**BAB II**

**TINJAUAN PUSTAKA**

1. **Tinjauan Studi**

Berikut adalah beberapa penelitian yang terkait dengan algoritma Naïve Bayes, secara garis besar tinjauan studi untuk penelitian ini meliputi :

1. Penelitian ini yang dilakukan oleh (Naisha Rahma Indraswari, 2018) yang berjudul Penggunaan Naïve Bayes untuk memprediksi Usia kelahiran umumnya kelahiran bayi sehat cukup bulan berada pada minggu 38-42 kehamiam. Namun ada banyak bayi yang terahir pada sian kelahiran yang kurang mencukupi bahkan lahir dalam usia kelahiran yang lewat waktu. Hal ini menjadi hal yang serius mengingat banyak terjadi kematian bayi akibat usia kelahiran yang kurang mencukupi atau lewat waktu. Penelitian ini bertujuan untuk membuat aplikasi prediksi yang nantinya akan membantu pasien dalam mengetahui usia kelahiranya dan mengantisipasi hal yang tidak diinginkan kedepanya. Metode yang digunakan merupakan metode Naïve Bayes dengan variabel inputan factor-faktor yang dialami oleh ibu hamil, diantanya: usia ibu, tekanan dara, jumlah bayi, riwayat persalinan, riwayat kehamilan. Hasil dari penelitian ini merupakan sebuah aplikasi yang dapat memprediksi usia kelahiran dengan nilai akurasi aplikasi tertinggi pada angka 78,69%, nilai *precision* tertinggi pada angka 70,14% dan nilai *recall* tetinggi pada angka 63,64%.
2. Penelitian ini yang di lakukan oleh (Supardi salmu, 2017). Yang berjudul Prediksi Tingkat Kelulusan Mahasiswa Tepat Waktu Menggunakan Naïve Bayes (Studi Kasus UIN Syarif Hidayattullah Jakarta). Proses belajar mengajar hingga proses wisuda di Universitas Islam Negeri (UIN) Syarif Hidayatullah Jakarta telah didukng oleh system informasi yang bernama Academic Information System (AIS). AIS mampu memberikan data dan informasi yang dibutukan terkait kegiatan akademik mahasiswa. Melalui system ini ditemukan bahwa jumlah mahasiswa baru lebi banyak dibandingkan mahasiswa yang lulus. Namun AIS belum dapat menghasikan pengetahuan (knowledge) yang lebi mendalam terkait kondisi ini. Salasatu cara untuk menghasilkan mutu tertinggi dari system perguruan tinggi adalah dengan menggali pengetahuan dari data bidang pendidikan. Sehingga diperlukan adanya pemanfaatan teknik data mining unttuk menggali pengetahuan dari data mahasiswa yang tersedia di AIS. Penelitian ini bertujuan untuk menemukan faktor-faktor yang mempengaruhi pediksi kelulusan mahasiswa melalui data kinerja akademik mahasiswa pada semester satu sampai dengan empat. Model yang digunakan yaitu Cross industry Standart Process for data mining (CRISP-DM) dengan mengimplementasikan algoritma Naïve Bayes untuk klasifikasi data. Melalui tahap data understanding didapatkan 12 atribut predictor yang akan dianalisis terhadap satu atrib 80,72%, dapat disimpulkan bahwa model ini dapat di jadikan acuan dalam memprediksi kelulsan mahasiswa. Pengetahuan ini dapat dimanfaatkan oleh pihak UIN sebagai langka preventif untuk menghindari penurunan kelulusan mahasiswa setiap tahunnya.
3. Penelitian ini yang dilakukan oleh (Momad Guntur, 2018) yang berjudul Prediksi Harga Emas dengan Menggunakan Metode Naïve Bayes dalam Ivestasi untuk Meminimalisasi Resiko fluktuasi harga emas dipengaruhi oleh banyak factor seperti kondisi perekonomian, laju inflansi, penawaran dan pemintaan serta masi banyak lagi. Algoritma Naïve Bayes mampu menghasilkan sebua klasifikasi yang digunakan untuk memprediksi peluang dimasa depan. Dengan menggunakan algoitma Naïve Bayes Classifer didapat sebuah prediksi harga emas yang bisa membantu pemanggil keputusan dalam menentukan apakah harus menjual atau membeli emas. Data emas yang diolah bersumber dari hasil observasi di website [www.pegadaian.co.id](http://www.pegadaian.co.id) dari tanggal 1 Desember 2017- 1 Januari 2018. Data emas diolah menggunakan software rapidminer. Tahapan pengolahanya adalah membaca data training, menghitung nilai mean dan standar deviasi, memasukan data uji dan mencari nilai desintas gauss lalu mencari nilai probailitas. Berdasakan hasil perhitungan yang tela dilakukan, metode Naïve Bayes Classifer mampu digunakan untuk memprediksi harga emas untuk 14 hari kedepan, data yang digunakan untuk untuk pengujian sebanyak 16 data dan diperoleh tingkat akurasi sebesar 75%, dengan hasil ini diharapkan dapat membantu para investor untuk mengambil keputusan dengan baik.

**2.2. Tinjauan Teori**

**2.2.1 Ujian Nasional**

Ujian Nasional adalah penilaian pada akhir proses pembelajaran disekolah (Depdiknas, 2003). Ujian nasional seringkali membawa efek psikologis bagi yang cukup berat bagi sebagian besar siswa ditingkat akhir, karena UN dianggap sebagai penentuan keberhasilan atau kegagalan siswa selama tiga atau enam tahun menempu masa pendidikannya. Jika mengalami kegagalan dalam menempu UN tentu saja akan menghambat proses kelulusan pada siswa tersebut, sehingga dibutukan sebua upaya untuk meminalkan kegagalan dalam Ujian Nasional

Sala satu cara untuk mengurangi kegalalan siswa dapat dilakukan suatu prediksi kelulusan Ujian Nasioanal berdasarkan nilai ujian nasional siswa kedalam sebuah system. Penggunaan metode *Naif Bayes* diharapkan dapat membantu memprediksi probailitas kelulusan Ujian Nasional yang akan didapatkan oleh masing-masing siswa, khususnya siswa Sekolah Menengah Atas. Sehingga pihak sekolah bias menyusun strategi yang tepat dalam meningkatkan kuantitas kelulusan Ujian Nasional siswa.

**2.2.2Data Mining**

*Data Mining* (Witten, 2011) didefinisikan sebagai proses penemuan pola dalam data. Menurut Daryl Pregibons dalam (Gorunescu, 2011) “Data mining adalah perpaduan dari ilmu statistik, kecerdasan buatan, dan penelitian bidang *database*”. Nama data mining berasal dari kemiripan antara pencarian informasi yang bernilai dari database yang besar dengan menambang sebuah gunung untuk sesuatu yang bernilai (Sumathi, 2006). Keduanya memerlukan penyaringan melalui sejumlah besar material, atau menyelidiki dengan cerdas untuk mencari keberadaan sesuatu yang disebut bernilai tadi. Istilah lain dari data (Han, 2006) yaitu *knowledge mining from databases*, *knowledge extraction, data/pattern analysis*, *data archeology*, dan data *dredging*. Banyak yang menggunakan data mining sebagai istilah popular dari KDD.

