**BAB II**

**TINJAUAN PUSTAKA**

1. **Tinjauan Studi**

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

1. Penelitian ini yang dilakukan oleh Amrin (2016) yang berjudul Data Mining Dengan Regresi Linier Berganda untuk Peramalan Tingkat Inflasi. Pada penelitian ini akan digunakan metode prediksi regresi linier berganda, untuk memprediksi tingkat inflasi bulanan di Indonesia. Dari hasil analisis data yang dilakukan disimpulkan bahwa model regresi linier berganda yang dihasilkan pada penelitian ini adalah Y= 0,241X1 + 0,164X2 + 0,271X3 + 0,07X4 + 0,040X5 + 0,060X6 + 0,169X7 – 0,010. Adapun nilai koefisien determinasi sebesar 0,997 (99,7%). Performa model regresi linier berganda yang dibentuk dari data training dan divalidasi pada data testing memberikan tingkat akurasi prediksi yang cukup baik dengan nilai *Mean Absolute Deviation* (MAD) sebesar 0.0380, *Mean Square Error* (MSE) sebesar 0.0023, dan *nilai Root* *Mean Square Error* (RMSE) sebesar 0.0481.
2. Penelitian ini yang di lakukan oleh Abdul Munir dkk (2017). Analisis Metode *Linier Regression* Untuk Prediksi Penjualan Jamur Pada Jamur Karunia Berbasis Web. Perkembangan teknologi informasi menjadikan komputer tidak hanya sebagai alat pengolah data, namun sebagai sarana pendukung untuk menyelesaikan segala pekerjaan baik dikantor maupun di rumah. Penggunaan komputer juga merambah pada dunia ekonomi dan perdagangan yang berperan sebagai alat bantu dalam memaksimalkan usaha. Produsen jamur Karunia merupakan unit usaha yang bergerak dalam produksi dan penjualan jamur tiram. Karena jumlah permintaan setiap bulan bervariasi, sulit menentukan jumlah produksi setiap bulan secara tepat untuk memaksimalkan laba. Untuk mengantisipasi hal tersebut perlu melakukan prediksi penjualan, salah satu alternatif pemanfaatan prediksi yang bertujuan untuk memprediksi tingkat penjualan pada tahun yang akan datang. Pada penelitian ini akan dianalisa menggunakan metode Regresi Linear Berganda, metode ini merupakan metode sebab akibat atau statistik yang berfungsi untuk menguji sejauh mana hubungan sebab akibat antara variabel faktor independen (X) terhadap variabel dependen (Y). Variabel dependen yang digunakan adalah volume penjualan jamur, sedangkan variabel independen jumlah jam kerja dan biaya promosi. Dari hasil analisa menunjukkan biaya promosi (X1) dan jumlah jam kerja (X2) berpengaruh untuk dapat menaikkan volume penjualan. Jika biaya promosi dan jam kerja dinaikkan dan dilakukan secara terus menerus, maka diharapkan dapat meningkatkan penjualan pada produsen jamur Karunia.
3. Penelitian yang dilakukan oleh M. Syafruddin, Lukmanul Hakim, Dikpride Despa (2014). Metode Regresi Linear Untuk Prediksi Kebutuhan Energi Lisrtik Jangka panjang (Studi kasus Provinsi Lampung). Meningkatnya pembangunan di Provinsi Lampung terutama di sektor perumahanbaik sederhana maupun rumah mewah yang membawa konsekuensi logis berupa peningkatan kebutuhan tenaga listrik. Sebuah studi komprehensif dalam rangkapenyedian tenaga listrik di Lampung menjadi kebutuhan yang mendesakdilakukan untuk membuat rencana operasi sistem tenaga listrik. Salah satu faktoryang sangat menentukan dalam membuat rencana operasi sistem tenaga listriktersebut adalah prediksi beban listrik yang akan ditanggung oleh sistem tenagalistik yang bersangkutan. Penelitian ini bertujuan untuk memprediksi kebutuhan energi listrik di Provinsi Lampung hingga Tahun 2030, diharapkan dapatdijadikan sebagai masukkan dalam melakukan perencanaan pembangunan sistemtenaga listrik. Prediksi kebutuhan energi listrik Lampung dibagi menjadi 4 sektoryaitu : sektor rumah tangga, bisnis, publik, dan industri. Proses perancanganprediksi kebutuhan energi listrik menggunakan 6 variabel dan dibagi menjadi 2 parameter, yaitu: parameter ekonomi (produk domestik regional bruto, jumlah penduduk, jumlah rumah tangga) dan parameter listrik (rasio elektrifikasi, faktor beban, losses). Dengan menggunakan metode regresi linier untuk memprediksivariabel-variabel di atas, diperoleh hasil prediksi daya listrik tersambung totalpada tahun 2028 sebesar 2.841,78 MVA (rata-rata pertumbuhannya sebesar 2,38 %), dan konsumsi energi listrik pada tahun 2023 sebesar 5.934,98 Gwh (rata-rata pertumbuhannya sebesar 3, 83 %).
4. **Tinjauan Teori**

