**BAB II  
LANDASAN TEORI**

**2. 1 Tinjauan Studi**

Berikut ini adalah beberapa penelitian terdahulu yang terkait dengan metode K-Means, yaitu :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No** | **Peneliti** | **Judul** | **Tahun** | **Hasil** |
| 1 | Widodo, Dina wahyuni | Implementasi algoritma K-Means Clustering untuk megengetahui bidang Skripsi mahasiswa multimedia | 2017 | Persiapan data dilakukan secara  manual yaitu menggunakan Ms. Excel.  Kolom ke-1 merupakan no.registrasi  mahasiswa, kolom ke-2 nama mahasiswa, kolom ke-3 nilai 1 yaitu rata-rata nilai dari mata kuliah proyek video digital, desain grafis, dan animasi komputer, serta kolom ke-4 nilai 2 yaitu nilai rata-rata dari mata kuliah algoritma dan pemrograman, struktur data, serta pemrograman webHasil data yang disajikan setelah mengalami proses awal (pre-processing) berjumlah 53 record dari jumlah data awal 83 record. |
| 2 | Rima Dias Ramadhani | Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Universitas Dian Nuswantoro | 2016 | Hasil dari penelitian ini digunakan sebagai salah satu dasar pengambilan keputusan untuk menentukan Kelompok berdasarkan cluster yang terbentuk oleh pihak admisi UDINUS |
| 3 | Johan Oscar Ong | Implementasi Algoritma *K-Means Clustering* Untuk Menentukan Strategi Marketing President University | 2015 | Hasil penelitian ini analisa data mining dilakukan dengan menggunakan metode *K-Means clustering.* Dengan menggunakan metode ini, datadata yang telah didapatkan dapat dikelompokan kedalam beberapa *cluster* berdasarkan kemiripan dari data-data tersebut, sehingga data-data yang memiliki karakteristik yang sama dikelompokan dalam satu *cluster* dan yang memliki karakteristik yang berbeda dikelompokan dalam *cluster* yang lain yang memiliki karakteristik yang sama. Dengan adanya pengelompokan-pengelompokan data seperti ini, diharapkan bagian marketing dapat melakukan pemasaran dengan strategi yang tepat untuk mendapatkan calon mahasiswa baru. |

**Tabel 2.1** Penelitian Terkait

**2.2 Tinjauan Pustaka**

**2.2.1 Usaha Mikro Kecil Menengah(UMKM)**

Usaha Mikro Kecil Menengah merupakan sektor usaha yang memegang peran penting dalam perekonomian Indonesia baik dilihat dari kontribusi terhadap PDB Nasional maupun dalam hal penyerapan tenaga kerja. IKM juga merupakan satu-satunya sektor ekonomi yang mampu bertahan dalam menghadapi terpaan krisis moneter yang memporakporandakan struktur ekonomi Indonesia. Seiring dengan era globalisasi saat ini, UMKM dituntut untuk melakukan pembenahan dan perubahan agar dapat meningkatkan daya saingnya.

Berdasarkan Undang-Undang Nomor 3 Tahun 2014, Peraturan Pemerintah Nomor 17 Tahun 1986 dan Keputusan Presiden Nomor 16 Tahun 1986 jenis Industri Kecil danMenengah dalam ruang lingkup industri pengolahan yaitu :

* 1. Industri Pangan

1. Industri Sandang
2. Industri Kimia dan Bahan Bangunan
3. Industri Logam dan Elektronika
4. Industri kerajinan

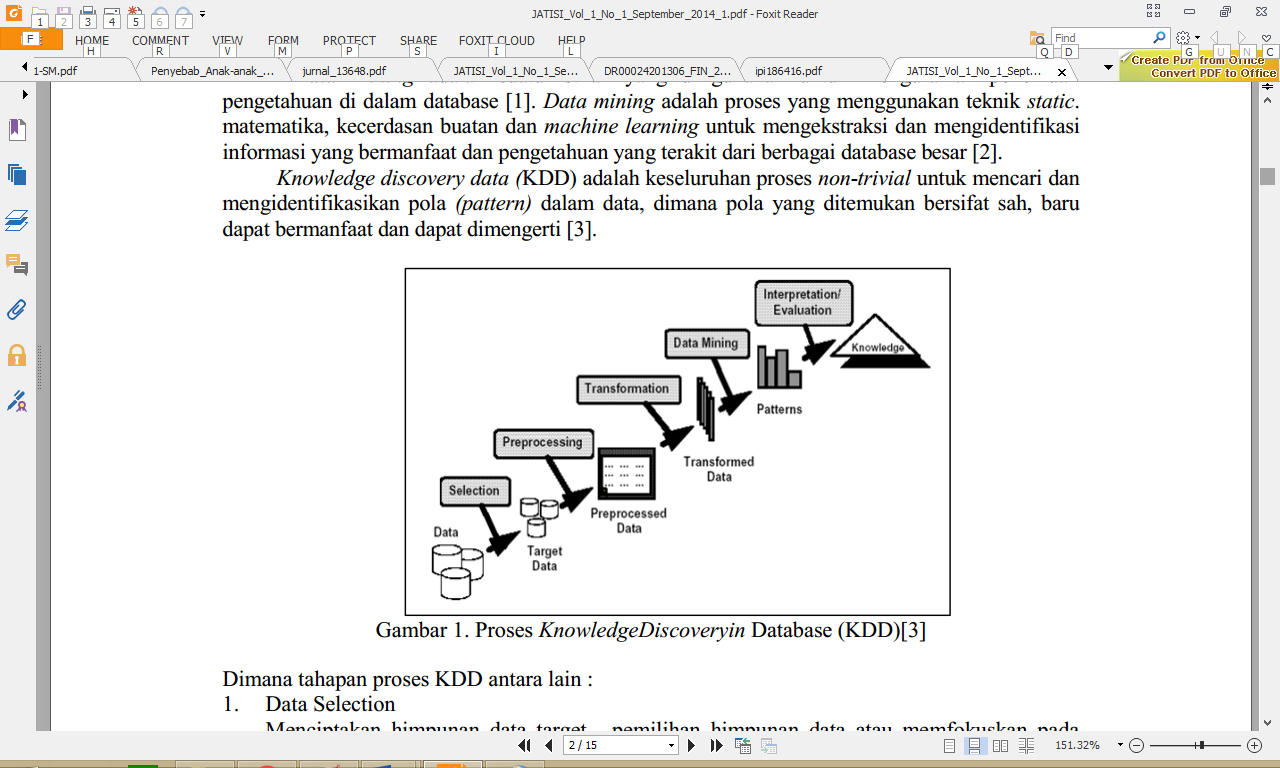
Klasifikasi Usaha Kecil dan Menengah (UMKM) di Provinsi Gorontalo merupakan solusi dari keberadaan industri kecil yang ada. Klasifikasi ini bertujuan untuk membantu mempermudah Pegawai Dinas Koperasi UKM Perindustrian dan Perdagangan di Provinsi Gorontalo