Data mining (Maimon, 2005) merupakan inti dari proses *Knowledge Discovery in Database* (KDD). KDD adalah proses terorganisir untuk mengidentifikasi pola yang valid, baru, berguna, dan dapat dimengerti dari sebuah data set yang besar dan kompleks.

**2.2.2.1 Proses Tahapan Data Mining**

Istilah data mining dan knowledge discovery databases (KDD) sering digunakan secara bergantian untuk menjelaskan proses penggalian informasi dalam database. Secara garis besar prosess KDD dapat dijelaskan sebagai berikut (Maimon, 2005):

1. Pembentukan pemahaman *domain* aplikasi

Pada tahap ini menentukan tujuan dari end-user dan bagian terkait dimana KDD dilakukan.

2. Memilih dan membuat data set dimana proses penemuan knowledge akan dilakukan.

Penentuan data yang akan digunakan untuk proses KDD dilakukan pada tahap ini. Mencari data yang tersedia, memperoleh data tambahan yang dibutuhkan, mengintegrasikan semua data untuk KDD ke dalam sebuah data set, termasuk atribut yang diperlukan dalam proses KDD.

3. *Preprocessing* dan *cleansing*

Dalam tahap ini kehandalan data ditingkatkan. Termasuk data clearing, seperti menangani data yang tidak lengkap, menghilangkan gangguan atau outlier.

4. Transformasi data

Pada tahap ini, membuat data menjadi lebih baik menggunakan metode reduksi dimensi dan transformasi atribut.

5. Memilih tugas data mining yang cocok

Pada tahap ini ditentukan tipe data mining yang akan digunakan, apakah klasifikasi, regresi, atau clustering, tergantung pada tujuan KDD dan tahap sebelumnya.

6. Memilih algoritma data mining

Pada tahap ini dilakukan pemilihan algoritma yang paling tepat untuk menemukan pola.

7. Penggunaan algoritma data mining

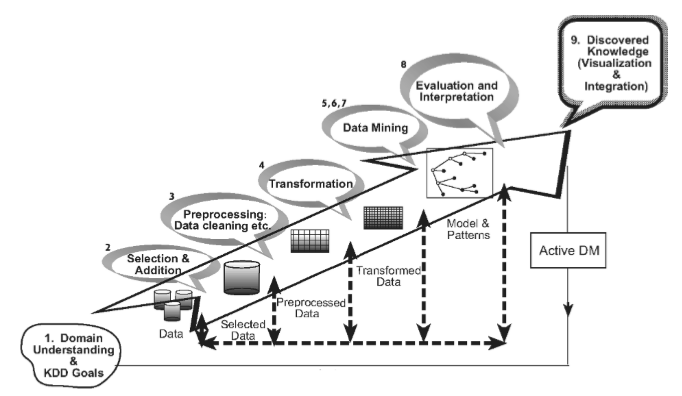
Pada tahap ini dilakukan implementasi dari algoritma data mining yang telah ditentukan pada tahap sebelumnya.

8. Evaluasi

Pada tahap ini dilakukan evaluasi dan penerjemahan dari pola yang didapat.

9. Penggunaan pengetahuan yang didapat

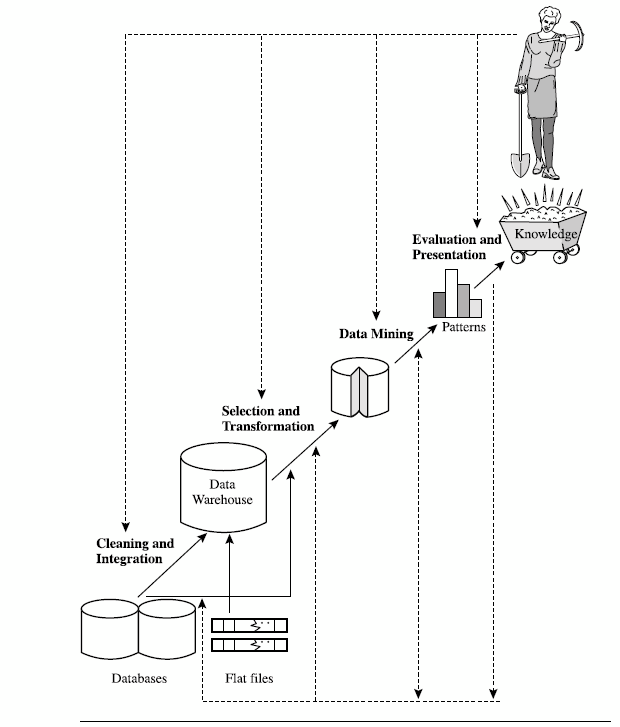
Pada tahap ini, pengetahuan dimasukkan ke sistem lain dan mengaktifkan sistem tersebut serta mengukur hasilnya. (Maimon, 2005)



(Maimon, 2005).

**Gambar 2.1.** Proses KDD

Sedangkan menurut (Han & Kamber, 2006dalam Andika Christian Dahlan, 2017) tahapan dalam KDD adalah sebagai berikut :



(Han & Kamber, 2006)

**Gambar 2.2.**  Tahap-Tahap Data Mining

1. Pembersihan data (data *cleaning*) Pembersihan data merupakan proses menghilangkan *noise* dan data yang tidak konsisten atau data tidak relevan. Pada umumnya data yang diperoleh, baik dari *database* suatu perusahaan maupun hasil eksperimen, memiliki isian-isian yang tidak sempurna seperti data yang hilang, data yang tidak valid atau juga hanya sekedar salah ketik. Selain itu, ada juga atribut-atribut data yang tidak relevan dengan hipotesa data mining yang dimiliki. Data-data yang tidak relevan itu juga lebih baik dibuang. Pembersihan data juga akan mempengaruhi performasi dari teknik data mining karena data yang ditangani akan berkurang jumlah dan kompleksitasnya.

2. Integrasi data (data *integration*) Integrasi data merupakan penggabungan data dari berbagai database ke dalam satu database baru. Tidak jarang data yang diperlukan untuk data mining tidak hanya berasal dari satu database tetapi juga berasal dari beberapa database atau file teks. Integrasi data dilakukan pada atribut-aribut yang mengidentifikasikan entitas-entitas yang unik seperti atribut nama, jenis produk, nomor pelanggan dan lainnya. Integrasi data perlu dilakukan secara cermat karena kesalahan pada integrasi data bisa menghasilkan hasil yang menyimpang dan bahkan menyesatkan pengambilan aksi nantinya. Sebagai contoh bila integrasi data berdasarkan jenis produk ternyata menggabungkan produk dari kategori yang berbeda maka akan didapatkan korelasi antar produk yang sebenarnya tidak ada.

3. Seleksi Data (Data *Selection*) Data yang ada pada database sering kali tidak semuanya dipakai, oleh karena itu hanya data yang sesuai untuk dianalisis yang akan diambil dari database. Sebagai contoh, sebuah kasus yang meneliti faktor kecenderungan orang membeli dalam kasus market basket analysis, tidak perlu mengambil nama pelanggan, cukup dengan id pelanggan saja.

4. Transformasi data (Data *Transformation*) Data diubah atau digabung ke dalam format yang sesuai untuk diproses dalam data mining. Beberapa metode data mining membutuhkan format data yang khusus sebelum bisa diaplikasikan. Sebagai contoh beberapa metode standar seperti analisis asosiasi dan clustering hanya bisa menerima input data kategorikal. Karenanya data berupa angka numerik yang berlanjut perlu dibagibagi menjadi beberapa interval. Proses ini sering disebut transformasi data.