**2.2.1 Lulusan**

Lulusan memiliki 1 arti. Lulusan berasal dari kata lulus. lulusan memiliki arti dalam kelas *nomina* atau kata benda sehingga lulusan dapat menyatakan nama dari seseorang, tempat, atau semua benda dan segala benda yang dapat dibendakan

Lulusan menurut Kamus Besar Bahasa Indonesia (KBBI) adalah yang sudah lulus dari ujian; tamatan

**2.2.2 Data 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, J & Kamber, M, 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..2.3 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

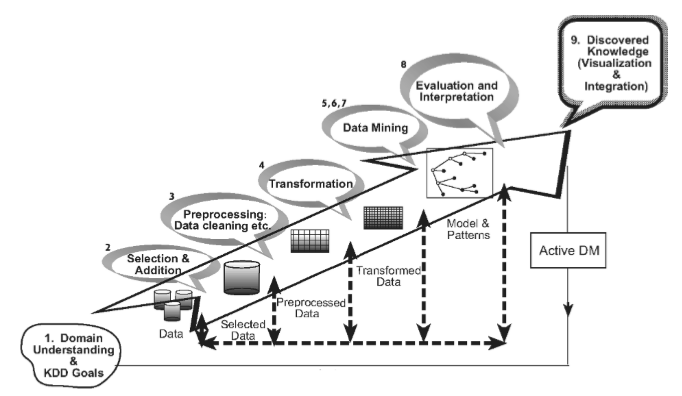
Pada tahap ini dilakukan implementasi dari algoritma data mining yang telah ditentukan pada tahap sebelumnya.

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)



**Gambar 2.1.** Proses KDD

Sumber :Maimon, 2005

Selama ini *data mining* berkembangdigunakan untuk menyelesaikan masalah menyangkutpendidikan.*Data mining* digunakan untuk menelusuri data yang ada untuk membangun sebuahmodel, kemudian menggunakan model tersebut agardapat mengenali pola data yang lain yang tidak beradadalam basis data yang disimpan. Salah satu teknik*data mining* adalah teknik klasifikasi. Teknik klasifikasiadalah teknik pembelajaran untuk prediksi suatu nilaidari target variabel kategori

Kegunaan data mining adalah untuk mengklasifikasikan pola yang harus ditemukandalam data mining. Secara umum, data mining dapat diklasifikasikan dalam duakategori yaitudeskriptif dan prediktif. (Whitten, I. H. dkk, 2011)

Adapun operasi-operasi dan teknik-teknik yang berhubungan:

1. Operasi *Predictive modeling* : (*classification, value prediction*)
2. *Database segmentation* : (*demographic clustering,neural clustering*)
3. *Link Analysis* : (*association discovery, sequential pattern discovery, similar timesequencediscovery*)
4. *Deviation detection*: (*statistics, visualization*

Hasil dari *data mining* sering kali diintegrasikan dengan *decision support system (DSS).* Sebagai contoh, dalam aplikasi bisnis informasi yang dihasilkan oleh *data mining* dapat diintegrasikan dengan tool manajemen kampanye produk sehingga promosi pemasaran yang efektif yang dilaksanakan dan dapat diuji. Integrasi demikian memerlukan langkah *postprocessing* yang menjamin bahwa hanya hasil yang *valid* dan berguna yang akan digabungkan dengan DSS. Salah satu pekerjaan dan *postprocessing* adalah visualisasi yang memungkinkan analist untuk mengeksplor data dan hasil *data mining* dari berbagai sudur pandang. Ukuran-ukuran statistik dan metode pengujian hipotesis dapat digunakan selama *postprocessing* untuk membuang hasil *data mining* yang palsu. Gambar 2.2 menunjukkan hubungan data mining dengan area-area lain.



