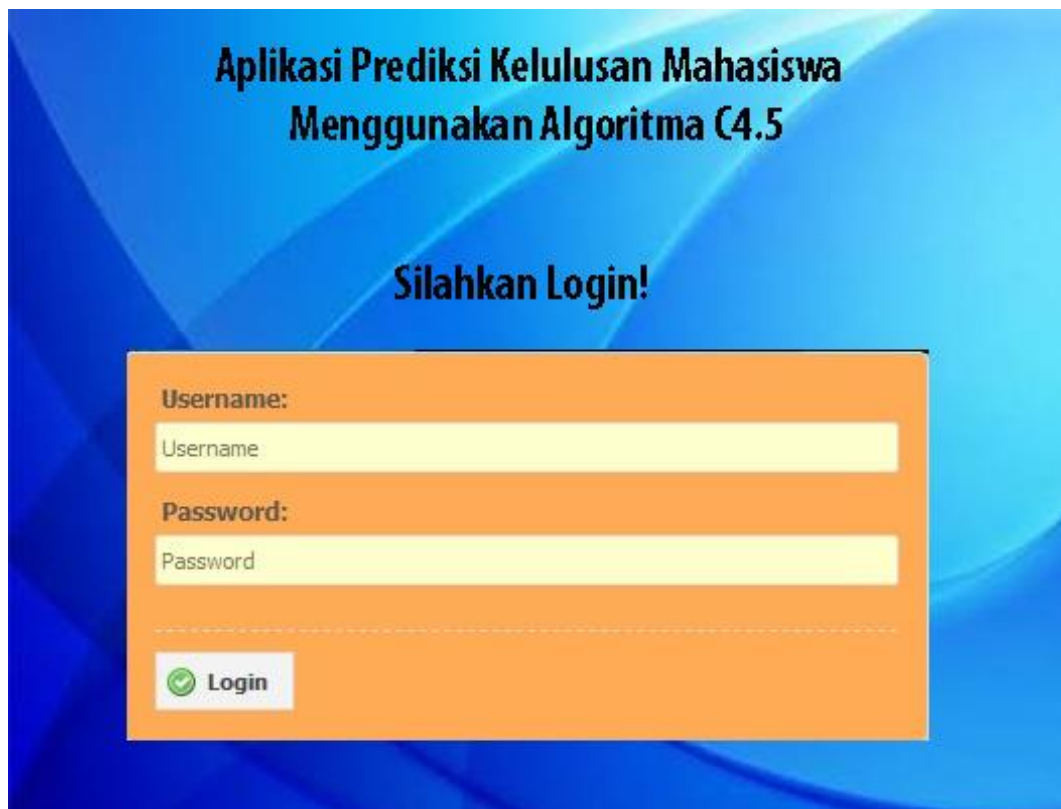
Spesifikasi perangkat lunak yang digunakan untuk menjalankan aplikasi Prediksi ini adalah XAMPP *Cotrol panel* V 3.2.2 dan *Windows* 7 64 bit.

5.2 Pengujian Sistem

Berdasarkan perancangan yang dilakukan telah diketahui bahwa terdapat beberapa tampilan *form* pada aplikasi yang bertujuan untuk mempermudah *user* dalam pengoperasian-nya yang terdiri dari *Login*, halaman dataset, halaman atribut data, halaman nilai atribut data mahasiswa dan halaman perhitungan data mahasiswa. Berikut ini adalah tahapan pengujian dari berbagai *form*.

1. Tampilan halaman login

Menu *login* pada sistem ini berguna untuk validasi data pengguna. Sebelum masuk ke menu utama, pengguna harus menginputkan *username* dan *password*. Setelah mengklik tombol masuk, sistem mengecek *database* dengan data *login* yang diinputkan oleh pengguna, termasuk level hak akses pengguna dalam menggunakan sistem. Jika data yang diinputkan benar, akan masuk ke tampilan menu utama. Tampilan menu *login* dapat dilihat pada gambar 5.18 di bawah ini.



Gambar 5.18 Tampilan halaman login

Halaman *login* di buat untuk mengamankan data user sehingga tidak sembarang orang yang dapat mengaksesnya, terdapat kotak *username* untuk mengisi *username user*, kotak *password* untuk mengisi *password user* dan tombol *login* untuk masuk ke dalam sistem.