5. Proses mining, Merupakan suatu proses utama saat metode diterapkan untuk menemukan pengetahuan berharga dan tersembunyi dari data.

6. Evaluasi pola (*pattern evaluation*), Untuk mengidentifikasi pola-pola menarik kedalam *knowledge based* yang ditemukan. Dalam tahap ini hasil dari teknik data mining berupa pola-pola yang khas maupun model prediksi dievaluasi untuk menilai apakah hipotesa yang ada memang tercapai. Bila ternyata hasil yang diperoleh tidak sesuai hipotesa ada beberapa alternatif yang dapat diambil seperti menjadikannya umpan balik untuk memperbaiki proses data mining, mencoba metode data mining lain yang lebih sesuai, atau menerima hasil ini sebagai suatu hasil yang di luar dugaan yang mungkin bermanfaat.

7. Presentasi pengetahuan (*knowledge presentation*), Merupakan visualisasi dan penyajian pengetahuan mengenai metode yang digunakan untuk memperoleh pengetahuan yang diperoleh pengguna. Tahap terakhir dari proses data mining adalah bagaimana memformulasikan keputusan atau aksi dari hasil analisis yang didapat. Ada kalanya hal ini harus melibatkan orang-orang yang tidak memahami data mining. Karenanya presentasi hasil data mining dalam bentuk pengetahuan yang bisa dipahami semua orang adalah satu tahapan yang diperlukan dalam proses data mining. Dalam presentasi ini, visualisasi juga bisa membantu mengkomunikasikan hasil data mining (Han, 2006).

**2.2.2.2Kelompok *Data Mining***

Berdasarkan tugasnya, data mining dikelompokkan menjadi (Larose, 2005):

1. Deskripsi

Mencari cara untuk menggambarkan pola dan trend yang terdapat dalam data. Sebagai contoh, seorang pengumpul suara mengungkap bukti bahwa mereka yang diberhentikan dari jabatannya saat ini, akan kurang mendukung dalam pemilihan presiden. Untuk deskripsi ini bisa dilakukan dengan *exploratory* data *analysis*, yaitu metode grafik untuk menelusuri data dalam mencari pola dan tren. (Larose, 2005)

2. Estimasi

Estimasi mirip seperti klasifikasi tapi variabel sasaran adalah numerik. Model dibuat menggunakan record yang lengkap, juga ada variabel targetnya. Kemudian untuk data baru, estimasi nilai variabel target dibuatberdasarkan nilai prediktor. Contoh, untuk estimasi tekanan darah pada pasien, variable prediktornya umur, jenis kelamin, berat badan, dan tingkat sodium darah. Hubungan antara tekanan darah, dan variable prediktor pada data training akan menghasilkan model kemudian diaplikasikan pada data baru. Untuk melakukan estimasi bisa digunakan *neural network* atau metode*statistic*seperti*point estimation*dan*conﬁdence interval estimations*, simple linear *regression* dan *correlation*, dan *multiple regression* (Larose, 2005).

3. Prediksi

Prediksi mirip seperti klasifikasi dan estimasi, tapi hasilnya untuk memprediksi masa depan. Contoh, memprediksi harga barang tiga bulan mendatang, memprediksi presentasi kenaikan angka kematian karena kecelakaan tahun mendatang jika kecepatan berkendara dinaikkan. Metode dan teknik untuk klasifikasi dan estimasi, jika cocok, bisa juga digunakan untuk prediksi, termasuk metode statistik. Algoritma untuk prediksi antara lain *regression tree* dan model *tree* (Han, 2006).

4. Klasifikasi

Dalam klasifikasi, sasarannya adalah variable kategori, misalkan atribut penghasilan, yang bisa dikategorikan menjadi tiga kelas atau kategori yaitu, tinggi, sedang, dan rendah. Model data mining membaca sejumlah besar record tiap record berisi informasi pada variable target. Contoh, dari sebuah data set misalkan mau mengklasifikasikan penghasilan seseorang yang datanya tidak terdapat pada dataset, berdasarkan karakteristik yang berhubungan dengan orang itu seperti, umur, jenis kelamin, dan pekerjaan. Tugas klasifikasi inicocok untuk metode dan teknik data mining. Algoritma akan mengolah dengan cara membaca data set yang berisi variable predictor dan variable taget yang telah diklasifikasikan, yaitu penghasilan. Di sini algoritma (*software*) “mempelajari” kombinasi variable mana yang berhubungan dengan penghasilan yang mana. Data ini disebut *training set*. Kemudian algoritma akan melihat ke data baru yang belum termasuk klasifikasi manapun. Berdasarkan klasifikasi pada data set, kemudian algoritma akan memasukkan data baru tersebut ke dalam klasifikasi yang mana. Misalkan seorang professor wanita berusia 63 tahun bisa jadi diklasifikasikan ke dalam kelas penghasilan tinggi. Algoritma klasifikasi yang banyak digunakan secara luas untuk klasifikasi antara lain, *decision tree*, *bayesian classifier*, dan *neural network* (Gorunescu, 2011)

5. *Clustering*

*Clustering* mengacu pada pengelompokkan record-record, observasi, atau kasus-kasus ke dalam kelas-kelas dari objek yang mirip. Pada *clustering*tidak ada variabel sasaran. Sebuah cluster adalah koleksi record yang mirip satu sama lain, dan tidak mirip dengan *record* pada *cluster*. Tidak seperti klasifikasi, pada *clustering* tidak ada variable target. *Clustering* tidak menglasifikasi atau mengestimasi atau memprediksi tetapi mencari untuk mensegmentasi seluruh data set ke subgroup yang *relative* sejenis atau *cluster*, dimana kemiripan record di dalam cluster dimaksimalkan dan kemiripan dengan record di luar cluster diminimalkan. Contoh *clustering*, untuk akunting dengan tujuan audit untuk mensegmentasi financial behaviour ke dalam kategori ramah dan curiga, sebagai alat reduksi dimensi ketika data set memiliki ratusan atribut, untuk clustering ekspresi gen, dimana kuantitas gen bisa terlihat mempunyai behavior yang mirip. Algoritma untuk *clustering* antara lain, *hierarchical agglomerative clustering, Bayesian clustering, self – organizing feature maps, growing hierarchical self - organizing maps* (Wu, 2009).

6. Asosiasi

Tugas asosiasi untuk *data mining* adalah kegiatan untuk mencari atribut yang “*go together*.” Dalam dunia bisnis, asosiasi dikenal sebagai affinity analysis atau market basket analysis, tugas asosiasi adalah membuka rules untuk pengukuran hubungan antara dua atribut atau lebih. Contoh asosiasi, prediksi degradasi dalam jaringan komunikasi, menemukan barang apa di supermarket yang dibeli bersama dengan barang lain yang tidak pernah dibeli bersama, menemukan proporsi kasus dimana obat baru akan memperlihatkan efek samping yang berbahaya. Untuk menemukanassociation rules, bisa dilakukan dengan algoritma a priori dan algoritmaGRI (*Generalized Rule Induction*) (Larose, 2005) .