**Gambar 2. 1**Data mining *sebagai pertemuan dari banyak disiplin ilmu*

Sumber :(Tan et al dalam Eko Prasetyo, 2012)

Secara khusus, data mining menggunakan ide-ide seperti (1) pengambilan contoh, estimasi, dan pengujian hipotesis, dari statistika dan (2) *algoritme*  pencarian, teknik pemodelan, dan teori pembelajaran dari kecerdasan buatan, pengenalan pola, dan *machine learning*. Data mining juga telah mengadopsi ide-ide dari area lain meliputi optimisasi, *evolutionary computing*, teori informasi, pemrosesan sinyal, visualisasi dan *information retrieval*. Sejumlah area lain juga memberikan peran pendukung dalam data mining, seperti sistem basis data yang dibutuhkan untuk menyediakan tempat penyimpanan yang efisien, *indexing* dan pemrosesan *query.*

* + 1. **Forecasting**

Peramalan (*forecasting*) merupakan suatu usaha untuk meramalkan keadaan dimasa mendatang melalui pengujian keadaan dimasa lalu (Prasetya, 2009).

*Forecasting* adalah proses analisis untuk memperkirakan masa depan dengan metode-metode tertentu dan mempertimbangkan segala variabel yang mungkin berpengaruh di dalamnya.

*Forecasting* berperan sangat penting dalam bisnis. Kemampuan untuk memprediksi secara akurat kejadian di masa depan menjadi dasar dalam pengambilan keputusan. Kemampuan *forecasting* banyak dipakai di bidang marketing, produksi, pengendalian inventori, dan banyak aktivitas bisnis lainnya.

Penentuan periode waktu peramalan tergantung pada situasi dan kondisi aktual dan tujuan peramalan. Periode waktu yang biasa digunakan adalah harian, mingguan, bulanan, triwulan, semesteran, dan tahunan. Semakin jauh periode di makan mendatang yang akan diramalkan maka hasil ramalan akan semakin kurang akurat.

Ada berapa macam tipe peramalan yang digunakan. Adapun tipe-tipe dalam peramalan (Heizer & Barry Render, 2005)adalah:

1. Times Series Model

Metode time series adalah metode peramalan secara kuantitatif dengan menggunakan waktu sebagai dasar peramalan.

2. Causal Model

Metode peramalan yang menggunakan hubungan sebab-akibat sebagai asumsi, yaitu bahwa apa yang terjadi di masa lalu akan terulang pada saat ini.

3. Judgemental Model

Bila time series dan causal model bertumpu pada kuantitatif, pada judgemental mencakup untuk memasukkan faktor-faktor kuantitatif/subjektif ke dalam metode peramalan. Secara khusus berguna bilamana faktor-faktor subjektif yang diharapkan menjadi sangat penting bilamana data kuantitatif yang akurat sudah di peroleh.

## Metode Regresi Linier Berganda

Regresi Linear atau merupakan metode statistik yang bertujuan untuk membentuk sebuah model antara variabel dependen (Y) dengan variabel independen (X). Regresi Linear yang memiliki satu variabel bebas disebut dengan Regresi Linear Sederhana, sedangkan Regresi Berganda diperuntukkan apabila memiliki lebih dari satu variabel bebas. Regresi Linear menggunakan garis kecenderungan apabila pola data menunjukkan suatu kecenderungan, baik berpola turun atau naik.