Pengujian :

Tabel 5.24 Pengujian halaman login

No	Tujuan	Input	Output diharapkan	Output sistem
1	Masuk kedalam menu <i>dashboard</i> aplikasi	<i>Username & password</i>	<i>Login</i> dan aplikasi menampilkan <i>dashboard</i>	*Sukses * <i>Dashboard</i> tampil sesuai dengan <i>output</i>

2. Tampilan Halaman Dataset Mahasiswa

No	Tahun	Semester	Prodi	NIM	Nama	IPK	TOEFL	Lengkap
1	2018	8 Genap	IF	C1A150001	B	3.0	400	1
2	2018	8 Genap	IF	C1A150002	C	3.0	420	1
3	2018	8 Genap	IF	C1A150003	C	3.0	400	1
4	2018	8 Genap	IF	C1A150004	B	2.9	400	0
5	2018	8 Genap	IF	C1A150005	C	3.2	350	0
6	2018	8 Genap	IF	C1A150006	B	3.0	400	1
7	2018	8 Genap	IF	C1A150007	C	3.0	420	1
8	2018	8 Genap	IF	C1A150008	A	3.0	410	1
9	2018	8 Genap	IF	C1A150009	B	3.0	400	1
10	2018	8 Genap	IF	C1A150010	B	3.0	420	1
11	2018	8 Genap	IF	C1A150011	C	3.0	400	1
12	2018	8 Genap	IF	C1A150012	B	3.2	420	1
13	2018	8 Genap	IF	C1A150013	C	3.0	400	1
14	2018	8 Genap	IF	C1A150014	C	3.0	350	0

Gambar 5.19 Tampilan Halaman Dataset Mahasiswa

Tampilan halaman dataset merupakan halaman awal aplikasi yang akan muncul setelah *user* berhasil melakukan *login* terlebih dahulu, dan sistem akan menampilkan data mahasiswa yang sebelumnya telah dimasukan, data tersebut berupa tahun, semester, prodi (Program Studi), NIM (Nomor Induk Mahasiswa), Nilai IPK dan Nilai TOEFL.

Pengujian :

Tabel 5.25 Pengujian Halaman Dataset

No	Tujuan	Input	Output yang diharapkan	Output Sistem
1	Pencarian	Klik untuk mencari data mahasiswa	Menampilkan data mahasiswa yang dicari oleh <i>user</i>	Sistem berhasil mencari data sesuai dengan yang dicari <i>user</i>
2	<i>Refresh</i>	Klik untuk memuat ulang data jika <i>user</i> memasukan data mahasiswa yang baru	Sistem akan memuat ulang halaman dan akan menambahkan data mahasiswa baru jika <i>user</i> memasukannya	Sistem berhasil memuat ulang halaman dan akan menambahkan data mahasiswa baru jika <i>user</i> memasukannya
3	Tambah	Klik untuk menambahkan data mahasiswa yang baru	Sistem akan menampilkan halaman <i>input</i> data.	Sistem berhasil menampilkan halaman <i>input</i> data

4	<i>import</i>	Klik untuk mengimport data baru	Sistem akan membaca <i>device</i> penyimpanan komputer dan berhasil mengimportkan data.	Sistem berhasil membaca <i>device</i> penyimpanan komputer dan berhasil mengimportkan data.
---	---------------	---------------------------------	---	---

3. Tampilan Halaman Nilai Atribut Data Mahasiswa

Algoritma C45 Atribut Nilai Atribut Dataset Perhitungan Password Logout				
Nilai Atribut				
Pencarian... Refresh + Tambah				
No	Nama Atribut	Nama Nilai Atribut		
1	Nilai IPK	Tinggi		
2	Nilai TOEFL	Rendah		
3	Nilai IPK	Menengah Atas		
4	Nilai TOEFL	Menengah Bawah		

Gambar 5.20 Tampilan Halaman Nilai Atribut

Tampilan halaman nilai atribut ini merupakan atribut yang sangat berpengaruh dalam menentukan kelulusan mahasiswa Fakultas Teknologi Informasi Universitas Bale Bandung.

Pengujian :

Tabel 5.26 Pengujian Halaman Nilai Atribut

No	Input	Output
1	Pencarian	Menampilkan nilai atribut yang ingin dicari
2	<i>Referseh</i>	Sistem memuat ulang halaman
3	Tambah	Untuk menambahkan data nilai atribut

4. Tampilan Halaman Perhitungan

Gambar 5.21 Tampilan halaman perhitungan

Tampilan halaman perhitungan ini merupakan tampilan untuk melakukan perhitungan data mahasiswa sehingga akan menghasilkan sebuah pohon keputusan.