**2.2.3Prediksi (*Forecasting*)**

*Forecasting* adalah prediksi (perkiraan) mengenai sesuatu yang belum terjadi. Forecasting atau peramalan adalah memperkirakan sesuatu pada waktu-waktu yang akan datang berdasarkan data masa lampau yang dianalisis secara ilmiah .Prediksi merupakan kegiatan untuk memperkirakan apa yang akan terjadi pada masa yang akan datang. Proses perkiraan pengukuran berdasarkan data yang relevan dengan masa lalu dan dianalisis secara ilmiah menggunakan metode statistika yang bertujuan untuk memperbaiki peristiwa yang akan terjadi pada masa yang akan datang. Dengan kata lain prediksi bertujuan untuk mendapatkan perkiraan yang bisa meminimumkan kesalahan prediksi (*forecast error*) yang biasanya diukur dengan *Standard Error Estimate* (SEE), *Mean Absolute PercentError* (MAPE) dan sebagainya.

**2.2.4** **Algoritma Naïve Bsyes**

Algoritma *Naive Bayes* merupakan salah satu algoritma yang terdapat pada teknik klasifikasi. *Naive Bayes* merupakan pengklasifikasian dengan metode probabilitas dan statistik yang dikemukan oleh ilmuwan Inggris *Thomas Bayes*, yaitu memprediksi peluang di masa depan berdasarkan pengalaman dimasa sebelumnya sehingga dikenal sebagai *Teorema Bayes*. Teorema tersebut dikombinasikan dengan *Naive* dimana diasumsikan kondisi antar atribut saling bebas. Klasifikasi *Naive Bayes* diasumsikan bahwa ada atau tidak ciri tertentu dari sebuah kelas tidak adahubungannya dengan ciri dari kelas lainnya.

Persamaan dari teorema *Bayes* adalah :

Keterangan :

X : Data dengan class yang belum diketahui

H : Hipotesis data merupakan suatu class spesifik

P(H|X) : Probabilitas hipotesis berdasar kondisi (posterioriprobability)

P(H) : Probabilitas hipotesis (prior probability)

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

P(X) : Probabilitas

Untuk menjelaskan teorema *Naive Bayes*, perlu diketahui bahwa prosesklasifikasi memerlukan sejumlah petunjuk untuk menentukan kelas apayang cocok bagi sampel yang dianalisis tersebut. Karena itu, teorema*bayes* di atas disesuaikan sebagai berikut :

Dimana Variabel *C* merepresentasikan kelas, sementara variabel *F1 ... Fn*merepresentasikan karakteristik petunjuk yang dibutuhkan untukmelakukan klasifikasi. Maka rumus tersebut menjelaskan bahwapeluang masuknya sampel karakteristik tertentu dalam kelas *C*(*Posterior*) adalah peluang munculnya kelas *C* (sebelum masuknyasampel tersebut, seringkali disebut *prior*), dikali dengan peluangkemunculan karakteristik karakteristik sampel pada kelas *C* (disebutjuga *likelihood*), dibagi dengan peluang kemunculan karakteristikkarakteristik sampel secara global ( disebut juga *evidence*). Karena itu,rumus diatas dapat pula ditulis secara sederhana sebagai berikut :

Nilai *Evidence* selalu tetap untuk setiap kelas pada satu sampel. Nilai dari *posterior* tersebut nantinya akan dibandingkan dengan nilai nilai*posterior* kelas lainnya untuk menentukan ke kelas apa suatu sampelakan diklasifikasikan.

Berikut contoh penerapan metode naïve bayes :

*Tabel 2. 1 Contoh Dataset Kelulusan Mahasiswa Pada Database*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **NO** | **Jenis**  **Kelamin** | **Status**  **Mahasiswa** | **Status**  **Pernikahan** | **IPK**  **Semester**  **1-6** | **Status**  **Kelulusan** |
| 1 | Laki-laki | Mahasiswa | Belum | 3.17 | Tepat |
| 2 | Laki-laki | Bekerja | Belum | 3.30 | Tepat |
| 3 | Perempuan | Mahasiswa | Belum | 3.01 | Tepat |
| 4 | Perempuan | Mahasiswa | Menikah | 3.25 | Tepat |
| 5 | Laki-laki | Bekerja | Menikah | 3.20 | Tepat |
| 6 | Laki-laki | Bekerja | Menikah | 2.50 | Terlambat |
| 7 | Perempuan | Bekerja | Menikah | 3.00 | Terlambat |
| 8 | Perempuan | Bekerja | Belum | 2.70 | Terlambat |
| 9 | Laki-laki | Bekerja | Belum | 2.40 | Terlambat |
| 10 | Perempuan | Mahasiswa | Menikah | 2.50 | Terlambat |
| 11 | Perempuan | Mahasiswa | Belum | 2.50 | Terlambat |
| 12 | Perempuan | Mahasiswa | Belum | 3.50 | Tepat |
| 13 | Laki-laki | Bekerja | Menikah | 3.30 | Tepat |
| 14 | Laki-laki | Mahasiswa | Menikah | 3.25 | Tepat |
| 15 | Laki-laki | Mahasiswa | Belum | 2.30 | Terlambat |

Jika seorang mahasiswa memiliki data uji sebagai berikut :

*Tabel 2. 2 Data Uji*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Jenis**  **Kelamin** | **Status**  **Mahasiswa** | **Status**  **Pernikahan** | **IPK**  **Semester**  **1-6** | **Status**  **Kelulusan** |
| Laki-laki | Mahasiswa | Belum | 2.70 | …? |

Dari data uji diatas untuk mencari klasifikasi kelulusan mahasiswa tersebut,

dilakukan dalam beberapa tahap :

1. Hitung jumlah class/ table

P(Y = Tepat) = 8/15

jumlah data “TEPAT” pada kolom “STATUS KELULUSAN” dibagi

jumlah data.

P(Y = Terlambat) = 7/15

jumlah data “TERLAMBAT” pada kolom “STATUS KELULUSAN”

dibagi jumlah data.

2. Hitung jumlah kasus yang sama dengan class yang sama

P(Jenis Kelamin = Laki-laki | Y= Tepat) = 5/8

jumlah data jenis kelamin “Laki-laki” dengan keterangan “TEPAT” dibagi

jumlah data “TEPAT”.

II-12

P(Jenis Kelamin = Laki-laki | Y= Terlambat) = 3/7

jumlah data jenis kelamin “Laki-laki” dengan keterangan

“TERLAMBAT” dibagi jumlah data “TERLAMBAT”.

P(Status Mahasiswa = Mahasiswa | Y= Tepat) = 5/8

jumlah data status mahasiswa “Mahasiswa” dengan keterangan “TEPAT”

dibagi jumlah data “TEPAT”.

P(Status Mahasiswa = Mahasiswa | Y= Terlambat) = 3/7

jumlah data status mahasiswa “Mahasiswa” dengan keterangan

“TERLAMBAT” dibagi jumlah data “TERLAMBAT”.

P(Status Pernikahan = Belum | Y= Tepat) = 4/8

jumlah data status pernikahan “Belum” dengan keterangan “TEPAT”

dibagi jumlah data “TEPAT”.

P(Status Pernikahan = Belum | Y= Terlambat) = 4/7

jumlah data status pernikahan “Belum” dengan keterangan

“TERLAMBAT” dibagi jumlah data “TERLAMBAT”.

P(IPK = 2.70 | Y= Tepat) = 0/8

jumlah data IPK “2.70” dengan keterangan “TEPAT” dibagi jumlah data

“TEPAT”.