Dalam mengelola data Kecil dan Menengah (IKM) serta memberikan tempat bagi pemilik UKM untuk ikut berpartisipasi dalam persaingan industri dan memberikan informasi kepada masyarakat tentang UKM apa saja yang ada di Provinsi Gorontalo secara uptodate dalam Sistem. Dari sistem ini menghasilkan tampilan informasi dari masing-masing UKM secara detail. Sistem ini merupakan pengembangan konsep pendataan Industri Kecil dan Menengah (UKM) yang ada di Dinas Koperasi Perindustrian dan Perdagangan di Provinsi Gorontalo.

**2.2.2 Data Mining**

Data mining adalah serangkaian proses untuk menggali nilai tambah dari suatu kumpulan data berupa pengetahuan yang selama ini Sedangkan menurut Olha Musa, Suharto(2015)[4], “Data mining adalah proses menambang (mining) pengetahuan dari sekumpulan data yang sangat besar”. Data mining merupakan suatu langkah dalam *knowledge discoveryin database* (KDD).

*Knowledge discovery data (*KDD) adalah keseluruhan proses *non*-*trivial* untuk mencari dan mengidentifikasikan pola *(pattern)* dalam data, dimana pola yang ditemukan bersifat sah, baru dapat bermanfaat dan dapat dimengerti (Edmira Rivani(2010)[5].



*Sumber :* Edmira Rivani(2010)[5]

**Gambar 2.1** Proses *Knowledge Discoveryin Database (KDD)*

Dewasa ini *data mining* berkembang digunakan untuk menyelesaikan masalah menyangkut pendidikan. *Data mining* digunkana untuk menelusuri data yang ada untuk membangun sebuah model, kemudian menggunakan model tersebut agar dapat mengenali pola data yang lain yang tidak berada dalam basis data yang disimpan. Salah satu teknik *data mining* adalah teknik klasifikasi. Teknik klasifikasi adalah teknik pembelajaran untuk prediksi suatu nilai dari target *variabel* kategori

Kegunaan data mining adalah untuk mengklasifikasikan pola yang harus ditemukan dalam data mining. Secara umum, data mining dapat diklasifikasikan dalam dua kategori yaitu deskriptif dan prediktif. [6]

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 time sequencediscovery*)
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.

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.



Sumber : (] Billi Rizky Jurista)[6]

**Gambar 2.2** *Data mining* sebagai pertemuan dari banyak disiplin ilmu

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 *Queri*.

## **2.2.3 Clustering**

Pada dasarnya clustering merupakan suatu metode untuk mencari dan mengelompokkan data yang memiliki kemiripan karakteriktik *(similarity)* antara satu data dengan data yang lain. *Clustering* merupakan salah satu metode *data mining* yang bersifat tanpa arahan *(unsupervised)*, maksudnya metode ini diterapkan tanpa adanya latihan *(taining)* dan tanpa ada guru *(teacher)* serta tidak memerlukan target output. Dalam data mining ada dua jenis metode clustering yang digunakan dalam pengelompokan data, yaitu *hierarchical clustering* dan *non-hierarchical clustering*[6]

Analisis Pengelompokan / *Clustering* merupakan proses membagi data dalam suatu himpunan ke dalam beberapa kelompok yang kesamaan datanya dalam suatu kelompok lebih besar daripada kesamaan data tersebut dengan data dalam kelompok lain [7]

Potensi clustering adalah dapat digunakan untuk mengetahui struktur dalam data yang dapat dipakai lebih lanjut dalam berbagai aplikasi secara luas seperti klasifikasi, pengolahan gambar, dan pengenalan pola

*Hierarchical clustering* adalah suatu metode pengelompokan data yang dimulai dengan mengelompokkan dua atau lebih objek yang memiliki kesamaan paling dekat. Kemudian proses diteruskan ke objek lain yang memiliki kedekatan kedua. Demikian seterusnya sehingga cluster akan membentuk semacam pohon dimana ada hierarki (tingkatan) yang jelas antar objek, dari yang paling mirip sampai yang paling tidak mirip. Secara logika semua objek pada akhirnya hanya akan membentuk sebuah *cluster*. Dendogram biasanya digunakan untuk membantu memperjelas proses hierarki tersebut[

Berbeda dengan metode *hierarchical clustering*, metode *non-hierarchical*  
*clustering* justru dimulai dengan menentukan terlebih dahulu jumlah cluster yang diinginkan (dua cluster, tiga cluster, atau lain sebagainya). Setelah jumlah cluster diketahui, baru proses cluster dilakukan tanpa mengikuti proses hierarki. Metode ini biasa disebut dengan *K-Means Clustering*[8]

## **2.2.4 Algoritma K-Means.**

K-Means merupakan salah satu metode pengelompokan data nonhierarki (sekatan) yang berusaha mempartisi data yang ada ke dalam bentuk dua atau lebih kelompok. Metode ini mempartisi data ke dalam kelompok sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan kedalam kelompok yang lain.