Pengujian :

Tabel 5.27 Pengujian Halaman Perhitungan

NO	Input	Keterangan
1	Tahun	Memasukan tahun mahasiswa saat melakukan study
2	Semester	Memasukan semester mahasiswa saat melakukan study
3	NIM	Memasukan nim mahasiswa
4	Nilai IPK	Memasukan Nilai IPK mahasiswa
5	Nilai TOEFL	Memasukan nilai TOEFL mahasiswa
6	Hitung	Mulai melakukan perhitungan jika semua data mahasiswa telah dimasukan.

BAB VI

KESIMPULAN DAN SARAN

6.1 Kesimpulan

Berdasarkan penelitian yang di lakukan penulis melalui beberapa tahapan yang di lakukan pada bab-bab sebelumnya maka penulis dapat menyimpulkan :

1. Dalam prediksi kelulusan mahasiswa menggunakan algoritma c4.5 pada data mahasiswa Fakultas Teknologi Informasi Universitas Bale Bandung, nilai IPK lebih dominan dalam prediksi kelulusan Mahasiswa dibandingkan dengan Nilai TOEFL.
2. Hasil dari kinerja pohon keputusan menghasilkan nilai rata-rata *precision* 65%, *recall* 61% dan Akurasi 70%, sehingga nilai *confidence* 0,20 menghasilkan tingkat akurasi yang lebih baik dibandingkan dengan nilai *condifence* 0,15.

6.2 Saran

Dalam proses pembuatan Aplikasi *prediksi kelulusan mahasiswa*, penulis masih banyak sekali memiliki kekurangan dan jauh dari kata sempurna, sehingga penulis berharap untuk peneliti selanjutnya dapat mengembangkan lagi aplikasi ini, diantaranya :

1. Aplikasi ini masih memiliki banyak kekurangan dari segi *interface* dan fungsionalitasnya, maka alangkah lebih baik jika dapat di kembangkan lagi sehingga dapat lebih memudahkan user dalam pengoperasian aplikasi ini.
2. Aplikasi ini dapat di tambahkan fitur ataupun performansinya, selain itu dapat dikembangkan menjadi berbasis mobile dengan bahasa pemrograman lain.

3. Untuk lebih meningkatkan akurasi dari hasil prediksi bisa menggunakan Algoritma *Pruning* yaitu algoritma pemangkasan dimana pohon keputusan yang telah dibentuk akan dipotong/dipangkas.

DAFTAR PUSTAKA

- Nuqson Masykur Huda. (2010). Aplikasi Data Mining Untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa.
- Mujib Ridwan, Hadi Suyono, M.Sarosa. (2013). PENERAPAN DATA MINING UNTUK EVALUASI KINERJA AKADEMIK MAHASISWA MENGGUNAKAN ALGORITMA NAIVE BAYES CLASSIFIER.
- Agus Romadhona, Suprapedi, H.Himawan. (2017). Prediksi kelulusan mahasiswa tepat waktu berdasarkan usia, jenis kelamin, dan indeks prestasi menggunakan algoritma decesion tree.
- Anik Andriani. (2013). SISTEM KEPUTUSAN BERBASIS DECISION TREE DALAM PEMBERIAN BEASISWA STUDI KASUS: AMIK BSI YOGYAKARTA. *Semiar nasional teknologi informasi dan komunikasi 2013 (SENTIKA 2013)*.
- eko prasetyo. (2014). *DATA MINING-Mengolah data menjadi informasi menggunakan matlab*. Yogyakarta: ANDI.
- Herdianto. (2013). Prediksi Kerusakan Motor Induksi Menggunakan Metode Jaringan Syaraf Tiruan Backpropagation. 8-9.
- Indah Puji Astuti. (2017). Prediksi Ketepatan Waktu Kelulusan Dengan Algoritma Data Mining C4.5.
- Isa Iskandar,Leli Hiryanto,Janson Hendryli. (n.d.). Judul : Prediksi Kelulusan Mahasiswa Menggunakan Algoritma Decision Tree C4.5 Dengan Teknik Pruning.
- Nanda Dimas Prayoga. (2017). Penerapan Algoritma C4.5 Dalam Memprediksi Kelulusan Tepat Waktu Pada Perguruan Tinggi (Studi Kasus : STMIK Royal Kisaran).
- Sarwono. (1978). *Pengertian mahasiswa menurut para ahli beserta peranannya lengkap*. Retrieved from <https://www.studineews.co.id/2016/12/pengertian-mahasiswa-menurut-para-ahli-beserta-peranannya.html>
- Studineews. (2016). *Pengertian mahasiswa menurut para ahli beserta peranannya lengkap*. Retrieved from www.studineews.co.id:https://www.studineews.co.id/pengertian-mahasiswa-menurut-para-ahli-beserta-peranannya.html
- Wenefrida Tulit Ina. (2013). Klasifikasi Data Rekam Medis Berdasarkan Kode Penyakit Internasional Menggunakan Algoritma C4.5.
- Wijaya. (2009). Model Driven Development