P(IPK = 2.70 | Y= Terlambat) = 1/7

jumlah data status pernikahan “Belum” dengan keterangan

“TERLAMBAT” dibagi jumlah data “TERLAMBAT”.

3. Kalikan semua hasil variable TEPAT dan TERLAMBAT

P(Kelamin = Laki-laki), (Status mahasiswa = Mahasiswa), (Status

pernikahan = belum), (IPK = 2.70)| TEPAT)

=5/8 . 5/8 . 4/8 . 0/8 . 8/15

= 0

P(Kelamin = Laki-laki), (Status mahasiswa = Mahasiswa), (Status

pernikahan = belum), (IPK = 2.70)| TERLAMBAT)

= 3/7 . 3/7 . 4/7 . 1/7 . 7/15

= 0,0069

II-13

Dari hasil penghitungan diatas, berdasarkan data latih dapat ditemukan bahwa

(P|TERLAMBAT) lebih besar daripada (P|TEPAT) maka klasifikasi untuk data uji diatas adalah “TERLAMBAT”

**2.2.4.1** **Analisis Hasil Akurasi *Prediksi***

Untuk menghitung kesalahan (error) dalam melakukan prediksi pada sistem ini, maka penulis menggunakan rumus MAPE (*Mean Absolute Presentage Error*).

Dimana :

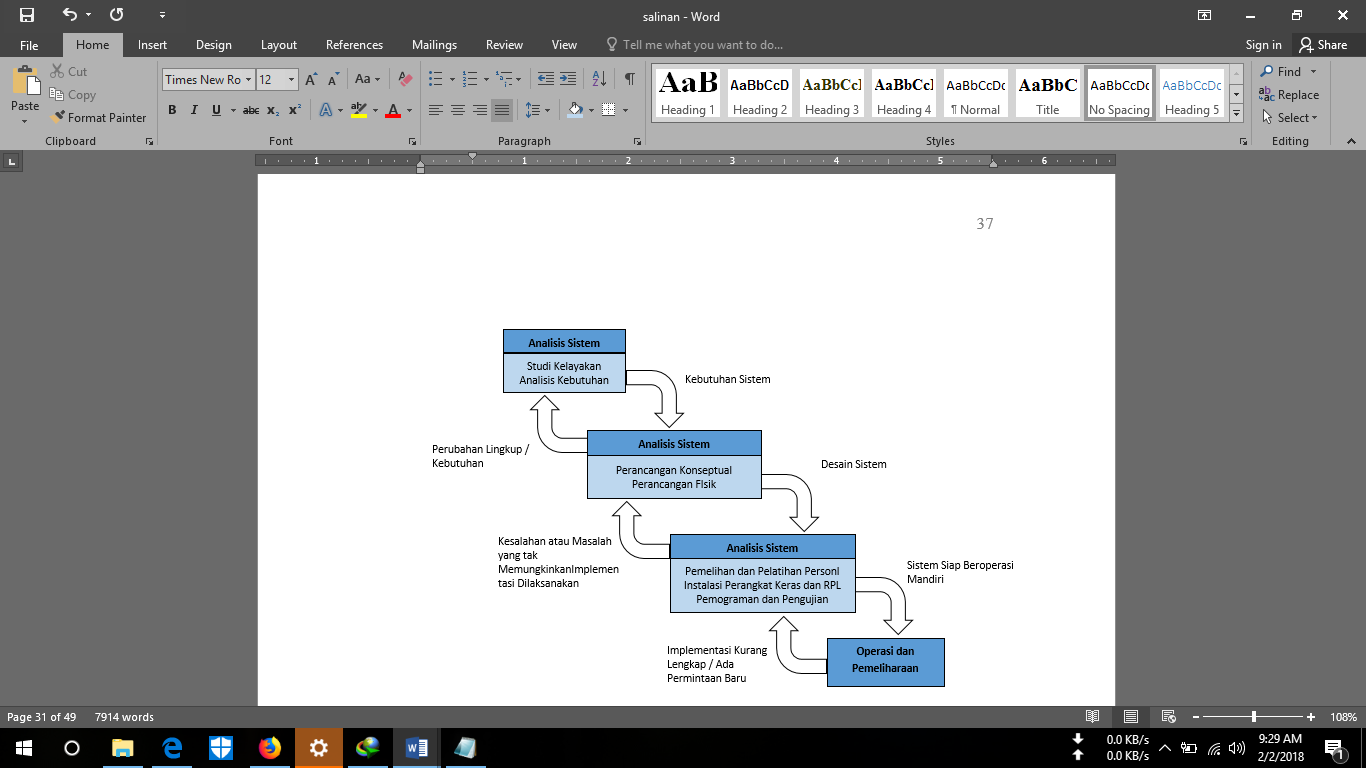
y’ : hasil prediksi

y : Data aktual

n : Jumlah data

**2.2.5 Siklus Hidup Pengembangan Sistem**

Menurut Jogiyanto (2005:41), Proses pengembangan sistem melewati beberapa tahapan dari mulai sistem itu direncanakan sampai dengan sistem tersebut diterapkan, dioperasikan dan dipelihara. Bila operasi yang sudah dikembangkan masih timbul kembali permasalahan-permasalahan yang tidak dapat diatasi dalam tahap pemeliharaan, maka perlu dikembangkan kembali suatu sistem untuk mengatasinya dan proses ini kembali ke tahap yang pertama, yaitu tahap perencanaan sistem. Siklus ini disebut dengan siklus hidup suatu sistem *(systems life cycle).* Daur atau siklus hidup dari pengembangan sistem merupakan suatu bentuk yang digunakan untuk menggambarkan tahapan utama dan langkah-langkah didalam tahapan tersebut dalam proses pengembangannya. Berikut langkah-langkah yang digunakan :



**Gambar 2.3**Siklus Hidup Pengembangan Sistem

**2.2.5.1 Perencanaan Sistem**

Kebijakanuntuk mengembangkan sistem informasi dilakukan oleh manajemen puncak karena menginginkan untuk meraih kesempatan-kesempatan yang ada yang tidak dapat diraih oleh sistem lama atau sistem yang lama mempunyai banyak kelemahan-kelemahan yang perlu diperbaiki. Setelah manajemen puncak menetapkan kebijakan untuk mengembangkan sistem informasi, sebelum sistem ini sendiri dikembangkan, maka perlu direncanakan terlebih dahulu dengan cermat. Perencanaan sistem ini menyangkut estimasi dari kebutuhan-kebutuhan fisik, tenaga kerja, dan dana yang dibutuhkan untuk mendukung pengembangan sistem ini serta untuk mendukung operasinya setelah diterapkan.

Selama fase perencanaan sistem, hal yang perlu dipertimbangkan adalah :

1. Faktor-Faktor Kelayakan (*Feasibility Factors*)yang berkaitan dengan kemungkinan berhasilnya sistem informasi yang dikembangkan dan digunakan.
2. Faktor-Faktor Strategis (*Strategic Factors*)yang berkaitan dengan pendukung sistem informasi dari sasaran bisnis dipertimbangkan untuk setiap proyek yang diusulkan. Nilai-nilai yang dihasilkan dievaluasi untuk menentukan proyek sistem mana yang akan menerima prioritas yang tertinggi

**2.2.5.2 Analisis Sistem**

Menurut Kusrini (2007 : 40), tahapan analisis sistem dimulai karena adanya permintaan terhadap sistem baru. Permintaan bisa datang dari seorang Pimpinan/Manajer di luar departemen sistem informasi yang melihat adanya masalah atau menemukan adanya peluang baru. Namun, adakalanya inisiatif pengembangan sistem baru berasal dari bagian yang bertanggung jawab terhadap pengembangan sistem informasi. Tujuan utama dari analisis sistem adalah menentukan hal-hal secara detail yang akan dikerjakan oleh sistem yang diusulkan.