Adapun kelebihan menggunakan metode Regresi Linier Berganda maka kita dapat menganalisis dengan menggunakan beberapa variabel bebas (X) sehingga hasil prediksi yang didapatkan lebih akurat dibandingkan dengan Regresi Linier sederhana yang hanya menggunakan satu variabel bebas (X), sedangkan kelemahannya adalah tidak mampu menunjukkan titik jenuh fungsi yang sedang diselidiki akibatnya selalu timbul kemungkinan kesalahan prediksi dan terdapat kemungkinan terjadinya multikoliniearitas pada variabel bebas. Akibatnya variable bebas tidak mampu menjelaskan variable tak bebas (hubungan antara X dan Y tidak bermakna). (Magdalena Iriani Kehi, dkk)

Regresi Linear merupakan bagian regresi yang mencakup hubungan linear satu peubah tak bebas Y dengan satu peubah bebas X. Analisis regresi adalah metode untuk menentukan hubungan sebab-akibat antara satu variabel dengan variabel-variabel yang lain. Variabel “penyebab” disebut dengan bermacam-macam istilah yaitu, variabel penjelas, variabel eksplanatorik, variabel independen, atau secara bebas, variabel X (karena seringkali digambarkan dalam grafik sebagai absis atau sumbu X). Variabel terkena akibat dikenal sebagai variabel yang dipengaruhi, variabel dependen, variabel terikat, atau variabel Y. Kedua variabel ini dapat merupakan variabel acak *(random*), namun variabel yang dipengaruhi harus selalu variabel acak.

Bentuk umum Regresi Linear Berganda sebagai berikut:

Dimana:

Y ̂ : nilai variabel Y hasil prediksi

Y : variabel tak bebas

X : variabel bebas

a : konstanta regresi

b : koefisien regresi

Untuk mencari nilai konstanta a dan koefisien regresi b dapat dihitung dengan cara pendekatan matriks, persamaan normal (substitusi), dan metode kuadrat terkecil. Pada penelitian ini penulis menggunakan metode kuadrat terkecil, dari persamaan 2, maka dapat dicari nilai *konstanta* a dengan rumus:

Untuk , digunakan rumus:

b1 dan b2 dicari dengan persamaan:

**2.2.6 Analisis Hasil Akurasi Prediksi**

Hasil dari prediksi apabila disajikan dalam diferensiasi teknik yang berbeda tentunya memiliki hasil yang berbeda. Perlu suatu konsep dalam menilai teknik mana yang paling optimum dalam memberikan nilai prediksi berdasarkan pola data tertentu. Konsep penilain ini akan menghasilkan berbagai macam metode yang bertujuan untuk menilai sejauh mana data aktual dengan hasil prediksi. MAPE (*Mean Absolute Percentage Error*) merupakan metode yang digunakan untuk menilai tingkat keakuratan. (Armstrong. JS, 1983) dalam Hendy Tannady dan Fan Andrew, 2013). Suatu model mempunyai kinerja sangat bagus jika nilai MAPE berada di bawah 10%, dan mempunyai kinerja bagus jika nilai MAPE  berada di antara 10% dan 20% (Zainun dan Majid, 2003).

Dimana :

y’ : hasil prediksi

y : Data aktual

n : Jumlah data

**2.2.7 Penerapan Algoritma Regresi Linear Berganda**

Sebelum melakukan analisis perhitungan prediksi penjualan jamur terlebih dahulu menyiapkan data mentah dengan tujuan agar data tersebut siap untuk diolah selanjutnya dengan metode Regresi Linear. Sumber data mentah diperoleh dari Koperasi Produsen Kurnia. Adapun data yang akan digunakan untuk memprediksi jumlah penjualan jamur Koperasi Produsen Kurnia dapat dilihat pada tabel 2.1

**Tabel 2.1.** Data Volume Penjualan Jamur Koperasi Produsen Karunia Tahun 2015

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| No. | Bulan (n) | X1(Jam) | X2(Rp.) | Y(Rp.) |
| Jam Kerja Pegawai | Biaya Promosi | Volume Penjualan |
| 1 | Januari | 240 | 120.500 | 5.000.000 |
| 2 | Februari | 236 | 250.000 | 5.400.000 |
| 3 | Maret | 238 | 210.000 | 3.500.000 |
| 4 | April | 240 | 275.000 | 3.100.000 |
| 5 | Mei | 237 | 320.000 | 5.300.000 |
| 6 | Juni | 241 | 120.000 | 3.000.000 |
| 7 | Juli | 237 | 150.000 | 4.100.000 |
| 8 | Agustus | 239 | 155.000 | 5.200.000 |
| 9 | September | 240 | 125.000 | 5.500.000 |
| 10 | Oktober | 241 | 175.000 | 4.800.000 |
| 11 | November | 235 | 130.000 | 3.400.000 |
| 12 | Desember | 236 | 150.000 | 4.500.000 |