Adapun tujuan pengelompokkan data ini adalah untuk meminimalkan fungsi objektif yang diatur dalam proses pengelompokan, yang pada umumnya berusaha meminimalkan variasi di dalam suatu kelompok dan memaksimalkan variasi antar kelompok.[9]

Menurut Santosa (2007)[7], langkah-langkah melakukan clustering dengan metode *K-Means* adalah sebagai berikut:

1. Pilih jumlah *cluster k*.
2. Inisialisasi *k* pusat *cluster* ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat cluster diberiduberi nilai awal dengan angka-angka random,
3. Alokasikan semua data/ objek ke cluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke *cluster* tertentu ditentukan jarak antara data dengan pusat *cluster*. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat *cluster*. Jarak paling antara satu data dengan satu *cluster* tertentu akan menentukan suatu data masuk dalam *cluster* mana. Untuk menghiutng jarak semua data ke setiap tiitk pusat cluster dapat menggunakan teori jarak Euclidean yang dirumuskan sebagai berikut:

(2.1)

......

dimana:  
*D (i,j)* = Jarak data ke *i* ke pusat cluster *j*

*Xk i*= Data ke *i* pada atribut data ke *k*

*Xk i*= Titik pusat ke *j* pada atribut ke *k*

1. Hitung kembali pusat *cluster* dengan keanggotaan *cluster* yang sekarang. Pusat *cluster* adalah rata-rata dari semua data/ objek dalam *cluster* tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.
2. Tugaskan lagi setiap objek memakai pusat *cluster* yang baru. Jika pusat cluster tidak berubah lagi maka proses *clustering* selesai. Atau, kembali ke langkah nomor 3 sampai pusat *cluster* tidak berubah lagi

## **2.2.5 Penerapan Metode K-Means Untuk Klustering**

Terdapat data sampel berupa 10 data sebagaimana yang ada pada Tabel 2.1. Data tersebut memiliki dua dimensi (fitur x dan fitur y) agar mudah di visualisasikan dalam koordinat kartesius.

**Tabel 2.2**  Data Set Sintetik Numerik 2 Dimensi

|  |  |  |
| --- | --- | --- |
| Data Ke-i | Fitur x | Fitur y |
| 1 | 1 | 1 |
| 2 | 4 | 1 |
| 3 | 6 | 1 |
| 4 | 1 | 2 |
| 5 | 2 | 3 |
| 6 | 5 | 3 |
| 7 | 2 | 5 |
| 8 | 3 | 5 |
| 9 | 2 | 6 |
| 10 | 3 | 8 |

|  |
| --- |
|  |

Berdasarkan data set tersebut, dilakukan proses pengelompokan menjadi 3 cluster (k = 3). berdasarkan k=3, maka ditentukan titik centroid sebanyak k berdasarkan titik-titik tertentu data set. Dapat dilakukan secara random ataupun langsung ditentukan. Maka dipilih data ke-2, 4 dan 6 sebagai centroid awal. Pengukuran jarak pada setiap data terhadap titik centroid dilakukan menggunakan perhitungan jarak Euclidean.

**Tabel 2.3.** Penentuan Centroid

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | Centroid | x | y | | 1 | 4 | 1 | | 2 | 1 | 2 | | 3 | 5 | 3 | |
|  |

Setelah nilai K dan centroid terinisialisasi, maka perlu ditentukan pula nilai fungsi objektif sebagai ambang batas iterasi yaitu sebesar 0.1. Sebagai permulaan, misalnya ditentukan nilai fungsi objektif sebesar 1000. Jadi jika pada iterasi tertentu nilai fungsi objektif didapatkan sudah berada di bawah ambang batas 0.1, maka iterasi akan berhenti.

Tahap berikutnya adalah melakukan iterasi untuk memperoleh pembaharuan titik centroid serta mendapatkan perubahan nilai fungsi objektif. Sepanjang nilai fungsi objektif masih berada di atas ambang batas yang ditetapkan, maka proses iterasi akan tetap beranjut.

Proses perhitungan dimulai dari **iterasi 1**. Awalnya, dihitung jarak dari setiap data (1-10) terhadap semua centroid yang ada. Dari hasil perhitungan jarak antara setiap data terhadap semua centroid, didapatkan nilai jarak terkecil terhadap satu centroid, maka centroid tersebut disebut sebagai centroid terdekat, dan data akan berafiliasi menjadi cluster dari centroid terdekat. Berikut ini contoh menghitung jarak antara data ke-1 dengan centroid 1,2 dan 3 menggunakan metode euclidean. (data dan centroid sudah saya tunjukkan pada bagian 1).

|  |
| --- |
| <https://3.bp.blogspot.com/-EELs1moot7g/WFCXki_cI3I/AAAAAAAAENQ/bWR8q3unnH0tk-8q-edOLpNDsRnxf2KGwCLcB/s1600/boleh.PNG> |

Dari perhitungan diatas, dapat diketahui bahwa data ke-1 lebih dekat jaraknya terhadap centroid ke-2. Oleh karena itu data ke-1 mengikuti cluster ke-2. Berdasarkan contoh diatas, ditentukan jarak terdekat pada semua data terhadap centroid. Jarak terdekat menentukan suatu data akan masuk ke cluster 1,2 atau 3.

**Tabel 2.4** Penentuan cluster pada semua data

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **Data ke-i** | **Jarak ke Centroid** | | | **Terdekat** | **Cluster** | | **1** | **2** | **3** | | 1 | 3 | 1 | 4.472136 | 1 | 2 | | 2 | 0 | 3.162278 | 2.236068 | 0 | 1 | | 3 | 2 | 5.09902 | 2.236068 | 2 | 1 | | 4 | 3.162278 | 0 | 4.123106 | 0 | 2 | | 5 | 2.828427 | 1.414214 | 3 | 1.41421356 | 2 | | 6 | 2.236068 | 4.123106 | 0 | 0 | 3 | | 7 | 4.472136 | 3.162278 | 3.605551 | 3.16227766 | 2 | | 8 | 4.123106 | 3.605551 | 2.828427 | 2.82842712 | 3 | | 9 | 5.385165 | 4.123106 | 4.242641 | 4.12310563 | 2 | | 10 | 7.071068 | 6.324555 | 5.385165 | 5.38516481 | 3 | |

Setelah dipastikan bahwa data 1 hingga data 10 sudah masuk ke salah satu cluster. Selanjutnya, ditentukan centroid baru berdasarkan data yang tergabung pada setiap clusternya. Dimulai dari cluster 1, terdapat dua data yang tergabung didalamnya (lihat gambar 2). Centroid baru didapatkan dari rata-rata yang diperoleh yaitu 5 dan 1.

**Tabel 2.5** Pentuan centroid baru untuk cluster 1

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | Data anggota | Fitur x | Fitur y | | 2 | 4 | 1 | | 3 | 6 | 1 | | Jumlah | 10 | 2 | | Rata-Rata | 5 | 1 | |

Kemudian untuk cluster ke-2, ada 5 data yang tergabung didalamnya (lihat gambar 2). Centroid baru didapatkan dari rata-rata yang diperoleh yaitu 1.6 dan 3.4.