Dalam menganalisis sistem pendukung keputusan akan dilakukan langkah-langkah pembuatan model, yaitu :

1. Proses studi kelayakan yang terdiri dari penentuan sasaran, pencarian prosedur, pengumpulan data, identifikasi masalah, identifikasi kepemilikan masalah, hingga akhirnya terbentuk sebuah pernyataan masalah.
2. Proses perancangan model. Dalam tahapan ini akan diformulasikan model yang akan digunakan serta kriteria-kriteria yang ditentukan. Setelah itu, dicari alternatif model yang bias menyelesaikan permasalahan tersebut. Langkah selanjutnya adalah memprediksi keluaran yang mungkin. Berikutnya, tentukan variabel-variabel model. Setelah beberapa altenatif model diberikan, pada tahap ini akan ditentukan satu model yang akan digunakan dalam sistem pendukung keputusan yang akan dibangun.

Di dalam tahap analisis sistem terdapat langkah-langkah dasar yang harus dilakukan oleh analis sistem, adalah sebagai berikut :

1. *Identify,* mengidentifikasi (mengenal) masalah merupakan langkah pertama yang dilakukan dalam tahap analisis sistem. Masalah dapat di definisikan sebagai suatu pertanyaan yang di inginkan untuk dipecahkan. Tahap identifikasi masalah sangat penting karena akan menentukan keberhasilan pada langkah-langkah selanjutnya.
2. *Understand,* adalah memahami kerja dari sistem yang ada. Langkah ini dapat dilakukan dengan mempelajari secara terinci bagaimana sistem yang ada beroperasi. Untuk mempelajari operasi dari sistem ini diperlukan data yang dapat diperoleh dengan cara melakukan penelitian.
3. *Analyze,* menganalisis sistem tanpa report.
4. *Report,* yaitu membuat laporan hasil analisis. Tujuan utama dari pembuatan laporan hasil analisis yaitu pelaporan bahwa analisis telah selesai dilakukan.

**2.2.5.3 Desain Sistem**

Dalam desain sistem, dibutuhkan alat bantu desain. Dalam tahapan ini, pengembang sistem bisa menentukan arsitektur sistemnya, merancang gambaran konseptual sistem, merancang database, perancangan interface, hingga membuat flowchart program. Salah satu alat bantu yang bisa digunakan dalam pembuatan sistem bantu keputusan adalah *Data Flow Diagram* (DFD). DFD adalah suatu model logika data atau proses yang dibuat untuk menggambarkan asal data dan tujuan data yang keluar dari sistem, tempat penyimpanan data, proses apa yang menghasilkan data tersebut, serta interaksi antara data yang tersimpan dan proses yang dikenakan pada data tersebut.

Menurut John Burch dan Gary Grudnitski, desain sistem dapat didefinisikan sebagai penggambaran, perencanaan, dan pembuatan sketsa atau pengaturan dari beberapa elemen yang terpisah ke dalam satu kesatuan yang utuh dan berfungsi. (Jogiyanto, 2005 : 196)

Tahap desain sistem mempunyai dua tujuan utama :

1. Untuk memenuhi kebutuhan kepada pemakai sistem.
2. Untuk memberikan gambaran yang jelas dan rancang bangun yang lengkap kepada pemrogram komputer dan ahli-ahli teknik lainnya.

Desain sistem dibagi dalam dua bagian, yaitu desain sistem secara umum *(general systems design)* dan desain sistem terinci *(detailed systems design).*

1. **Desain Sistem Secara Umum (*general systems design*)**

Pada tahap desain secara umum, komponen-komponen sistem informasi yang dirancang dengan tujuan dikomunikasikan kepada user bukan untuk pemrograman. Komponen sistem informasi yang di desain adalah model, output, input, database, teknologi, dan kontrol. (Jogiyanto,2005 : 211)

a**.** Desain Model Secara Umum

Analisis sistem dapat mendesain model dari sistem informasi yang di usulkan dalam bentuk *physical* sistem dan *logical* model. Bagan alir sistem merupakan alat yang tepat digunakan untuk menggambarkan *physical systems,* logical model dapat digambar dengan diagram arus data. (Jogiyanto,2005 : 211)

Bagan alir sistem merupakan bagan yang menunjukan arus pekerjaan secara keseluruhan dari sistem. Bagan alir sistem digambar dengan simbol-simbol sebagai berikut :

**Tabel 2.3**Daftar Simbol Bagan Alir Dokumen

| No. | Nama Simbol | Simbol | Keterangan |
| --- | --- | --- | --- |
|  | Terminal |  | Menunjukkan untuk memulai dan mengakhiri  Suatu proses |
|  | Dokumen |  | Menunjukkan dokumen input dan output baik itu proses manual, mekanik, atau computer |
|  | Kegiatan Manual |  | Menunjukan pekerjaan manual |
|  | Simpanan Offline | N  A  C | Menunjukkan file non-komputer yang diarsip urut angka (*numerical*), huruf (*alphabetical*), atau tanggal (*chronological*) |
|  | Proses |  | Menunjukkan kegiatan proses dari operasi program komputer |
|  | Operasi Luar |  | Menunjukkan operasi yang dilakukan diluar operasi computer |
|  | Hard Disk |  | Menunjukkan *input* dan *output* menggunakan *harddisk* |
|  | Keyboard |  | Menunjukkan *input* yang menggunakan *on-line keyboard* |
|  | Display |  | Menunjukkan *output* yang ditampilkan di monitor |
|  | Hubungan Komunikasi |  | Menunjukkan proses transmisi data melalui channel komunikasi |
|  | Garis Alir |  | Menunjukkan arus dari proses |
|  | Penjelasan |  | Menunjukkan penjelasan dari suatu proses |
|  | Penghubung |  | Menunjukkan penghubung ke halaman yang masih sama atau ke halaman yang lain |

(Sumber: Jogiyanto HM, 2005 : 802)

Untuk mempermudah penggambaran suatu sistem yang ada atau sistem yang baru yang akan dikembangkan secara logika dan tanpa memperhatikan lingkungan fisik data tersebut mengalir atau lingkungan fisik dimana data tersebut akan disimpan, maka digunakan Diagram Arus Data (DAD) atau *Data Flow Diagram* (DFD).