## (Sumber : Abdul Munir dkk, 2017)

Berdasarkan data pada tabel 1 diatas, maka diketahui variabel dependen (Y) = Volume Penjualan, sedangkan variabel independen (X1) = Jam Kerja dan (X2) = Biaya Promosi. Analisa metode Regresi Linear Berganda dimulai dengan menghitung nilai *konstanta* a dan *koefisien regresi* b1 dan b2. Untuk mencari nilai *konstanta* a dan *koefisien regresi* b1 dan b2 digunakan tabel bantu sebagai berikut:

**Tabel 2.2**. Menentukan Nilai *Konstanta* Dan *Koefisien Regresi*



Berdasarkan tabel 2 diatas maka di dapatkan:

Maka diperoleh nilai *konstanta* a dan *koefisien regresi* b1 dan b2:

4400000 – (-15764,682 \* 238,333 – 1,262 \* 181708,333)

8386559,871

Sehingga didapat model persamaan regresi dari hasil perhitungan kasus diatas adalah sebagai berikut:

Setelah model persamaan Regresi Linear didapat, maka tahap selanjutnya adalah melakukan prediksi penjualan jamur untuk periode mendatang. Berikut contoh hasil perhitungan prediksi menggunakan metode Regresi Linear. Prediksi bulan Januari tahun 2016 dengan X1 = 240 dan X2 = 120.500

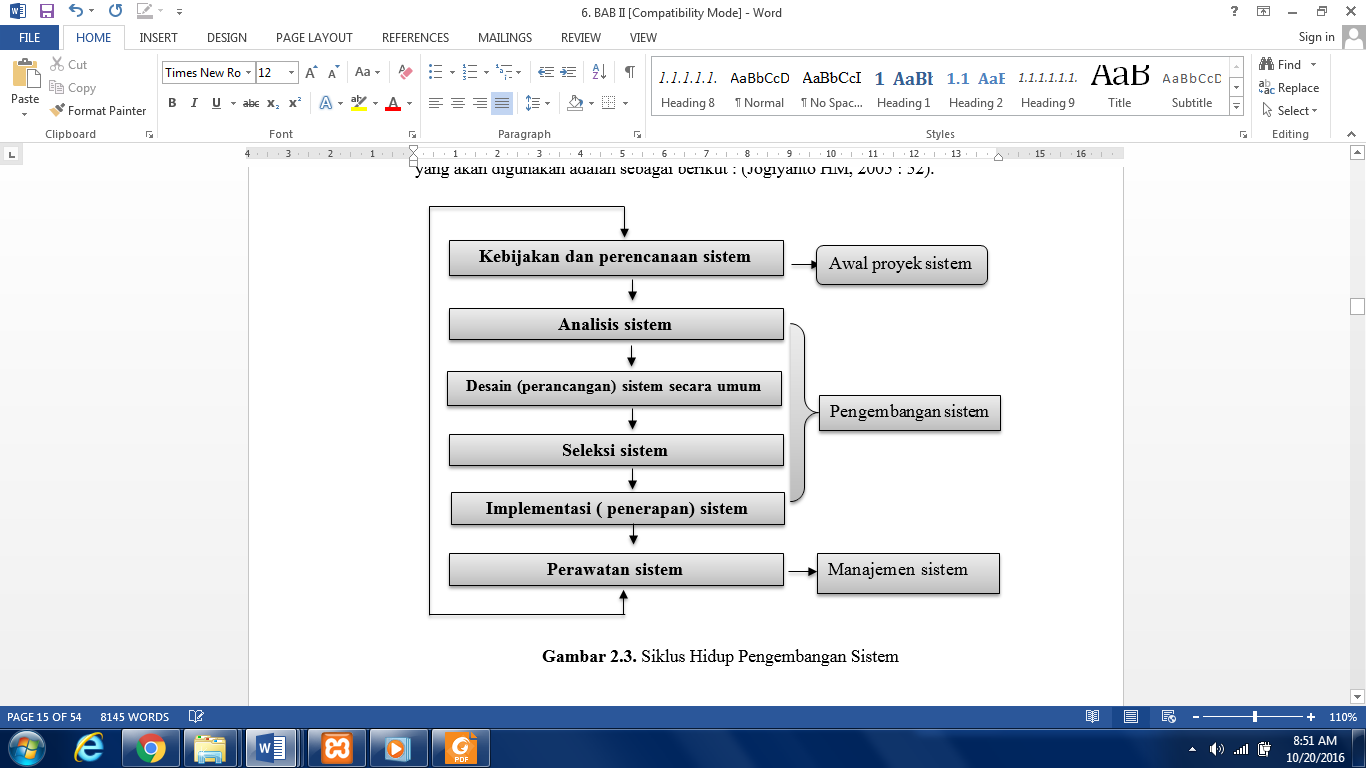
**2.2.8 Siklus Hidup Pengembangan Sistem**