**Tabel 2.6** Penentuan centroid baru untuk cluster 2

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | Data anggota | fitur x | Fitur y | | 1 | 1 | 1 | | 4 | 1 | 2 | | 5 | 2 | 3 | | 7 | 2 | 5 | | 9 | 2 | 6 | | jumlah | 8 | 17 | | Rata-Rata | 1.6 | 3.4 | |

Yang terakhir adalah untuk cluster ke-3, ada 3 data yang tergabung didalamnya (lihat gambar 2). Centroid baru didapatkan dari rata-rata yang diperoleh yaitu 3.66667 dan 5.33333.

**Tabel 2.7** Penentuan centroid baru untuk cluster 3

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | Data anggota | Fitur x | Fitur y | | 6 | 5 | 3 | | 8 | 3 | 5 | | 10 | 3 | 8 | | Jumlah | 11 | 16 | | Rata-Rata | 366.667 | 533.333 | |
|  |

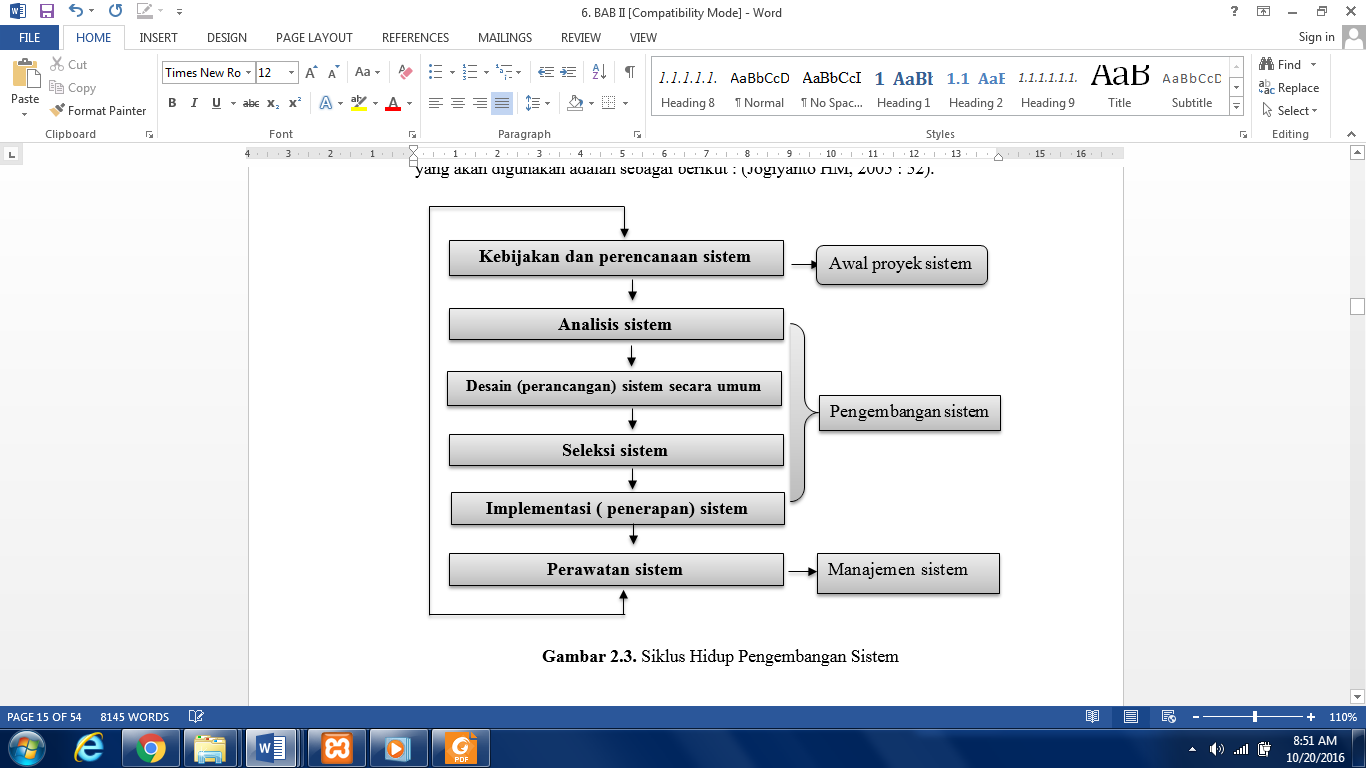
Rata-rata yang didapatkan dari ketiga cluster tersebut merupakan centroid baru yang didapat. Centroid baru digunakan untuk perhitungan pada iterasi ke-2.

## **2.2.6 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 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.

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



Sumber : Muchdarsyah Sinugun [10]

**Gambar 2.3** Siklus hidup pengembangan sistem

Ide dari *systems life cycle* adalah sederhana dan 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.

## **Analisa 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.

M. Rosyid Ridlo[7] 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.

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.

1. 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 Yuniasih Dan Suwanto (2015),[11]

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 inernal).

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

## **Desain Sistem**

Setelah tahap analisis sistem selesai dilakukan, maka analisis sistem telah mendapat gambaran dengan jelas apa yang harus dikerjakan. Tiba waktunya sekarang bagi analisis sistem untuk memikirkan bagaiamana membentuk sistem tersebut. Tahap ini disebut dengan desain sistem (*system design* ).

M. Rosyid Ridlo[7] mengungkapkan :” *System design* adalah spesifikasi atau instruksi solusi yang teknis dan berbasis komputer untuk persyaratan bisnis yang diidentifikasikan dalam analisis sistem.”

Desain sistem adalah spesifikasi atau intruksi solusi yang teknis dan berbasis komputer untuk persyaratan bisnis yang diidentifkasikaan dalam analisis sistem.

Driver teknologi sekarang ( dan dimasa depan ) paling berimpak pada proses dan keputusan desain sistem. Banyak organisasi mengidentifikasikan arsitektur teknologi informasi umum yang didasarkan pada driver-driver teknologi ini.

Tahap desain sistem mempunyai dua tujuan utama, yaitu :

1. Untuk memenuhi kebutuhan kepada pemakai system.
2. Untuk memberikan gambaran yang jelas dan rancang bangun yang lengkap kepada pemogram computer dan ahli-ahli teknik lainnya.