**Tabel 2.4**Daftar Simbol Diagram Alir Dokumen

| No | Simbol | Keterangan |
| --- | --- | --- |
|  |  | Simbol Proses, Menunjukan informasi dari masukan menjadi keluaran |
|  |  | Eksternal Entity, merupakan kesatuan dilingkungan luar system yang dapat berupa orang, organisasi atau system lain yang berada di lingkungan luarnya yang akan memberikan input seta menerima output dari system |
|  |  | Aliran atau arus data, menggambarkan gerakan paket data atau informasi dari suatu bagian kebagian yang lain, dimana penyimpanan mewakili lokasi penyimpana data |
|  |  | Penyimpanan, digunakan untuk memodelkan kumpulan data atau paket data |

(Sumber : Jogiyanto, 2005 : 700-807)

b.Desain Output Secara Umum

Output adalah produk dari sistem informasi yang dapat dilihat. Output terdiri dari macam-macam jenis seperti hasil di media kertas, dan hasil di media lunak. Disamping itu output dapat berupa hasil dari suatu proses yang akan digunakan oleh proses lain dan tersimpan di suatu media seperti tape, disk, atau kartu. Yang dimaksud dengan output pada tahap desain ini adalah output yang berupa tampilan di media kertas atau di layar video. (Jogiyanto,2005 : 213)

c. DesainInput Secara Umum

Alat input dapat digolongkan ke dalam 2 golongan, yaitu alat input langsung *(online input device)* dan alat input tidak langsung *(offline input device).* Alat input langsung merupakan alat input yang langsung dihubungkan dengan CPU, sedangkan alat input tidak langsung adalah alat input yang tidak langsung dihubungkan dengan CPU. (Jogiyanto, 2005 : 214)

d.Desain Database Secara Umum

Basis data (database) adalah kumpulan dari data yang saling berhubungan satu dengan yang lainnya, tersimpan diluar komputer dan digunakan perangkat lunak tertentu untuk memanipulasinya. Sistem basis data adalah suatu sistem informasi yang mengintegrasikan kumpulan dari data yang saling berhubungan satu dengan yang lainnya dan membuatnya tersedia untuk beberapa aplikasi yang bermacam-macam di dalam suatu organisasi. (Jogiyanto, 2005 : 217)

**2. DesainSistem Secara Rinci (*Detailed systems design***)

a.Desain Output Terinci

Desain output terinci dimaksudkan untuk mengetahui bagaimana dan seperti apa bentuk output-output dari sistem yang baru. Desain output terinci terbagi atas dua, yaitu desain output berbentuk laporan di media kertas dan desain output dalam bentuk dialog di layar terminal. (Jogiyanto,2005 : 362)

1. Desain output dalam bentuk laporan : dimaksudkan untuk menghasilkan output dalam bentuk laporan dimedia kertas. Bentuk laporan yang paling banyak digunakan adalah dalam bentuk tabel dan berbentuk grafik atau bagan. (Jogiyanto, 2005 : 362)
2. Desain output dalam bentuk dialog layar terminal : merupakan rancang bangun dari percakapan antara pemakai sistem atau user dengan komputer. Percakapan ini dapat terdiri dari proses memasukkan data ke sistem, menampilkan output informasi kepada user, atau keduanya.

b.Desain Input Terinci

Masukan merupakan awal dimulainya proses informasi. Bahan mentah dari informasi adalah data yang terjadi dari transaksi-transaksi yang dilakukan oleh organisasi. Data hasil dari transaksi merupakan masukan untuk sistem informasi. Hasil dari sistem informasi tidak lepas dari data yang dimasukan. Desain input terinci dimulai dari desain dokumen dasar sebagai penangkap input yang pertama kali. Jika dokumen dasar tidak di desain dengan baik, kemungkinan input yang tercatat dapat salah bahkan kurang. (Jogiyanto,2005 : 375)

Fungsi dokumen dasar dalam penanganan arus data :

1. Dapat menunjukkan macam dari data yang harus dikumpulkan dan ditangkap.
2. Dapat dicatat dengan jelas, konsisten, dan akurat.
3. Dapat mendorong lengkapnya data, disebabkan data yang dibutuhkan disebutkan satu persatu di dalam dokumen dasarnya.

c.Desain Database Terinci

Database merupakan kumpulan dari data yang saling berhubungan satu dengan yang lainnya, tersimpan di simpanan luar komputer dan digunakan perangkat lunak tertentu untuk memanipulasinya. Database merupakan salah satu komponen yang penting di dalam sistem informasi, karena berfungsi sebagai basis penyedia informasi bagi para pemakainya. Penerapan database dalam sistem informasi disebut *database system.* (Jogiyanto,2005 : 400)

**2.2.5.4 Seleksi Sistem**

Tahap ini merupakan tahap untuk memilih perangkat yang akan digunakan untuk sistem informasi. Pengetahuan dibutuhkan oleh pemilih sistem diantaranya adalah pengetahuan tentang siapa yang menyediakan teknologi ini, cara pemilikannya, dan sebagainya. Pemilihan sistem yang harus paham dengan teknik-teknik evaluasi untuk menyelesaikan sistem.

**2.2.5.5 Implementasi Sistem**

Menurut Kusrini (2007 : 43), Implementasi sistem merupakan tahapan untuk meletakkan sistem supaya siap untuk dioperasikan. Pada tahapan ini terdapat banyak aktifitas yang dilakukan, yaitu :

1. Pemrograman dan pengetesan program

Pemrograman merupakan kegiatan menulis program yang akan dieksekusi oleh komputer. Kode program harus berdasarkan dokumentasi yang disediakan oleh analis sistem hasil dari desain sistem.

1. Instalasi perangkat keras dan lunak

Proses pemasangan perangkat keras dan instalasi perangkat lunak yang sudah ada.

1. Pelatihan kepada pemakai

Manusia merupakan faktor yang diperlukan dalam sistem informasi. Jika ingin sukses dalam sistem informasi, maka personil-personil yang terlibat harus diberi pengertian dan pengetahuan tentang sistem informasi dan posisi serta tugas mereka.

1. Pembuatan dokumentasi

Dokumentasi adalah melakukan pencatatan terhadap setiap langkah pekerjaan pembuatan sebuah program yang dilakukan dari awal sampai selesai.

**2.2.5.6 Perawatan Sistem**

Perawatan sistem informasi adalah suatu upaya untuk memperbaiki, menjaga, menanggulangi, mengembangkan sistem yang ada. Perawatan ini di perlukan untuk meningkatkan efisiensi dan efektivitas kinerja sistem yang ada agar dalam penggunaannya dapat optimal. Beberapa alasan mengapa kita perlu memelihara sistem yang ada yaitu: agar dapat meningkatkan sistem / kinerja sistem, dan menyesuaikan dengan perkembangan, agar sistem yang ada tidak tertinggal.

Aplikasi yang professional dalam SDLC dan teknik maupun perangkat modeling yang mendukungnya adalah hal-hal keseluruhan yang terbaik yang dapat seseorang lakukan untuk meningkatkan maintainabilitas sistem.

Jenis-jenis perawatan sistem meliputi :

* 1. Perawatan korektif: adalah pemeliharaan yang mengkoreksi kesalahan-kesalahan yang ditemukan pada sistem, pada saat sistem di jalankan/berjalan.
  2. Pemeliharaan adaptif: yaitu pemelihaaan yang bertujuan untuk menyesuaikan perubahan yang terjadi.
  3. Pemeliharaan perfektif: pemeliharaan ini bertujuan untuk meningkatkan cara kerja suatu sistem.
  4. Pemeliharaan preventif: pemeliharaan ini bertujuan untuk menangani masalah-masalah yang ada.