Pengembangan sistem informasi yang berbasis komputer dapat merupakan tugas kompleks yang membutuhkan banyak sumber daya dan dapat memakan waktu berbulan-bulan bahkan bertahun tahun untuk menyelesaikannya. Proses pengembangan sistem melewati beberapa tahapan dari mulai sistem itu direncanakan sampai dengan sistem tersebut diterapkan, dioperasikan dan dipelihara. Bila operasi sistem yang sudah dikembangkan masih timbul kembali permasalahan-permasalahan yang kritis serta tidak dapat diatasi dalam tahap pemeliharaan sistem, maka perlu dikembangkan kembali suatu sistem untuk mengatasinya dan proses ini kembali ketahap 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.

Ide dari *systems life cycle* adalah sederhanadan masuk akal. Di *systemslife cycle*, tiap-tiap bagian dari pengembangan sistem dibagi menjadi beberapa tahapan kerja. Tiap-tiap tahapan ini mempunyai karakteristik tersendiri. Tahapan utama siklus hidup pengembangan sistem dapat terdiri dari tahapan perencanaan sistem, analisis sistem, desain sistem, seleksi sistem, implementasi sistem dan perawatan sistem. Tahapan-tahapan seperti ini sebenarnya merupakan tahapan di dalam pengembangan sistem teknik.

Siklus hidup pengembangan sistem dengan langkah-langkah utamanya yang akan digunakan adalah sebagai berikut :



**Gambar 2. 3** Siklus hidup pengembangan sistem

Sumber : Jogiyanto HM, (2005 : 52 )

**2.2.8.1 Perencanaan Sistem**

Perencanaan sistem merupakan kata lain dari sebuah konsep, dimana dalampengembangan suatu sistem/perangkat lunak konseptualisasi ini dilakukan denganmaksud tujuan tertentu.

Hariyanto (2004 : 353) mengungkapkan :“Tujuan konseptualisasi adalah untuk menghasilkan spesifikasi perilakusistem yang disepakati antara pembeli dan pengembang, pemakai danstakeholder lain serta merupakan kontrak resmi pengembang dan client,juga menjadi dokumen yang menuntun pemrogram dalam implementasisistem”.

Perencanaan atau *planning* adalah hal-hal yang menyangkut studi tentang kebutuhan pengguna atau (*user’s spesification*), studi kelayakan *(feasibility study*)baik secara teknis maupun secara teknologi serta penjadwalan pengembangansuatu proyek sistem informasi dan/atau perangkat lunak. Yang mana pada tahapperencanaan ini pengembang melakukan observasi untuk mengenalicalon pengguna dari sistem informasi/perangkat lunak yang akan dikembangkannantinya.

**2.2.8.2 Analisis Sistem**

Analisa sistem (*System Analisa* ) dapat didefinisikan sebagai penguraian dari suatu sistem informasi yang utuh kedalam bagian-bagian komponennya dengan maksud untuk mengidentifikasikan dan mengevaluasi permasalahan-permasalahan, hambatan-hambatan yang terjadi dan kebutuhan-kebutuhan yang diharapkan sehingga dapat diusulkan perbaikan-perbaikannya.

Analisa sistem adalah spesialis yang mempelanjari masalah dan kebutuhan sebuah organisasi untuk menentukan bagaimana orang, data, proses dan teknologi informasi dapat mencapai kemajuan terbaik untuk bisnis.

Analisis sistem adalah *Stakeholder* yang berperan sebagai fasilitator atau pelatih, menjebatani jurang komunikasi yang dapat secara alamiah berkembang antara pemilik dan pengguna *system nonteknis* atau desainer dan perkembangan sistem teknis.