Perancangan sistem adalah suatu keinginan membuat desain teknis yang berdasarkan evaluasi yang dilakukan pada kegiatan analisis. Perancangan disini dimaksudkan sebagai proses pemahaman dan perancangan suatu sistem berbasis computer yang akan menghasilkan komputerisasi.

Dengan demikian, suatu kegiatan perancangan sistem bertujuan untuk menghasilkan suatu sistem komputerisasi. Komputerisasi adalah suatu kegiatan atau sistem pengolahan data dengan menggunakan komputersebagai alat bantu. Perancangan sistem dilakukan setelah tahap analisis sistem selesai dilaksanakan yang kemudian akan menghasilkan output berupa kebutuhan yang akan dijadikan dasar untuk merancang sistem tersebut.

Perancangan sistem terbagi dua, yaitu :

1. Perancangan konseptual.

Perancangan konseptual sering kali disebut dengan perancangan logis. Pada perancangan ini, kebutuhan pemakai dan pemecahan masalah yang teridentifikasi selama tahap analisis sistem mulai dibuat untuk diimplementasikan. Ada tiga langkah penting yang dilakukan dalam perancangan konseptual, yaitu evaluasi alternatif rancangan, penyiapan spesifikasi rancangan, dan penyiapan laporan rancangan sistem secara konseptual.

Menurut Yuniasih Dan Suwanto (2015),[11]evaluasi yang dilakukan mengandung hal-hal berikut :

1. Bagaiamana alternatif-alternatif tersebut memenuhi sasaran sistem dan organisasi dengan baik ?
2. Bagaimana alternatif-alternatif tersebut memenuhi kebutuhan pemakai dengan baik ?
3. Apakah alternatif-alternatif tersebut layak secara ekonomi ?
4. Apa saja keuntungan dan masing- masing ?

Setelah alternatif rancangan dipilih, tahap selanjutnya adalah penyiapan spesikasi rancangan, yang mencakup elemen- elemen sebagai berikut :

1. Keluaran.

Rancangan laporan mencakup frekuensi laporan (harian, mingguan, dsb ), isi laporan , dan laporan cukup ditampilkan pada layar atau perlu dicetak.

1. Penyiapan data.

Dalam hal ini, semua data yang diperlukan untuk membentuk laporan ditentukan lebih detail,termasuk ukuran data dan letaknya dalam berkas.

1. Masukan.

Rancangan masukan meliputi data yang perlu dimasukan kedalam sistem.

1. Prosedur pemrosesan dan operasi.

Rancangan ini menjelaskan bagaimana data dimasukan diproses dan disimpan dalam rangka untuk menghasilkan laporan.

1. Perancangan fisik.

Pada perancangsn ini, rancangan yang masih bersifat konsep diterjemahkan dalam bentuk fisik sehingga terbentuk spesifikasi lengkap tentang modul sistem dan antarmuka antar modul, serta rancangan basis data secara fisik.

Beberapa hasil akhir setelah tahap perancangan fisik berakhir :

1. Rancangan keluaran.

Rancangan keluaran berupa bentuk laporan dan rancangan dokumen

1. Rancangan masukan.

Rancangan masukan berupa rancangan layar untuk pemasukan data.

1. Rancangan antarmuka pemakai dengan sistem.

Rancangan ini berupa rancangan interaksi antara pemakai dan sistem.Misalnya : berupa menu, ikon, dan lain-lain.

1. Rancangan *platform.*

Rancangan ini berupa rancangan yang menentukan *hardware*(perangkat keras) dan *software* (perangkat lunak) yang akan digunakan.

1. Rancangan ini berupa rancangan-rancangan berkas dalam basis data, termasuk penentuan kapasitas masing-masing.
2. Rancangan modul.

Rancangan ini berupa rancangan program yang dilengkapi dengan algortima (cara modul/program bekerja).

1. Rancangan control.

Rancangan ini berupa rancangan kontrol-kontrol yang dugunakan dalam sistem seperti validasi, otorisasi,audit data.

1. Dokumentasi.

Berupa hasil dokumentasi hingga tahap perancangan fisik.

1. Rencana pengujian.

Berupa rencana yang dipakai untuk menguji sistem.

1. Rencana konversi.

Berupa rencana untuk menerapkan sistem baru terhadap sistem lama.

Dalam perancangan sistem yang baik melalui tahapan-tahapan sebagai berikut:

1. Identifikasi masalah yaitu mengidentifikasi masalah yang ada secara rinci agar tidak timbul masalah lain selain masalah utama.
2. Menentukan input, proses dan output yang diinginkan yaitu menginginkan hasil dari perancangan sistem yang dibuat sesuai dengan prosedur.
3. Menentukan algoritma.
4. Mengimplementasikan dengan bahasa pemograman tertentu.
5. Desain sistem dapat dibagi dua bagian,yaitu desain sistem secara umum (general system design) dan desain sistem terinci (detailed system design).

## **Desain sistem secara umum**

Tujuan dari desain sistem secara umum adalah untuk memberikan gambaran secara umum kepada user tentang sistem yang baru,yang mana merupakan persiapan dari desain sistem secara rinci.Desain secara umum dilakukan oleh analisis sistem untuk mengidentifikasikankomponen-komponen sistem informasi yang akan didesain secara rinci oleh pemograman komputer dan ahli teknik lainya.

Pada tahap inikomponen-komponen sistem informasi di rancang untuk dikomunikasikan kepada user.Komponen sistem informasi yang didesain adalah model, output - input,database,teknologi dan kontrol.

## **Desain Sistem Terinci (*Detailed system design*)**

1. 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.

1. Desain Output dalam bentuk laporan

Desain ini dimaksudkan untuk menghasilkan output dalam bentuk laporan di media kertas.Bentuk laporan yang paling banyak digunakan adalah dalam bentuk tabel dan berbentuk grafik atau bagan.

1. Desain Output dalam bentuk dialog layar terminal.

Desain ini merupakan rancangan bangun dari percakapan antara pemakai sistem (user) dengan komputer.Percakapan ini dapat terdiri dari proses memasukkan data ke sistem,menampilkan output informasi kepada user,atau keduanya.

Beberapa strategi dalam membuat layar dialog terminal:

1. Dialog pertanyaan/jawaban.
2. Menu.

Menu banyak digunakan karena merupakan jalur pemakai yang mudah dipahami dan mudah digunakan.Menu berisi beberapa alternatif atau option atau option atau pilihan yang di sajikan kepada user.Pilihan menu akan lebih baik bila dikelompokan fungsinya.