**2.2.6 Teknik Pengujian Sistem**

Pengujian sistem adalah elemen kritis dari jaminan kualitas perangkat lunak dan mempresentasikan kajian pokok dari spesifikasi, desain, dan pengkodean. Tujuan dari pengujian ini adalah diharapkan dengan minimal tenaga dan waktu untuk menemukan berbagai potensi kesalahan dan cacat.Harus didasarkan pada kebutuhan berbagai tahap pengembangan, desain dan dokumen lain atau program yang dirancang untuk menguji struktur internal, dan menggunakan contoh-contoh ini untuk menjalankan program untuk mendeteksi kesalahan.Pengujian sistem informasi harus mencakup pengujian perangkat lunak, pengujian perangkat keras dan pengujian jaringan.Pengujian hardware, jaringan pengujian berdasarkan indikator kinerja spesifik yang akan digunakan di sini pengujian lebih jauh adalah pengujian perangkat lunak.

**2.2.6.1 *White Box***

Pengujian *white-box (glass box)*, adalah metode desain *test case* yang menggunakan struktur kontrol desain prosedural untuk memperoleh *test case*. Dengan menggunakan metode pengujian *white-box*, perekayasa sistem dapat melakukan *test case* untuk memberikan jaminan bahwa :

1. Semua jalur independen pada suatu modul ditelusuri minimal 1 (satu) kali.
2. Semua jalur keputusan logis *True/False* dilalui.
3. Semua *loop* dieksekusi pada batas yang tercantum dan batas operasionalnya.
4. Struktur data internal digunakan agar validitas terjamin.

Pengujian *white-box* bisa dilakukan dengan pengujian *basis path*, metode ini merupakan salah satu teknik pengujian struktur kontrol untuk menjamin semua statemen dalam setiap jalur independen program dieksekusi minimal 1 kali dan tidak menjumpai *error message*. Perhitungan jalur independen dapat dilakukan melalui metrik *Cyclomatic Complexity.* Sebelum menghitung nilai *Cyclomatic Complexity, harus* diterjemahkan desain prosuderal ke grafik alir, kemudian dibuat *flow graphnya*, seperti pada gambar di bawah ini (Roger S. Pressman, 2002 : 536).



**Gambar 2.4**Contoh Bagan Alir

****

**Gambar 2.5**ContohGrafik Alir

Keterangan :

1. *Node* adalah lingkaran yang merepresentasikan satu atau lebih statemen prosedural.
2. *Edge* adalah anak panah pada grafik alir.
3. *Region* adalah area yang membatasi edge dan node.
4. Simpul Predikat adalah simpul atau node yang berisi kondisi yang ditandai dengan dua atau lebih edge yang berasal darinya.

Darigambar *flowgraph* di atas didapat :

*Path* 1 =1– 11

*Path* 2 =1– 2 – 3 – 4 – 5 – 10– 1–11

*Path* 3 =1– 2 – 3 – 6 – 8 – 9 – 10– 1 – 11

*Path* 4 =1– 2 – 3 – 6 – 7 – 9–10–1–11

*Path* 1,2,3,4 yang telah didefinisikan diatas merupakan *basis set*untuk diagram alir.

*Cyclomatic complexity* digunakan untuk mencari jumlah *path*dalam satu *flowgraph*.Dapat dipergunakan rumusan sebagaiberikut :

1. Jumlah region grafikalir sesuaidengan*cyclomaticcomplexity*.

2. *Cyclomatixcomplexity*V(G) untukgrafikalir dihitung dengan rumus:

***V(G) =E– N +2*** …………………. (2.12)

Dimana :

E= jumlah*edge*pada grafikalir

N= jumlah *node*pada grafikalir

*Cyclomatixcomplexity*V(G) jugadapatdihitung dengan rumus :

***V(G) =P +1*** ………………….. (2.13)

DimanaP =jumlah*predicate node* pada grafikalir

Dari Gambar di atas dapat dihitung*cyclomaticcomplexity*:

1.*Flowgraph*mempunyai 4 region

2.V(G) =11 *edge*– 9 *node* +2 =4

3.V(G) =3 *predicatenode* +1 =4

Jadi*cyclomaticcomplexity*untuk*flowgraph*adalah4.

**2.2.6.2 *Black Box***

Pengujian *Black-Box* berusaha menemukan kesalahan dalam kategori :

1. Fungsi tidak benar atau hilang.
2. Kesalahan antar muka.
3. Kesalahan pada struktur data (pengaksesan basis data).
4. Kesalahan inisialisasi dan akhir program.
5. Kesalahan performasi.

Pengujian ini berfokus pada persyaratan fungsional perangkat lunak dan merupakan komplemen dari pengujian *White-Box*. Hal ini dapat dicapai melalui :

1. Pengujian *Graph-based*: dimulai dengan membuat grafik sekumpulan node yang mempresentasikan objek (misal *New File*, Layar baru dengan atributnya), link (hubungan antar objek), *node-weight* (misal nilai data tertentu seperti atribut layar, perilaku), dan link-weight (karakteristik suatu link, misal menu select).
2. *Equivalence Partitioning*: membagi domain input untuk pengujian agar diperoleh kelas-kelas kesalahan (misal kelompok data karakter, atau atribut yang lain).
3. Analisis Nilai Batas: pengujian berdasarkan nilai batas domain input.
4. Pengujian Perbandingan: disebut juga pengujian *back-to-back* yang diterapkan pada pada suatu versi perangkat lunak atau perangkat lunak redundan untuk memastikan konsistensinya.
   * 1. **Perangkat Lunak Pendukung**

Adapun perangkat lunak pendukung yang digunakan oleh penulis dalam membangun sistem ini ada beberapa diantaranya adalah:

**Tabel 2.5**Perangkat Lunak Pendukung

|  |  |  |
| --- | --- | --- |
| **No.** | **Perangkat Lunak Pendukung** | **Kegunaan** |
| 1. | PHP | Bahasa Pemrograman yang digunakan untuk membuat program. |
| 2. | Database MySQL | Sebuah perangkat lunak yang digunakan dalam pengoperasian basis data. |

**2.2.8 Kerangka Pikir**

1. SMA Negeri 1 Gorontalo Membutuhkan Aplikasi untuk Memprediksi Tingkat Kelulusan
2. Algoritma Naïve Bayes Dapat Digunakan Untuk Memprediksi Tingkat kelulusan

**Peluang**

1. Bagaimana cara merekayasa Sistem Prediksi Tingkat kelulusan Menggunakan Metode *Nive Bayes* ?
2. Bagaimana hasil penerapan Metode *Naïve Bayes* untuk Memprediksi Tingkat kelulusan SMA pada SMA N 1 Kota Gorontalo ?

**Masalah**

Membangun Aplikasi Data Mining untuk Memprediksi Prediksi kelulusan

**Solusi**

1. Sistem Berjalan
2. Sistem Diusulkan

**Analisa Sistem**

1. PHP
2. MySQL

**Pembangunan Sistem**

1. Desain Model
2. Desain User Interface(Desain Output, Desain Input, Desain Menu Utama)
3. Desain Database
4. Desain Teknologi

**Desain Sistem**

1. White Box.
2. Black Box.

**Pengujian**

SMA Negeri 1 Kota Gorontalo

**Implementasi**

1. Merekayasa aplikasi data mining untuk Memprediksi Tingkat kelulusan SMA dengan Metode *Naïve Bayes*.
2. Menerapkan Metode *Naïve Bayes* untuk memprediksi Tingkat kelulusan SMA

**Tujuan**

**Gambar 2.6** Kerangka Pikir