Whitten, et al. (2004 :33) mengungkapkan “ *System analysis* adalah study domain masalah bisnis untuk merekomendasikan perbaikan dan menspesifikasi persyaratan dan prioritas bisnis untuk solusi”.

Impak teknologi objek sangat berarti dalam dunia analisis dan desain sistem. Sebelum ada teknologi objek, kebanyakan bahasa pemrograman didasarkan pada apa yang disebut metode yang terstruktur ( *structured method*). Contohnya COBOL bahasa yang domain 0, C, Fortan, Pascal, dan PL/i. Maka, metode analisis dan desain berorientasi objek telah muncul sebagai pendekatan terpilih untuk membangun kebanyakan sistem informasi saat ini.

Sebagai tambahan keahlian analisis dan desain sistem formal, seorang analis harus mengembangkan atau memilki keahlian lain, pengetahuan, dan karakter untuk menyelesaikan pekerjaan. Hal ini termasuk:

1. Pengalaman dan keahlian pemrograman komputer.
2. Sulit untuk membayangkan bagaimana para analisis sistem dapat dengan cukup mempersiapkan bisnis dan spesifikasi teknis untuk programer jika mereka tidak memilki pengalaman programan. Kebanyakan analis system harus menguasai satu atau lebih bahasa pemrograman tingkat tinggi.
3. Pengetahuan umum proses dan teknologi bisnis.

Analis sistem harus mampu berkomunikasi dengan para ahli bisnis untuk memperoleh pemahaman masalah dan kebutuhan mereka. Untuk analis, paling tidak sebagian dari pengetahuan ini datang hanya dari pengalaman. Pada saat yang sama analis yang terinspirasi harus mengambil manfaat dari setiap kesempatan untuk menyelesaikan mata kuliah teori bisnis dasar.

Tahap analisis merupakan tahap yang kritis dan sangat penting, karena kesalahan didalam tahap ini akan meyebabkan juga kesalahanditahap selanjutnya. Tahap analisa sistem mencakup studi kelayakan analisis kebutuhan.

1. Studi Kelayakan.

Studi kelayakan digunakan untuk menentukan kemungkinan keberhasilan solusi yang diusulkan. Tahapan berguna untuk memastikan bahwa solusi yang diusulkan tersebut benar-benar dapat dicapai dengan sumber daya dan dengan memperhatikan kendala yang terdapat pada perusahan serta dampak terhadap lingkungan sekeliling. Tugas-tugas yang tercakup dalam studi kelayakan meliputi:

1. Penentuan masalah dan peluang yang dituju sistem.
2. Pembentukan sasaran sistem baru secara keseluruhan.
3. Pengidentifikasian para pemakai sistem.
4. Pembentukan lingkup sistem.

Selain itu, selama dalam tahapan studi kelayakan sistem analisis juga melakukan tugas-tugas sebagai berikut :

1. Pengusulan perangkat lunak dan perangkat keras untuk sistem baru.
2. Pembuatan analisis untuk membuat atau membeli aplikasi.
3. Pembuatan analisis biaya/manfaat.
4. Pengkajian terhadap resiko proyek.

Studi kelayakan diukur dengan memperhatikan aspek teknologi, ekonomi, faktor organisasi dan kendala hukum, etika, dan yang lain (Turban, *et, al,* 1999 dalam Abdul Kadir, 2003:403 ).

1. Analisis kebutuhan.

Analisis kebutuhan dilakukan untuk menghasilkan spesifikasi kebutuhan ( disebut juga spesifikasi fungsional ). Spesifikasi kebutuhan adalah spesifikasi yang rinci tengtang hal-hal yang akan dilakukan sistem ketika diimplementasikan. Spesifikasi ini sekaligus dipakai untuk membuat kesepakatan antara pengembang sistem, pemakai yang kelak akan menggunakan sistem, manajemen, dan mitra kerja yang lain (misalnya auditor inetrnal).

Analisis kebutuhan ini diperlukan untuk menentukan keluaran yang akan dihasilkan sistem, masukan yang diperlukan sistem, lingkup proses yang digunakan untuk mengolah masukan menjadi keluaran, volume data yang akan ditangani sistem, jumlah pemakai dan kategori pemakai, serta kontrol terhadap sistem.