1. Desain input Terinci.

Masukan merupakan awal dimulainya proses informasi.Bahan mentah dari informasi adalah data yang terjadi dari transaksi-transaksi yang dilakukan loleh organisasi.Data hasil dari transaksi merupakan masukan untuk sistem informasi.Hasil dari sistem informasi tidak lepas dari data yang dimasukkan. Desain *Inpu*t terinci dimulai dari desain dokumen dasar tidak didesain desain dengan baik, kemungkinan input yang tercatat dapat salah bahkan kurang.

Fungsi dokumen dasar dalam penanganan arus data:

1. Dapat menunjukan macam dari data yang harus dikumpulkan dan ditangkap.
2. Data dapat dicatat dengan jelas,konsisten dan akurat.
3. Dapat mendorong lengkapnya data,disebabkan data yang dibutuhkan disebutkan satu persatu di dalam dokumen dasarnya.
4. Desain Database Terinci.

Basis data (database) merupakan kumpulan dari data yang saling berhubungan satu dengan yang lainnya,tersimpan di simpan luar komputer dan digunakan perangkat lunak tertentu untuk manipulasinya.Databse merupakan salah satu komponen yang penting di sistem informasi,karena berfungsi sebagian penyedia informasi bagi para pemakainya.penerapan database dalam sistem informasi disebut database system.

1. Desain Teknologi.

Tahap desain terbagi atas dua yaitu desain teknologi secara umum di rinci. Pada tahap ini kita menentukan teknologi yang akan di pergunakan dalam menerima input, menjalankan model, menyimpan dan mengakses data,menghasilkan dan mengirimkan keluaran dan membantu pengendalian dari sistem secara keseluruhan.

1. Tahap Desain

Tahap desain terbagi menjadi dua,yaitu desain model secara umum dan terinci. Tahap desain model secara umum berupa desain sistem secara fisik dan logika. Desain fisik dapat di gambarkan dengan bagian alir sistem bagian alir dokumen, dan desain secara logika digambarkan dengan diagram dengan arus data(DAD), pada tahap desain model terinci,modelakan didefinisikan secara terinci. urut-urutan langkah proses ini diwakili oleh suatu program komputer.

**Tabel 2.8** Bagan Alir Sistem

|  |  |  |  |
| --- | --- | --- | --- |
| **NO** | **NAMA SIMBOL** | **SIMBOL** | **KETERANGAN** |
| 1. | Simbol Dokumen |  | Menunjukkan dokumen *input* dan *output* baik itu proses manual, mekanik, atau komputer |
| 2. | Simbol Kegiatan Manual |  | Menunjukan pekerjaan manual |
| 3. | Simbol Simpanan Offline |  | Menunjukkan file non-komputer yang diarsip urut angka (*numerical*), huruf (*alphabetical*), atau tanggal (*chronological*) |
| 4. | Simbol Kartu Plong |  | Menunjukkan *input* dan *output* yang menggunakan kartu plong (*punched card*). |
| 5. | Simbol Proses |  | Menunjukkankegiatan proses dari operasi program komputer |
| 6. | Simbol Operasi Luar |  | Menunjukkan operasi yang dilakukan di luar proses operasi komputer |
| 7. | Simbol Hard Disk |  | Menunjukkan *input* dan *output*  menggunakan *harddisk* |
| 10. | Simbol Keyboard |  | Menunjukkan *input* yang menggunakan *on-line keyboard* |
| 11. | Simbol Display |  | Menunjukkan *output* yang ditampilkan di monitor. |
| 12. | Simbol Hubungan Komunikasi |  | Menunjukkan proses transmisi data melalui *channel* komunikasi. |
| 13. | Simbol Garis Alir |  | Menunjukkan arus dari proses |
| 14. | Simbol Penjelasan |  | Menunjukkan penjelasan dari suatu proses |
| 15. | Simbol Penghubung |  | Menunjukkan penghubung ke halaman yang masih sama atau ke halaman yang lain |

Sumber : Muchdarsyah Sinugun (2015),[10]

Untuk mempermudah penggambaran suatu sistem yang ada atau sistem yang baru yang akan dikembangkan secara logika tanpa memperhatikan lingkungan fisik di mana data tersebut mengalir atau lingkungan fisik di mana data tersebut akan disimpan, maka digunakan Diagram Arus Data (DAD) atau *Data Flow Diagram* (DFD). Dalam menggambarkan sistem perlu dilakukan pembentukan simbol, berikut ini simbol-simbol yang sering digunakan dalam DAD :

1. *External entity* (kesatuan luar) atau *boundary* (batas sistem).

Setiap sistem pasti mempunyai batas sistem *(boundary)* yang memisahkan suatu sistem dengan lingkungan luarnya. Sistem akan menerima *input* dan menghasilkan *output* kepada lingkungan luarnya. Kesatuan luar *(external entity)* merupakan kesatuan di lingkungan luar sistem yang dapat berupa

orang, organisasi atau sistem lain yang berada di lingkungan luarnya yang akan memberikan *input* serta menerima *output* dari sistem. [10]

**Gambar 2.4** Notasi kesatuan luar di DAD

1. *Data flow* (arus data).

Arus data ini menunjukkan arus atau aliran data yang dapat berupa masukkan untuk sistem atau hasil dari proses sistem. . [10]

**Gambar 2.5** Nama Arus Data di DAD

1. *Process* (proses).

Suatu proses adalah kegiatan atau kerja yang dilakukan orang, mesin atau komputer dari hasil suatu arus data yang masuk ke dalam proses untuk dihasilkan arus data yang akan keluar dari proses.[10]



**Gambar 2.6** Contoh Notasi Proses

1. *Data store* (simpanan data).

Simpanan data pada DFD dapat disimbolkan dengan sepasang garis horisontal paralel yang tertutup disalah satu ujungnya.[12]

Media Nama Data store

**Gambar 2.7** Notasi Simpanan Data di DAD

## **Implementasi Sistem**

Tahapan implementasi merupakan tahap dimana dilakukan transformasi/ penerjemahan dari bahasa modeling ke suatu bahasa pemrograman. hal ini merupakan tugas dari pemprogram, pada pengembangan sistem/perangkat lunak berorientasi objek penerjemahan dari setiap diagram-diagram DAD yang telah di rancang pada tahap perancangan harus diterjemahkan ke dalam bahasa pemrograman sama persis dengan diagram-diagram yang ada guna menghindari terjadinya perubahan fungsi/tujuan dari pengembangan sistem/perangkat lunak.