Lampiran 1 Data Mahasiswa

NO	TAHUN	SEM	SEM2	PRODI	NIM	NAMA	IPK	TOEFL	LENGKAP	KETLULUS
1	2018	GENAP		8 IF	C1A150001	A	3,0	400	1	LULUS
2	2018	GENAP		8 IF	C1A150002	B	3,0	420	1	LULUS
3	2018	GENAP		8 IF	C1A150003	A	3,0	400	1	LULUS
4	2018	GENAP		8 IF	C1A150004	B	2,9	400	0	TIDAK LULUS
5	2018	GENAP		8 IF	C1A150005	C	3,2	350	0	TIDAK LULUS
6	2018	GENAP		8 IF	C1A150006	A	3,0	400	1	LULUS
7	2018	GENAP		8 IF	C1A150007	C	3,0	420	1	LULUS
8	2018	GENAP		8 IF	C1A150008	A	3,0	410	1	LULUS
9	2018	GENAP		8 IF	C1A150009	B	3,0	400	1	LULUS
10	2018	GENAP		8 IF	C1A150010	B	3,0	420	1	LULUS
11	2018	GENAP		8 IF	C1A150011	C	3,0	400	1	LULUS
12	2018	GENAP		8 IF	C1A150012	B	3,2	420	1	LULUS
13	2018	GENAP		8 IF	C1A150013	C	3,0	400	1	LULUS
14	2018	GENAP		8 IF	C1A150014	C	3,0	400	1	LULUS
15	2018	GENAP		8 IF	C1A150015	B	2,9	400	0	TIDAK LULUS
16	2018	GENAP		8 IF	C1A150016	C	3,0	350	0	TIDAK LULUS
17	2018	GENAP		8 IF	C1A150017	B	2,8	400	1	LULUS
18	2018	GENAP		8 IF	C1A150018	C	3,0	420	1	LULUS
19	2018	GENAP		8 IF	C1A150019	A	3,0	430	1	LULUS
20	2018	GENAP		8 IF	C1A150020	B	3,2	350	0	TIDAK LULUS
21	2018	GENAP		8 IF	C1A150021	C	3,3	400	1	LULUS
22	2018	GENAP		8 IF	C1A150022	A	3,0	410	1	LULUS
23	2018	GENAP		8 IF	C1A150023	B	2,9	400	1	LULUS
24	2018	GENAP		8 IF	C1A150024	B	3,3	350	0	TIDAK LULUS
25	2018	GENAP		8 IF	C1A150025	C	3,2	430	1	LULUS
26	2018	GENAP		8 IF	C1A150026	B	3,3	400	1	LULUS
27	2018	GENAP		8 IF	C1A150027	C	3,0	380	0	TIDAK LULUS
28	2018	GENAP		8 IF	C1A150028	C	3,0	400	1	LULUS
29	2018	GENAP		8 IF	C1A150029	B	3,1	350	0	TIDAK LULUS
30	2018	GENAP		8 IF	C1A150030	C	3,0	400	1	LULUS

31	2018	GENAP	8 IF	C1A150031B	3,2	410	0 TIDAK LULUS
32	2018	GENAP	8 IF	C1A150032C	3,3	400	1 LULUS
33	2018	GENAP	8 IF	C1A150033A	3,0	430	1 LULUS
34	2018	GENAP	8 IF	C1A150034B	3,2	420	1 LULUS
35	2018	GENAP	8 IF	C1A150035C	3,2	410	1 LULUS
36	2018	GENAP	8 IF	C1A150036A	3,2	430	1 LULUS
37	2018	GENAP	8 IF	C1A150037B	3,0	350	0 TIDAK LULUS
38	2018	GENAP	8 IF	C1A150038B	3,0	400	1 LULUS
39	2018	GENAP	8 IF	C1A150039C	3,0	400	1 LULUS
40	2018	GENAP	8 IF	C1A150040B	3,1	420	1 LULUS
41	2018	GENAP	8 IF	C1A150041C	3,2	410	1 LULUS
42	2018	GENAP	8 IF	C1A150042C	3,0	410	1 LULUS
43	2018	GENAP	8 IF	C1A150043B	3,0	400	1 LULUS
44	2018	GENAP	8 IF	C1A150044C	3,3	410	1 LULUS
45	2018	GENAP	8 IF	C1A150045B	3,3	350	0 TIDAK LULUS
46	2018	GENAP	8 IF	C1A150046C	3,0	400	1 LULUS
47	2018	GENAP	8 IF	C1A150047A	3,2	410	1 LULUS
48	2018	GENAP	8 IF	C1A150048B	3,3	430	1 LULUS
49	2018	GENAP	8 IF	C1A150049C	3,3	420	1 LULUS
50	2018	GENAP	8 IF	C1A150050A	3,2	400	1 LULUS
51	2018	GENAP	8 IF	C1A150051B	3,0	400	1 LULUS
52	2018	GENAP	8 IF	C1A150052A	3,0	410	1 LULUS
53	2018	GENAP	8 SI	C1B15001 A	3,1	400	1 LULUS
54	2018	GENAP	8 SI	C1B15002 B	3,2	430	1 LULUS
55	2018	GENAP	8 SI	C1B15003 B	3,0	430	1 LULUS
56	2018	GENAP	8 SI	C1B15004 C	3,3	420	1 LULUS