Didalam tahap analisis ini sistem terdapat langkah-langkah dasar yang harus dilakukan oleh analisis sistem, yaitu sebagai berikut :

1. *Identify,* yaitu mengidentifikasi masalah.

Mengidentifikasi (mengenai) masalah merupakan langkah pertama yang dilakukan dalam tahap analisis sistem. Masalah (*problems*) dapat didefinisikan sebagai suatu pertanyaan yang diinginkan untuk dipecahkan. Tahap indentifikasi sebagai suatu pertanyaan yang diinginkan untuk dipecahkan. Tahap identifikasi masalah sangat penting karena akan menentukan keberhasilan pada langkah-langkah selanjutnya.

1. *Understand,* yaitu memahami kerja dari sistem yang ada.

Langkah kedua dari tahap analisis sistem adalah memahami kerja dari sistem yang ada. Langkah ini dapat dilakukan dengan mempelajari operasi dari sistem ini diperlukan data yang dapat diperoleh dengan cara melakukan penelitian.

1. *Analyze,* yaitu menganalisis sistem tanpa report.

Langkah ini dilakukan berdasarkan data yang telah diperoleh dari hasil penelitian yang telah dilakukan.

1. *Report,* yaitu membuat laporan hasil analisis.

Tujuan utama dari pembuatan laporan hasil dilakukan ;

1. Pelaporan bahwa analisi telah selesai dilakukan.

Meluruskan kesalah pengertian mengenai apa yang telah ditemukan dan dianalisis oleh analis sistem tetapi tidak sesuai menurut manajemen.

**2.2.8.3 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. **DesainSistem 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)

**2.2.8.4 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.

**2.2.8.5 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.

**2.2.8.6Perawatan 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.

**2.2.9 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.

**2.2.9.1 *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*** …………………. (2.5)

Dimana :

E= jumlah*edge*pada grafikalir

N= jumlah *node*pada grafikalir

*Cyclomatixcomplexity*V(G) jugadapatdihitung dengan rumus :

***V(G) =P +1*** ………………….. (2.6)

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.

**2.2.9.2 *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. | Microsoft Visual Basic Net 2010 | Bahasa Pemrograman yang digunakan untuk membuat program. |
| 2. | Database MySQL | Sebuah perangkat lunak yang digunakan dalam pengoperasian basis data. |
|  | Crystall Report for Visual Studio | Digunakan untuk pembuatan laporan. |

* 1. **Kerangka Pikir**

1. Kaprodi Teknik Informatika membutuhkan aplikasi untuk Prediksi Jumlah Lulusan Berdasarkan Angkatan
2. Algoritma Regresi Linier dapat digunakan Untuk Prediksi Jumlah Lulusan Berdasarkan Angkatan

**Peluang**

1. Bagaimana cara merekayasa sebuah aplikasi penerapan data mining algoritma Regresi linierUntuk Prediksi Jumlah Lulusan Berdasarkan Angkatan ?
2. Bagaimana hasil penerapanRegresi linier Untuk Prediksi Jumlah Lulusan Berdasarkan Angkatan?

**Masalah**

**2.4 Kerangka Pikir**

Membangun sebuah Aplikasi Data Mining Prediksi Jumlah Lulusan Berdasarkan Angkatan

**Solusi**

1. Sistem Berjalan
2. Sistem Diusulkan

**Analisa Sistem**

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

**Desain Sistem**

1. Microsoft Visual Basic Net 2010
2. MySQL
3. Crystall Report for Visual Studio

**Pembangunan Sistem**

1. White Box.
2. Black Box.

**Pengujian**

1. Merekayasa aplikasi penerapan data mining algoritma Regresi Linier **Untuk Prediksi Jumlah Lulusan Berdasarkan Angkatan**
2. Menerapkan hasil dari algoritma Regresi Linier Untuk **Prediksi Jumlah Lulusan Berdasarkan Angkatan** pada Fakultas Ilmu Komputer Universitas Ichsan Gorontalo

**Tujuan**

Fakultas Ilmu Komputer Universitas Ichsan Gorontalo

**Implementasi**

**Gambar 2.6** Kerangka Pikir