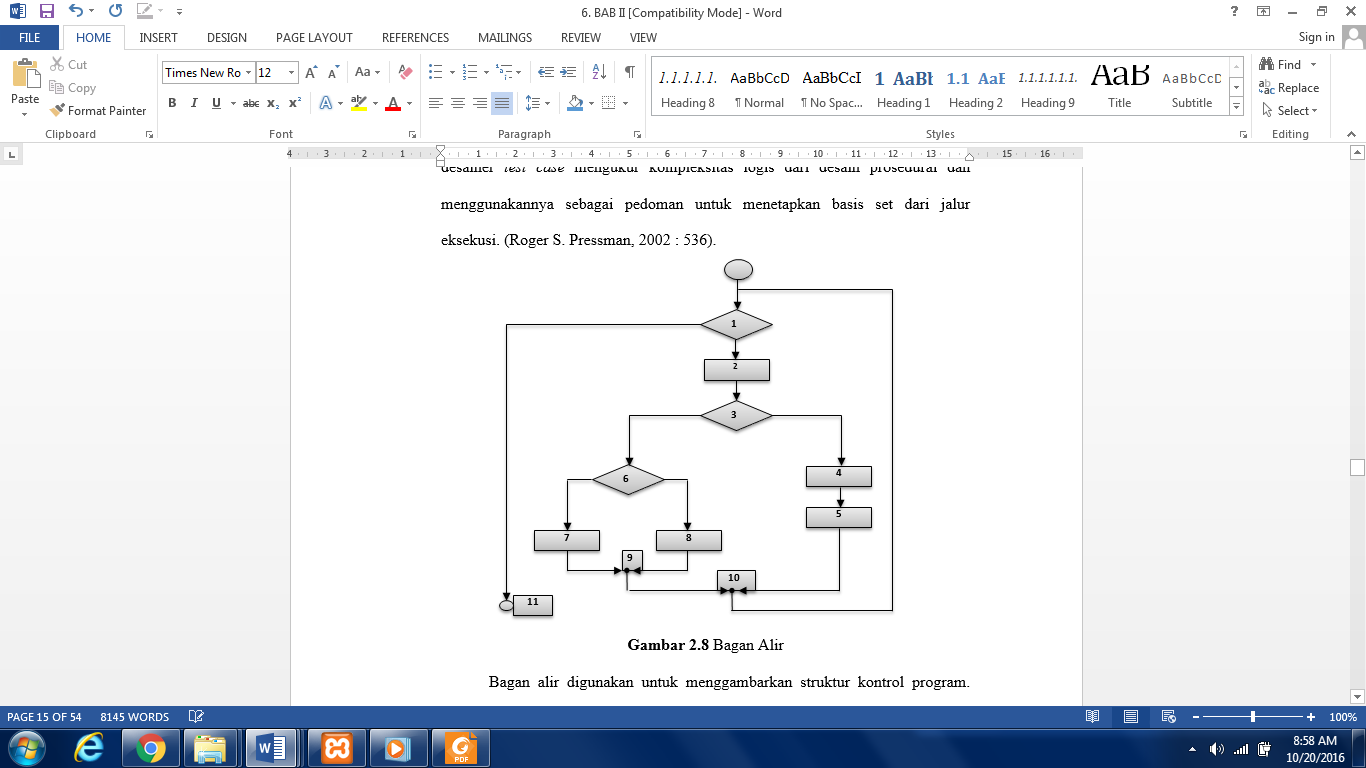
## **Teknik Pengujian Sistem**

## **2.4.1 White Box Testing**

White box testing adalah metode desain test case yang menggunakanstruktur control desain procedural untuk mendapatkan test case. Denganmenggunakan metode white box analisis sistem akan dapat memperoleh *test case*yang meliputi :

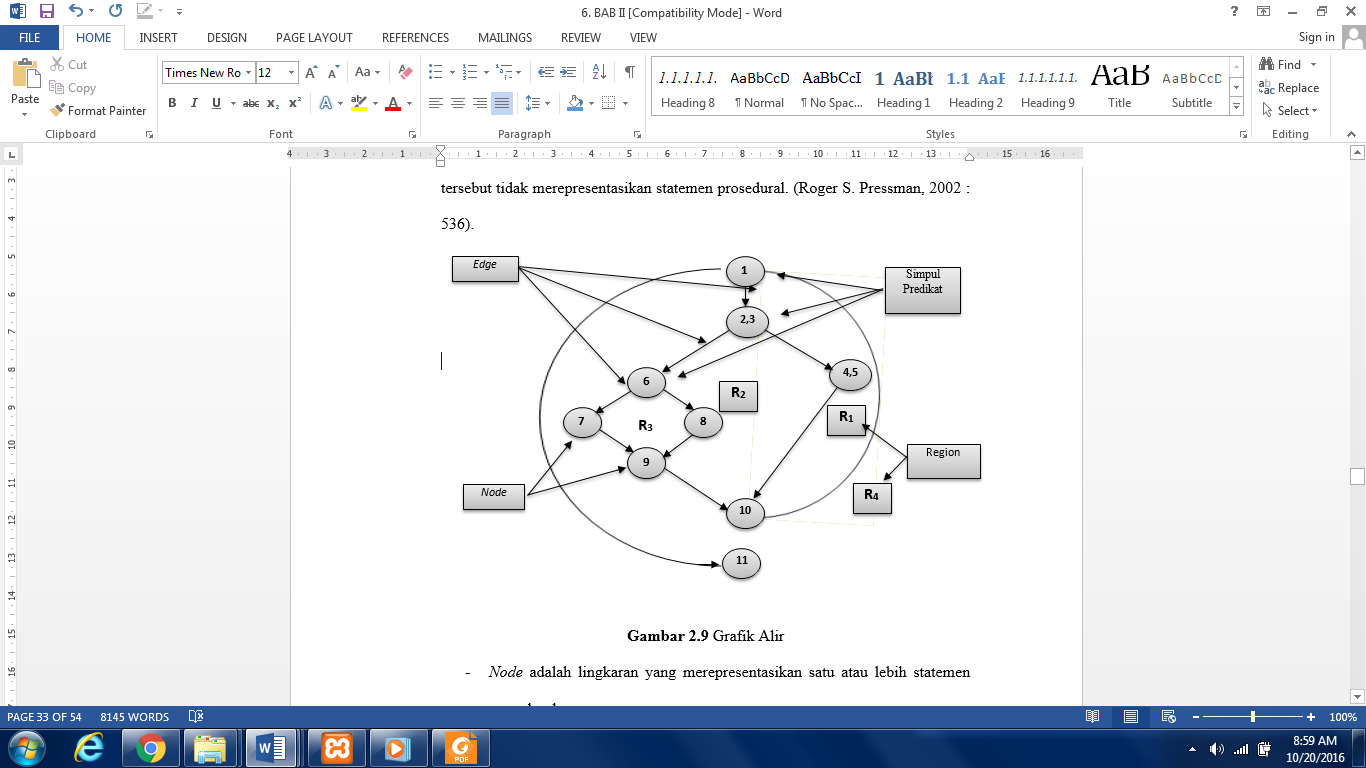
1. Menjamin seluruh independent path di dalam modul yang di kerjakan sekurang kurangnya sekali.
2. Mengerjakan seluruh keputusan *logical*.
3. Mengerjakan seluruh loop sesuai dengan batasannya.
4. Mengerjakan seruruh struktur data internal untuk menjamin validitas.