Lampiran 2 Source Code Aplikasi

Implementasi Algoritma C4.5

```
CREATE DATABASE IF NOT EXISTS db_dm;
DROP TABLE IF EXISTS c45_attributes;
CREATE TABLE IF NOT EXISTS c45_attributes(
    id INT AUTO_INCREMENT PRIMARY KEY,
    name VARCHAR(50),
    target INT DEFAULT '0'
);
DROP TABLE IF EXISTS c45_instances;
CREATE TABLE IF NOT EXISTS c45_instances(
    id INT AUTO_INCREMENT PRIMARY KEY,
    id_attribute INT NOT NULL,
    name VARCHAR(50)
);

DROP TABLE IF EXISTS c45_samples;
CREATE TABLE IF NOT EXISTS c45_samples(
    id INT AUTO_INCREMENT PRIMARY KEY,
    name VARCHAR(50)
);

DROP TABLE IF EXISTS c45_sample_details;
CREATE TABLE IF NOT EXISTS c45_sample_details(
    id INT AUTO_INCREMENT PRIMARY KEY,
    id_sample INT NOT NULL,
    id_instance INT NOT NULL
);

DROP TABLE IF EXISTS c45_decision_tree;
CREATE TABLE IF NOT EXISTS c45_decision_tree(
    id INT AUTO_INCREMENT PRIMARY KEY,
    id_attribute INT NOT NULL,
    id_instance INT NOT NULL,
    id_target INT DEFAULT '0'
);

$dbhost = 'localhost';
$dbuser = 'root';
$dbpass = '';
$dbname = 'db_dm';
//-- koneksi ke database server dengan extension mysqli
$db = new mysqli($dbhost,$dbuser,$dbpass,$dbname);
//--
- hentikan program dan tampilkan pesan kesalahan jika koneksi ga
gal
if ($db->connect_error) {
    die('Connect Error ('.$db->connect_errno.')'.'$db-
>connect_error');
}
```

```
funksi Entropy
function entropy($S){
    $entropy=0;
    foreach($S as $s){
        $p= $s/array_sum($S);
        $entropy+=(-$p)*log($p);
    }
    return $entropy;
}
//-- funksi Gain
function gain($S,$Si){
    $gain=entropy($S);
    foreach($Si as $si){
        $gain-=( $si/array_sum($Si))*entropy($Si);
    }
    return $gain;
}
```

Lampiran 3 Biodata Penulis



Muhamad Ghasa Saputra lahir di Bandung, pada 21 Maret 1997. Anak ketiga dari tiga bersaudara, yang terlahir dari pasangan Isnaeni dan Juhendi. Mulai mengenyang pendidikan di TK NURUL SYIFA (2003), kemudian dilanjutkan ke SDN 1 KERTAMULYA (2004-2009), kemudian melanjutkan pendidikan ke SMP KRIDA UTAMA PADALARANG (2009-2012), serta melanjutkan masa SMA di SMA PASUNDAN MAJALAYA (2012-2015). Untuk mendapatkan gelar sarjana penulis melanjutkan S1 penulis melanjutkan pendidikan di Universitas Bale Bandung Fakultas Teknologi Informasi Jurusan Teknik Informatika. Penulis juga pernah aktif sebagai anggota di UKM Bale Percussion di FTI UNIBBA pada tahun 2015-2016, selain itu penulis pernah aktif sebagai anggota HIMA-IF sampai tahun 2018.