Untuk melakukan pengujian *Test Case* terlebih dahulu dilakukan penerjemahan *Flowchart* kedalam notasi *Flowgraph*.



Sumber (Roger S. Pressman)[12]

Gambar 2.8 Bagan Alir



Sumber (Roger S. Pressman, 2002)[12]

Gambar 2.9 Grafik Alir

Ada beberapa istilah saat pembuatan Flowgraph, yaitu :

1. *Node* yaitu lingkaran pada Flowgraph yang menggambarkan satu atau lebih perintah procedural.
2. *Edge* yaitu tanda panah yang menggambarkan aliran kontrol dan setiap Node harus mempunyai tujuan Node.
3. *Region* yaitu daerah yang dibatasi oleh node dan edge dan untuk menghitung region daerah di luar flowgraph juga harus dihitung.
4. *Predicate node* yaitu kondisi yang terdapat pada node dan mempunyai karasteristik dua atau lebih edge lainnya.
5. *Cyclomatic complexity* yaitu metrik perangkat lunak yang menyediakan ukuran kuantitatif dari kekompleksan logical program dan dapat digunakan untuk mencari jumlah path dalam satu flowgraph.
6. *Independen path* yaitu jalur yang melintasi atau melalui program dimana sekurang-kurangnya terdapat proses perintah yang baru atau kondisi yang baru.

Rumus-rumus menghitung jumlah independen path dalam satu *flowgraph*yaitu :

1. Jumlah region flowgraph mempunyai hubungan dengan cyclomatic comlexity (CC)
2. V (G) untuk flowgraph dapat dihitung dengan rumus :

a. V (G) = E – N + 2 ………………………………………..Persamaan (2.2)

Dimana :

E = Jumlah edge pada flowgraph.

N = Jumlah node pada flowgraph.

b. V (G) = P + 1 .…………………………………………....Persamaan (2.3)

Dimana :

P = Jumlah predicate node pada flowgraph.

Dalam pelaksanaanya teknik pelaksanaanya pengujian White box inimempunyai tiga langkah yaitu:

1. Menggambar *flowgraph* yang ditransfer dari *flowchart*.
2. Menghitung *cyclomatic complexity* untuk *flowgraph* yang telah dibuat
3. Menentukan jalur pengujian dari *flowgraph* yang berjumlah sesuai dengan *cyclomatic complexity* yang telah ditentukan.

## **Black Box Testing**

Menurut Ladjamudin (2006)[13] pengujian black box berfokus padapersyaratan fungsional perangkat lunak.Konsep pengujian black box (kotak hitam) digunakan untukmerepresentasikan sistem yang cara kerja didalamnya tidak tersedia untukdiinspeksi karena item-item yang diuji dianggap gelap karena logik nya tidakdiketahui, yang diketahui hanya apa yang masuk dan apa yang keluar dari kotak hitam.



Gambar 2.10 Notasi Graph Bases Testing

## **Tools Pendukung**

Perangkat lunak pendukung yang digunakan penulis dalam membangun sistem ini yaitu PHP dan MySQL, seperti pada tabel di bawah ini :

**Tabel 2.9**  Perangkat Lunak Pendukung

|  |  |  |
| --- | --- | --- |
| **NO** | **TOOLS** | **KEGUNAAN** |
| 1 | PHP | Adalah Sebuah bahasa *scripting* yang terpasang pada HTML. Yang bertujun untuk memungkinkan perancang web menulis halaman web dinamik dengan cepat. |
| 2 | MySQL | Salah satu pengolah database yang menggunakan SQL (*Strukture Query Language*) sebagai bahan dasar untuk mengakses databasenya. Yang memiliki keuntungan seperti *open source* dan memiliki kemampuan menampung kapasitas yang besar. |

## **Kerangka Pemikira****n**

**MASALAH**

1. Pihak Dinas Perindustrian & Perdangan Masih sulit mengelompokan UMKM dalam pemberian bantuan
2. Belum adanya sistem yang digunakan untuk pemberian bantuan pada usaha Mikro, Kecil, Menengah

**Identifikasi Model Algoritma** *K-Means Clustering*



Observasi

Pengumpulan Data Set

Parameter *K-Means Clustering*

d(x,y) = ||x-y||2 = 2

**System Development**

Analisa Sistem

Desain Sistem

- diagram konteks(DFD)

- diagram berjenjang(DFD)

- Diagram Arus data level 0,

- Kamus data

- Desain Output(Visio)

- Desain Input(Visio)

- Desain basis data(struktur data

1. Untuk membangun metode dalam mengelompokan Usaha Mikro Kecil Menengah.
2. Mengimplementasikan Algoritma K-Means dalam mengelompokan pemberian bantuan pada Usaha Mikro Kecil Menengah

**TUJUAN**

Pengujian Sistem

* Program (white box testing)
* Interface (black box testing)

Konstruksi Sistem

- Programing PHP

* Database MySQL

**Gambar 2.11**  Bagan Kerangka berpikir